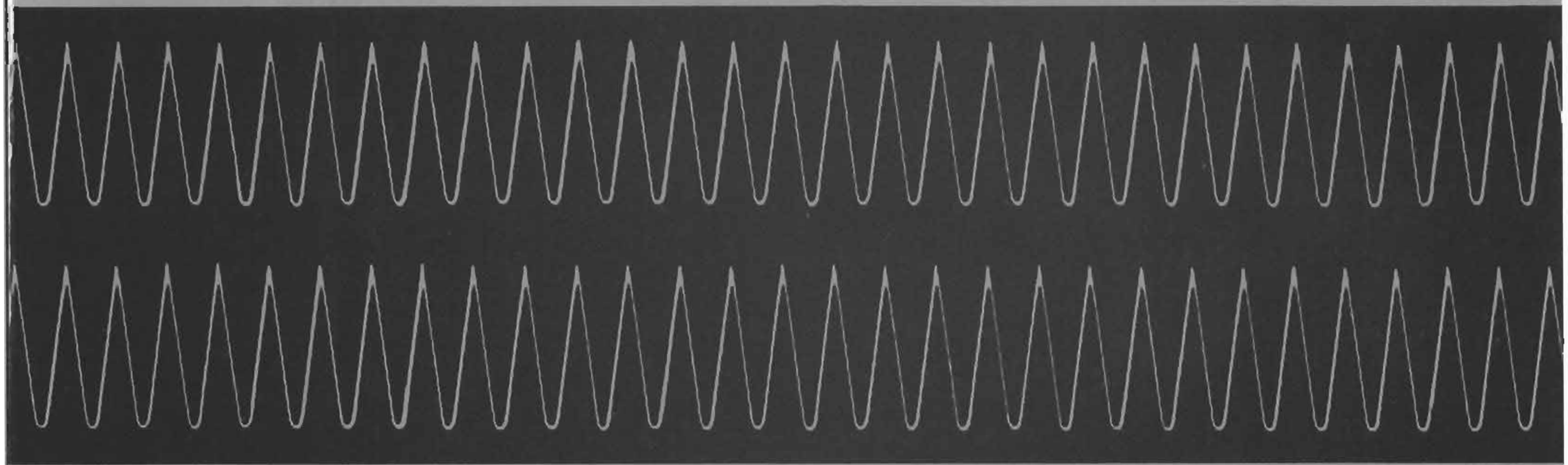
Proceedings of the Conference on the Planning and Production of Speech in Normal and Hearing-Impaired Individuals:

A Seminar in Honor of S. Richard Silverman



ASHA REPORTS

NUMBER 15 A PUBLICATION OF THE AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION

PROCEEDINGS OF THE CONFERENCE ON THE PLANNING AND PRODUCTION OF SPEECH IN NORMAL AND HEARING-IMPAIRED INDIVIDUALS:

A SEMINAR IN HONOR OF S. RICHARD SILVERMAN

PROCEEDINGS OF THE CONFERENCE ON THE PLANNING AND PRODUCTION OF SPEECH IN NORMAL AND HEARING-IMPAIRED INDIVIDUALS:

A SEMINAR IN HONOR OF S. RICHARD SILVERMAN

Edited by

JUDITH L. LAUTER, PH.D.

with the editorial collaboration of:

Craig Bowe, Ph.D.
Donald R. Calvert, Ph.D.
Kenneth Grant, M.A.
Roanne Karzon, Ph.D.
James D. Miller, Ph.D.
Janet Weisenberger, Ph.D.

ASHA Reports 15 (ISSN 0569-8553)
AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION
Rockville, Maryland
October 1985

AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION

President
Chair, Legislative Council
Chair, Executive Board

Richard M. Flower, Ph.D.
University of California—San Francisco

EXECUTIVE BOARD

President-Elect
Betty Jane Philips, Ed.D.
Boys Town National Institute
Omaha, Nebraska

Past President
David E. Yoder, Ph.D.
University of Wisconsin—Madison

Vice President for Administration
Sandra C. Holley, Ph.D.
Southern Connecticut State University
New Haven

Vice President for Clinical Affairs
Gilbert R. Herer, Ph.D.
Children's Hospital National Medical Center
Washington, DC

Vice President for Education and Scientific Affairs
Steffi B. Resnick, Ph.D.
J. F. Kennedy Institute
Baltimore, Maryland

Vice President for Planning
Carol P. Leslie, Ph.D.
Cincinnati Speech and Hearing Center
Cincinnati, Ohio

Vice President for Professional and Governmental Affairs
Nancy G. Becker, Ph.D.
Cook County Hospital
Chicago, Illinois

Vice President for Standards and Ethics
Jane R. Madell, Ph.D.
New York League for the Hard of Hearing
New York, New York

Executive Director
Frederick T. Spahr, Ph.D.

PUBLICATIONS BOARD

Robert C. Bilger, Ph.D., **Chair**
Edward G. Conture, Ph.D.
Tanya M. Gallagher, Ph.D.
Katherine Harris, Ph.D.
Thomas J. Hixon, Ph.D.
John H. Mills, Ph.D.
David Prins, Ph.D.
Gerald M. Siegel, Ph.D.
Frederick T. Spahr, Ph.D. (*ex officio*)

ASHA REPORTS NUMBER 15

Editor
Judith L. Lauter, Ph.D.

ASHA Reports No. 15 Editor (1979-1983)
Theodore J. Glattke, Ph.D.

Series Editor
Arnold M. Small, Ph.D., 1984–

Business Manager
Frederick T. Spahr, Ph.D.

Publications/Marketing Division
Alfred K. Kawana, Director

Production Editor
Sharyn Schlesinger

ACKNOWLEDGMENTS

We gratefully acknowledge contributors to the S. Richard Silverman Fund of Central Institute for the Deaf who made the conference possible and helped support this publication.

PREFACE

Although oral education of hearing-impaired children and adults is based on a tradition that is centuries old, the science of speech production in hearing or hearing-impaired individuals is a young one. In the last few years interest in the details of speech planning and production has blossomed, drawing on new technologies such as electromyography, and new strategies such as analysis of speech errors.

The Silverman Seminar on the Planning and Production of Speech was organized in October 1983 to bring together two groups of experts: (a) those with experience in the day-to-day exigencies of teaching hearing-impaired children to speak, and (b) those whose research is directed to understanding the processes of speech planning and production in speakers with normal hearing. A combination of formal paper presentations, and working-group papers and discussions, was formulated to encourage the interaction of teachers and researchers.

The result is a collection of data, observations, questions, and answers that should be of interest to a wide range of readers. Those who are interested in how language is organized both as a system and within the human brain should find these proceedings of interest, as should teachers of the hearing-impaired, and other therapists who seek to understand the variety of disorders that can occur within the complicated interactive systems used to produce speech.

Judith L. Lauter, Ph.D.

CONFERENCE PARTICIPANTS

Myra Aubuchon, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Arthur Boothroyd, Ph.D.	City University of New York, New York, New York
Craig A. Bowe, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Donald R. Calvert, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Christine Clark, B.A.	Central Institute for the Deaf, St. Louis, Missouri
William Clark, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Saundra Daugherty, M.A.	Central Institute for the Deaf, St. Louis, Missouri
Donald H. Eldredge, M.D.	Central Institute for the Deaf, St. Louis, Missouri
A. Maynard Engebretson, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Ann E. Geers, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Ruth Geier, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Robert H. Gilkey, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Kenneth Grant, M.A.	Central Institute for the Deaf, St. Louis, Missouri
Christine Gustus, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Katherine S. Harris, Ph.D.	City University of New York, New York, New York
Arnold Heidbreder, B.S.	Central Institute for the Deaf, St. Louis, Missouri
Ira J. Hirsh, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Roanne G. Karzon, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Ray D. Kent, Ph.D.	University of Wisconsin, Madison, Wisconsin
Mary Lou Koelkebeck, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Victoria Kozak, M.A.	Central Institute for the Deaf, St. Louis, Missouri
Ginger Kuehn, M.A.T.	Central Institute for the Deaf, St. Louis, Missouri
Susan Langtry, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Judith L. Lauter, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Robert A. Lutfi, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Peter F. MacNeilage, Ph.D.	University of Texas, Austin, Texas
Nancy S. McGarr, Ph.D.	Haskins Laboratories, New Haven, Connecticut
David Mason, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
James D. Miller, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Jean Moog, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Michael Mudrovic, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Arthur R. Niemoeller, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Mary J. Osberger, Ph.D.	Boys Town Institute for Communication Disorders, Omaha, Nebraska
Mary Ellen Nevins, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Colin Painter, Ph.D.	Washington University School of Medicine, St. Louis, Missouri
David P. Pascoe, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Gerald R. Popelka, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri
Ellen Rajtar, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Deborah Servi, M.A.	University of Massachusetts, Amherst, Massachusetts
Stefanie R. Shattuck-Hufnagel, Ph.D.	Massachusetts Institute of Technology, Cambridge, Massachusetts
S. Richard Silverman, Ph.D.	University of Florida, Gainesville, Florida
Margaret W. Skinner, Ph.D.	Washington University School of Medicine, St. Louis, Missouri
Joan Slein, M.A.	Central Institute for the Deaf, St. Louis, Missouri
Karen Stein, M.A.	Central Institute for the Deaf, St. Louis, Missouri
George E. Stelmach, Ph.D.	University of Wisconsin, Madison, Wisconsin
Kenneth N. Stevens, Ph.D.	Massachusetts Institute of Technology, Cambridge, Massachusetts
Marcia Tash, M.S.	Central Institute for the Deaf, St. Louis, Missouri
Cindy Thomas, B.S.	Central Institute for the Deaf, St. Louis, Missouri
Janet M. Weisenberger, Ph.D.	Central Institute for the Deaf, St. Louis, Missouri

CONTENTS

Preface
Judith L. Lauter vii

Conference Participants ix

A. Invited Papers

Chapter 1 The Planning and Production of Speech
Donald R. Calvert 1

Chapter 2 How We Teach the Deaf to Speak: A Survey
Jean S. Moog 3

Chapter 3 Residual Hearing and the Problem of Carry-Over in the Speech of the Deaf
Arthur Boothroyd 8

Chapter 4 Planning and Production of Speech: An Overview
Peter F. MacNeilage, Michael G. Studdert-Kennedy, and Bjorn Lindblom 15

Chapter 5 Attention, Motor Control, and Automaticity
George E. Stelmach and Barry G. Hughes 22

Chapter 6 Developing and Disordered Speech: Strategies for Organization
Ray D. Kent 29

Chapter 7 Speech Production and Acoustic Goals
Kenneth N. Stevens 38

Chapter 8 Context Similarity Constraints on Segmental Speech Errors: An Experimental Investigation of the Role of Word Position and Lexical Stress
Stefanie Shattuck-Hufnagel 43

Chapter 9 The Role of Production Variability in Normal and Deviant Developing Speech
Katherine S. Harris, Judith Rubin-Spitz, and Nancy S. McGarr 50

B. Papers Presented to Working Groups

Chapter 10 Respiratory Function in Speech Production by Normally-Hearing and Hearing-Impaired Talkers: A Review
Judith L. Lauter 58

Chapter 11 Memory Issues in Speech Planning and Production
Craig A. Bowe 61

C. Summaries of Working-Group Discussions

Group #1 Speech Planning and Production: Step-By-Step Definition of the Task
Peter F. MacNeilage (Chairman), Jean Moog, Mary Ellen Nevins, Katherine Harris, Colin Painter, Stefanie Shattuck-Hufnagel, and Judith L. Lauter (Editor) 70

Group #2 The Role of Memory in Speech Planning and Production
George E. Stelmach (Chairman), Myra Aubuchon, Ann Geers, Ruth Geier, Ginger Kuehn, James D. Miller, Mary Joe Osberger, Joan Slein, and Craig A. Bowe (Editor) 72

Group #3 The Role of Sensory Factors in Speech Acquisition, Monitoring, and Production
Arthur Boothroyd (Chairman), Donald Eldredge, Robert Gilkey, Robert Lutfi, Gerald Popelka, Ellen Rajtar, Karen Stein, Marcia Tash, and Kenneth Grant (Editor) 72

Group #4 Instructional Strategies for Teaching Deaf Children to Speak
Ray D. Kent (Chairman), Chris Clark, William Clark, Saundra Daugherty, Mary Lou Koelkebeck, Nancy McGarr, David Mason, Michael Mudrovic, Cindy Thomas, and Roanne Karzon (Editor) 74

Group #5 Strategies for Speech Training and Speech Monitoring Aids for the Deaf
Kenneth N. Stevens (Chairman), Margaret W. Skinner, A. Maynard Engebretson, Christine Gustus, Arnold Heidbreder, Victoria Kozak, Arthur R. Niemoeller, Deborah Servi, and Janet M. Weisenberger (Editor) 75

D. Conference Summary of the Silverman Seminar
Ira J. Hirsh 79

Chapter 1

THE PLANNING AND PRODUCTION OF SPEECH

Donald R. Calvert

Central Institute for the Deaf, St. Louis, MO

The planning and production of speech is viewed quite differently by research scientists and by teachers of speech and language. It is the subject of contemplative study, reflection, and cautious observation by scientists on the one hand, and the subject of applied pedagogy, urgency, and daily action by teachers of speech on the other. What is it these two groups have to tell each other and to learn from each other about their common subject? That is the question that underlies this seminar and Report in honor of Dr. S. Richard Silverman.

The casual observer of a group of normal-hearing teenagers chattering in a school cafeteria might judge that no planning is necessary for production, and that the process is essentially spontaneous, with the appropriate pragmatics, semantics, syntax, morphology, and phonology erupting simultaneously and automatically; so rapid and effortless does the act of natural speech production appear.

We know a few ways to examine this apparently inscrutable process in order to resolve its characteristic parts and sequences, usually listening in or peeking in when the process isn't working very well. For example, we can observe "slips of the tongue," or elicit them with tongue-twisting phrases, and then make inferences about the forward planning of morphemes and phonemes. In a recent oral report, for example, one of our graduate students planning to say "the basis of judgment" said instead "the basement." Adults with brain trauma that causes aphasia provide still another look at the process in disarray. I recall one of my earliest clients at Letterman Hospital, an elderly woman who had completely lost her ability to produce speech for several months after a stroke, bursting into song accompanying a recording of a familiar Stephen Foster melody I played for her. The experience started the partial recovery of her ability to plan and produce speech. We have also been able to infer some things about the process while we watch it in slow motion as babies acquire their adult capacity through trial, error, and revision.

Hearing-impaired children, who lack one of the essential ingredients for naturally acquiring facility in planning and producing spoken language, may be yet another source of information. Furthermore, those who would have hearing-impaired children produce fluent spoken language, their teachers, have had to contrive instructional strategies and pursue sequences of procedures intended to lead the child systematically from his first perception of spoken language to the production of his own spontaneous speech. Some children learn it well—others do not. Teachers need to know more about instructional strategies that help develop spoken language efficiently when the typical process breaks down.

Schematically, the process of a hearing child's natural acquisition of spoken language production compared to a hearing-impaired child's primed development, might look like Figure 1. Spoken language is available to the normal child's ear and eye as a stimulating input from the time of birth. After a period of attending and absorbing, the child begins acquiring facility for producing spoken language by attempting numerous practice productions—sometimes imitating what others say, sometimes responding to questions or other language stimuli, and sometimes initiating his own speech. In time, the child formulates and produces fluent spoken language.

The hearing-impaired child may have available the same quantity and quality of spoken language stimulation, but the auditory perception of those stimuli may not begin until later when appropriate hearing aids are worn. Even then, the quantity and quality of spoken language perceived by the ear and eye are likely to be reduced. Experience suggests that stimulation from this limited amount of spoken language will not be sufficient to induce the child to begin or continue his own practice productions. Over a longer period of time, the teacher contrives situations that induce the child to produce

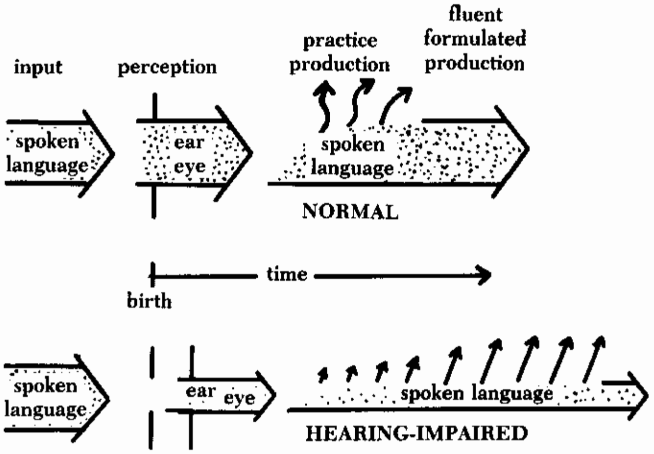


FIGURE 1. Sequence of spoken language acquisition from speech perception to fluent formulated speech production for normal and hearing-impaired children.

speech, often imitating the teacher's production, frequently responding to her questions or written symbols or display of stimulus objects, and constantly receiving reward whenever he initiates speech. In time, the child reaches his own best level of formulating and producing spoken language.

In order to develop efficient instructional strategies and to systematize their teaching procedures better, teachers need to know more about the planning and production of speech. They need to consider sensory factors, especially when sensory information is reduced and may be enhanced by sensory aids for speech reception, speech training, and speech monitoring. They also need to know the steps in planning and producing speech, both the units and sequences that occur in natural speech production, and those most useful in teaching a child deprived of normal audition.

Another pedagogical interest of teachers is the role of memory in speech and language. Recognizing the advantages of the "generation effect" on memory and recall, teachers of deaf children typically require their children to participate actively in the language learning process. For the simple act of imitation, we understand that only "echoic" short-term memory is required. But for spontaneous spoken language that is intended for communication, the child must not only retrieve words from his semantic bank of long-term memory, but at times formulate new phrases that he has neither seen nor heard. As Norma Rees describes this phenomenon of creative formulation, the child is often called upon to "say more than he knows." It is the teacher's job to build on short-term memory, aiming toward longer memory span, toward long-term memory, to overlearned memory with automatic retrieval, and at the same time help the child establish linguistic rules that permit original formulation. Which among these targets gets emphasized by teachers of hearing-impaired children often distinguishes among schools, methods, or general strategies for teaching.

Imitation involving only "immediate echolalia" is used in the classroom to practice the automatic production of articulation and speech rhythm, gradually extending the length and complexity of the units imitated. It is also a step toward "delayed echolalia" in which the child is asked to turn around to the class and repeat from memory the just imitated or recently read phrases. Repeated practice with delayed imitation in which spoken language is associated with meaning is followed by contrived opportunities for long-term recall and formulation of language. Questions asked by the teacher are used as semantic and syntactic priming to suggest some of the language the child will need in order to formulate the answering statement. Repeated picture sequences, written words, and classroom activities are used as stimuli to elicit formulated language responses, the "carry-over" requiring children to induce and recall the rules of the language with minimal teacher prompting. Sometimes the rules of language structure are made explicit by visual word-order devices, but our approach at the Central Institute for the Deaf (CID) tends more toward the generation of rules through inductive reasoning derived from use of effective examples.

It makes little difference to the teacher of deaf children whether memory is considered to be like Socrates' ball of wax with engrams traced upon it, like a telephone switchboard that calls on various regions of the central nervous system for particular details, or whether memory is considered a scattered and statistical-like activity of the brain. It is automatic recall for pragmatic language behavior that counts, and finding the most efficient instructional strategies to derive language competence is the teacher's goal.

At this 50th anniversary of Dick Silverman's quest for this goal, for he first entered this field in 1933, it is fitting that we take stock of what we know; and that scientists and teachers convene to share what we have to share. At this Silverman Seminar, we shall hear some general statements about what we know of the planning and production of speech and how we teach deaf children; and then break into working groups that include both educators and scientists to focus on five major aspects of our concern: (a) steps involved in the planning and production of speech, (b) the role of memory (c) and the role of sensory factors in the process, (d) the uses of speech training and speech monitoring aids, and (e) instructional strategies for developing spoken language. An overview will be presented by Dr. MacNeilage in the Silverman lecture "The Planning and Production of Speech."

Chapter 2

HOW WE TEACH THE DEAF TO SPEAK: A SURVEY

JEAN S. MOOG

Central Institute for the Deaf, St. Louis, MO

Those of us involved in the day to day teaching of speech to deaf children often have people comment on their impression of what our work must be like. We often hear words such as "rewarding," "requires patience," "how wonderful," "you must love it" . . . and when they learn that we teach the children to talk, we suddenly move to a slightly higher pedestal and are asked such things as "how do you do it!", "I thought that was impossible," "I did not know deaf people could talk," and so forth. We, the teachers, then patiently explain that of course it is possible and that in fact the children we teach, learn to talk and to understand when others talk to them and that, though it is hard work, it is worth it.

When talking among ourselves, we often complain about the intensity of the work, the discouragement when progress is slow, the exhilaration when one of our students makes some headway, produces his first "bababa" or speaks his first word. Perhaps somewhat like the athletes on Wide World of Sports we experience alternately "the thrill of victory, the agony of defeat."

What keeps us going is the children who demonstrate so clearly that learning to talk is within the reach of almost every hearing-impaired child, even children with severe to profound hearing impairments, such as those we teach at CID. Because every day in our school we see proof that deaf children can learn to talk, I had no hesitation in accepting this invitation to describe how we do it. I approached the preparation of this presentation with the utmost confidence but I decided that in order to bone up, it might be helpful to review some of the relatively recent literature, particularly in regard to the planning and production of speech.

After plodding through several texts and articles, I reviewed my notes to find that I had written such things as: "The wonder is that anyone at all can generate the sounds of normal speech," and "Talking requires an extremely intricate coordination of three major functional subdivisions of the speech production mechanism—tens of body parts, more than a hundred muscles, and millions of nerve cells." At that point I decided to stop reading.

It occurred to me for the first time that what I was asked to describe—in fact what we do every day—teach the deaf to speak—might scientifically be proved impossible. From the sound of this paragraph, speaking is much too complex to teach and an absurd undertaking if the students are deaf. So I shut the book and thought to myself, "too much knowledge is a dangerous thing." If I had read that paragraph 25 years ago, I certainly would have gone into another line of work.

As the author stated, "It is a wonder that anyone at all can generate the sounds of normal speech." But then, whether it is a wonder or not, we as hearing persons do talk and so do many deaf persons.

In this audience we have an unusual mix of people. Some of you are teachers in our school who spend most of your waking hours either teaching deaf children or thinking about teaching them or trying not to think about teaching them. Your daily work involves improving children's speech productions and responding to whatever comes up in the classroom. Some of you are researchers who have worked with one or two or perhaps a few more deaf children for brief periods of time, usually in a laboratory setting with very controlled conditions and usually investigating or developing training procedures for some isolated aspect of speech. You have studied a few deaf children for short periods of time and then sent them back to the classroom for "teaching." Some of you are scientists who have become experts in some aspect or other of speech production but have never seen, heard, or spoken to a deaf child.

My job this morning is to explain, describe, and demonstrate a little about how we teach deaf children to speak so you will have some information in common with us for discussions during the next two days. I thought to myself last week as I was working on this paper, "Teaching deaf children to talk is easy. What's hard is explaining how we do it."

WHAT THE TEACHER NEEDS TO KNOW

What is required of a teacher—what must one know to teach deaf children to speak? The phonetician says we must know how each consonant is produced. We should know, for example, how the mouth should be shaped, where the articulators should be placed, and whether voicing occurs. For the vowels we should know how much the tongue should be elevated and whether the velopharyngeal port should be open or closed. The acoustic phoneticist does not know how we can teach unless we know the acoustic characteristics of each sound. He thinks we ought to have the ability to read and understand spectrographs and apply that information as we teach. We should know for each child which sounds are available acoustically so that we may plan speech instruction based on those acoustical factors. Psycholinguists say it is important that we understand the development of language in normal-hearing children and apply that knowledge to the instructional strategies we use with deaf children. The physiologist wants us to understand such things as the way the muscles work to control the articulators, the workings of the larynx, the effect of the oral cavity on frequency in the vocal tract, and the control of air flow over the vocal chords. Audiologists say an understanding of how the ear works is most important and particularly what a hearing aid can and cannot do for impaired hearing. The psychologist insists that we understand about learning, that we know how the child learns and how best to keep him motivated. Furthermore, we are expected to know how to organize the material we teach into an effective learning sequence and be skilled in presenting material at the appropriate level for maximum learning. And the necessity of being sensitive to the cognitive level is also in there somewhere.

No wonder there is a shortage of teachers! No one can learn, remember, and apply all that stuff! It might be great if we could be expert in all these areas but then again it might only confuse us. Some knowledge in each of these areas is essential and most of our teachers know a little bit about most areas and a lot about some areas. But much of what we do, we do because it works, the children's spoken language improves and often we are not even sure why. We hope that this seminar will help us know why some of what we do is successful and perhaps suggest that we can do better than we are now doing. Maybe just talking about what we do will help us understand it better.

A look at the way normal-hearing children learn to talk helps us understand the profound effect that hearing impairment has on the learning of spoken language. In a seemingly effortless way the normal-hearing child learns to talk simply from hearing those around him talk. Infants begin with random vocalization and progress through babbling and jargon stages and finally by age 2 or 2½, they are producing sentences. In the beginning their speech productions are gross approximations of the sounds they are trying to produce but gradually children are able to refine their productions until these productions become very much like the speech in the environment. This gradual refinement of speech results from the child's monitoring his own productions and modifying them to more closely match the model he hears. This acquisition of spoken language appears to occur without conscious effort.

The deaf child, however, does not hear very well the spoken language around him. Even with a hearing aid he receives a very deficient and distorted version of what is spoken. This deficit in hearing affects the child's spoken language development in several ways: (a) He receives a deficient model from which to learn; (b) he is not able to monitor very well his own production which affects his ability to modify it to more closely match the model; and (c) he does not get the same pleasurable reinforcement as the hearing child from hearing himself babble and jargon, and therefore these stages come to a halt prematurely. As a result, the hearing-impaired child does not practice producing the phonemes and syllables and he therefore fails to develop the motor facility for producing speech that is unconsciously developed by the normal-hearing child during the babbling and jargon stages of language development.

In devising instructional procedures for teaching deaf children to speak, two very different skills must be considered as targets of that instruction. On the one hand the child must develop the skills necessary for producing speech and on the other hand the child must develop the ability to generate language to express ideas and to use his acquired speech skills to express this language in spoken form.

Teachers of deaf children at CID and elsewhere have been relatively successful in developing the speech skills of deaf children, particularly at the phoneme, syllable, and word levels. However, the task that proves to be more difficult and one which has received relatively little attention, is developing the ability to produce those sounds automatically and spontaneously in the context of generated language which is used for the purpose of communicating.

CHARACTERISTICS OF THE LEARNER AND THE MATERIAL TO BE LEARNED

Our objectives then are to develop in children the ability to produce speech and then to use that speech in communicating. Now let us turn our attention to two other important factors which must be considered in selecting effective instructional objectives: (a) The learner and (b) the material to be learned.

First, consider the learner, in this case the deaf child. It is important to realize that there is a fairly wide range in hearing ability even in a relatively homogeneous group of severely to profoundly hearing-impaired children such as we have at CID. Within this group, the children with the best hearing ability are able to hear some words well enough to identify them while those with the least amount of hearing are able only to perceive durational differences in spoken utterances and must learn to produce speech primarily through vision. Teaching techniques must be adapted to capitalize on useable hearing and to compensate for lack of it.

Variability in children is also apparent in motor abilities for speech production. Some children are able to produce speech sounds rather easily in imitation of a spoken model, yet others require that initially we manipulate their mouths to help them make the sounds.

We also must take into account variations in learning style. Some children learn to talk in relatively unstructured settings with only a minimum of practice and drill while others require that all sounds be presented in practice activities designed to practice each target sound. Some children rely primarily on the acoustic input while others depend on visual information and for some the written form is helpful or even necessary.

Some children have an analytic learning style and learn better when material is presented step by step and they are permitted to progress systematically from smaller to larger units; others have a more synthetic style of learning and are able to learn better when they progress from the whole to the parts, starting with larger units and backtracking to smaller units only when necessary. These differences in hearing ability, in motor ability, and in learning style need to be taken into account in selecting instructional strategies for particular children.

Characteristics of the Linguistic Information

Our experience also suggests that at least three factors affecting the child's ability to produce speech are characteristics of the spoken material rather than characteristics of the child. One factor is the size of the speech unit used for developing speech skills. There are advocates of just about every size unit of speech as being the starting point for instruction, from the phoneme to the syllable to the word to whole sentences. Our experience indicates that usually, the more easily the child learns, the larger the unit can be, and that when the child is having difficulty it is helpful to reduce the size of the unit. If the child can learn at the word level or in the context of phrases and sentences, then it is economical and usually more effective to begin there. If the child experiences difficulty, reducing the size of the unit usually will help.

A second factor which seems to affect the child's ability to produce speech is the complexity of the language the child is attempting to produce. Words are easier to produce than phrases and simple sentences are easier than complex sentences. Usually the child's speech production skills deteriorate as he moves from simpler to more complex language materials.

A third factor which appears to contribute to differences of difficulty in producing speech is the degree to which the child is required to generate new language. For example, a child's speech is usually more precise or more intelligible in an imitated production than when he is spontaneously conversing. His speech is better when producing practiced sentences than when generating new sentences. The more the child's attention must be focused on the language the more the quality of his speech is diminished.

A REVIEW OF METHODOLOGIES FOR TEACHING SPEECH TO THE DEAF

Instructional strategies for teaching speech are affected by ideas about how a child learns and theories about the most effective ways to compensate for the hearing deficit. I'd like to take about 6 or 7 min for a look at the techniques proposed for teaching the deaf during this century. This is not intended as a review of the subject, not even a sketchy one; rather it is my impression of how some of the techniques for teaching speech to deaf children developed.

Before 1950 and the advent of good hearing aids, most procedures relied heavily on visual cues. Teachers were analytic in their approach, attempting to build speech and language skills step by step. In teaching speech they started with the small unit of a phoneme or syllable and gradually built to words.

Children learned phonemes, first in isolation, then in syllables and finally in words. We did not know as much about phonology as we do now and did not realize the degree to which the production of vowels and consonants is influenced by the phonetic context. In addition because so little attention was focused on the auditory channel, the durational patterns of deaf speech were sometimes distorted. Particularly, there was a tendency toward elongated vowels and overarticulated consonants.

In teaching language, educators started with nouns, added verbs, and then built to longer sentences. Children were expected to memorize correct word order and were taught to think "First comes the subject word (who), then the verb (what did do?) and then the direct object (what) or the prepositional phrase (where). Constructing a sentence or generating one involved producing the right sort of word to fit in the particular slot. This type of teaching led to very stilted language and robot-like speech and brought about the reaction in deaf education techniques of the 1950s and 1960s.

In the 1950s and early 60s a natural language approach was more in vogue and it was accompanied by an auditory global approach to speech and language instruction. As hearing aids improved, as we gained more knowledge about linguistics and phonology, and as we learned to capitalize more on sensory information of all kinds, the emphasis of our teaching changed. By using the auditory channel we were able to get better voice quality from the children. We were able to get most of the children to produce vowels that matched more closely the intended vowel and the children learned to use lipreading and auditory cues altogether. Even our deafest children were able to match durational patterns as they learned first to receive them in a spoken model and then to produce them themselves. As more information could be processed through the ear, learning became more efficient and many children were able to work with larger units than previously. The learning of deaf children began to resemble more closely the learning of normal-hearing children, especially for those deaf children who were now able to hear some aspects of words through their hearing aids.

A Natural Language Approach was advocated in which children were exposed to a great deal of speech and language and it was expected that with enough exposure deaf children would learn just as normal-hearing children did. In some programs, educators stopped working on the development of speech sounds altogether, assuming that with time the deaf children would gradually modify their productions to be like the targets they were attempting just as normal-hearing children do and that specific speech training would no longer be necessary. Unfortunately, this did not turn out to be the case and deaf children continued to need instruction on producing the speech sounds they could not hear.

For children who had difficulty learning, remedial procedures were devised. Remedial efforts with deaf children have focused almost exclusively on reducing the size of the unit, tightening up the structure, and increasing the amount of drill provided. The most extreme example probably is the Association Method developed at CID by Mildred McGinnis. In her work with aphasic children and hard-to-teach deaf children she began with the phoneme, built to syllables, phoneme by phoneme, and then built to words, phoneme by phoneme. To learn the word "boat," for example, the child learned *b*, then *o-e*, then *b*, *o-e*, *bo-e*, then *t*, then *b*, *o-e*, *t* and finally *boat*. McGinnis believed that all modalities should be used and for each phoneme, syllable, and word the child was expected to listen to it, lipread it, say it, write it, and read it. The development of speech skills progressed slowly step by step. Emphasis was on precision of articulation and new sounds were added only when previously learned sounds had been mastered. Many tasks were designed to increase the child's memory for more sounds in sequence to form words and then more words in sequence to form sentences.

In the late 1960s and 1970s educators began incorporating into their proposed instructional programs knowledge acquired from psycholinguistic research. Today we are struggling with syntactic, semantic, and pragmatic considerations as we develop speech and language curricula for deaf children.

Also during this period the use of residual hearing commanded increased attention. Disagreement developed among educators concerning the best ways to train the child to listen effectively to make maximum use of auditory information. Educators have disagreed about the extent to which deaf children can learn to use listening for the acquisition of spoken language. At one extreme was the acoupedic or unisensory approach, which recommended that visual cues be eliminated entirely in order to better develop the residual hearing of the child. In this approach lipreading and other visual cues are removed and efforts are directed toward teaching the child to listen. In the acoupedic approach, listening is the primary sense through which deaf children learn to talk.

However, the number of deaf children who have successfully acquired spoken language skills using the acoupedic approach is relatively small. The failure of many deaf children to receive sufficient acoustic input to learn spoken language through the unisensory approach may have caused some educators to focus attention on the visual channel to compensate for the deficient hearing. In the 1970s there was renewed interest in manual communication to supplement lipreading and hearing and in the last 10 years there has been an explosion of growth in total communication programs. In total communication programs manual signs and finger spelling are used to supplement spoken language.

Another approach emphasizing the importance of visual cues to compensate for hearing deficit has been cued speech. Cued speech employs a systematic use of hand cues to supplement lipreading to clarify homophenous sounds.

Vibrotactile aids represent a very different attempt to compensate for the hearing deficit. Tactile aids have been recommended to supplement lipreading for children who demonstrate very limited auditory speech perception abilities. However, vibrotactile aids are still in their infancy in terms of practical use in the classroom or as a practical aid in communicating. Our hope for the future is the discovery of some sensory aid or new instructional approach that will compensate for hearing deficit and make more complete linguistic information available to deaf children. In the meantime, we must do the best we can with the sensory aids and instructional techniques presently available to us. I would like to share with you some of the strategies we have found effective at CID and share with you some of the problems we are still struggling with.

Perhaps the most important feature of our program and the one contributing most to our success in teaching speech is the intensity of instruction. All of our teachers are experienced speech teachers and all are firmly committed to teaching deaf children to speak. Instruction in spoken language goes on all day in all classroom activities. Children are taught in small groups of two or three children to a teacher for most of the day. Severely and profoundly hearing-impaired children probably need this kind of intensive instruction in order to learn spoken language no matter what specific teaching techniques are used.

At CID we attempt to work from both directions at the same time, from the whole to the part as we encourage children to approximate words, phrases, and sentences as best they can and from the part to the whole as we provide instruction and practice at the phoneme, syllable, and word levels.

When we are attempting to develop new sounds or to practice difficult ones, we usually work at the syllable level and backtrack to the phoneme level only if we have to. We also use these small units of speech to practice such skills as control of pitch, good voice placement, and improved matching of durational patterns. This kind of practice of specific skills in which we are stretching the child to better speech production skills is what constitutes one part of the speech lesson. The focus is on developing the motor skills for producing more intelligible speech. As with any motor skill repetitive practice is necessary and it is often exasperating as the child gets the sound, loses it, and gets it again. Then he finally has it at the syllable level only to lose it at the word level or has it at the word level but loses it as he tries to incorporate it into a sentence.

When the child is not in speech class, speech corrections are made at a level that requires less effort from the child and that does not interfere with communication. The child's spoken language including both speech and language are worked on together, often in the context of other subject areas.

When we are working on larger units, at the phrase or sentence level, the typical interaction goes something like this:

1. The child says something.
2. The teacher selects aspects of what the child says as targets for improvement.
3. The teacher cues the child to those aspects and uses a variety of techniques to help him improve his production.
4. The child's production is improved or the teacher tries another technique.
5. The teacher tries to get the child to produce the sentence again, incorporating the improved aspect.

As I mentioned earlier, when we move to connected speech, one factor that affects the accuracy of the child's production is the size of the speech unit. Syllables are usually produced more clearly than words, words are produced more clearly than phrases, and short phrases are produced more clearly than long sentences.

In addition to length of unit, another factor is the degree to which the child is generating the language. We have labeled three levels of spoken production as: (a) imitated production, (b) prompted production, and (c) spontaneous production. Speech productions are best at the imitated level and deteriorate as the child moves to prompted production and further deteriorate at the level of spontaneous production. The more the child has to think about the language, the less attention can be given to producing the speech.

VIDEO TAPES DEMONSTRATING TEACHING TECHNIQUES

As demonstrated on the video tapes, teachers use a variety of techniques to improve the child's spoken production, including some of the following:

1. Repeating what the child says, perhaps emphasizing an aspect that needs improving.
2. Pulling out a word or phrase and working on it separately.
3. Reducing the size of the speech unit.
4. Reducing the syntactic complexity of the sentence.
5. Helping the child with the language by prompting him or modeling so that all he must do is imitate.
6. Using a visual cue, such as calling attention to the mouth or tongue position or using the printed word to cue the child to the sounds or words he is omitting or distorting.
7. Describing the child's error by saying (e.g., "You forgot the first sound.")
8. Describing what the child should do by saying (e.g., "Move your tongue forward.")
9. Imitating the child's error, perhaps even exaggerating it. Imitating the error also helps the teacher figure out what the child is doing wrong and may suggest a correction technique.
10. A most important correction technique may simply be saying, "I don't understand. Tell me again." This gives the child an opportunity to correct himself. Some children clarify by improving their articulation, some by simplifying the language, and some by providing other cues.

Some teachers have suggested that once a child can produce a sound, he should be required to produce it whenever it occurs. In this way carryover will be achieved and the child will not be practicing incorrect production of sounds he has demonstrated he can produce correctly. However, such teachers are disappointed when children do not use in spontaneous conversation all the sounds they demonstrate the ability to produce in speech class. When all aspects of speech production are not yet automatic, teachers find that as soon as one piece gets fixed, another piece falls apart and if the teacher pushes too hard for perfection, the child will feel defeated.

Sometimes a speech lesson can get too intense. The teacher must constantly make decisions about what to correct, how much correction a child can tolerate before he will think "The heck with it." "I'd rather not talk at all." The teacher must manipulate the level of the speech task to be within the child's reach. She must make sure the child is successful because success will stimulate the child to continue trying.

During the majority of the school day, teachers correct spoken language in the context of conversation and content lessons. The teacher selects aspects of the child's speech which merely need reminders. At this level the child is capable of making the correction but may need to be alerted to the fact that he has made an error.

A critical feature of good teaching which I have not yet talked about is the ear of the teacher, that is the teacher's ability to carefully listen to the child's production, analyze the child's error, and figure out how best to correct it. The teacher must take into account the appropriateness of the language in relation to that child's abilities, the degree of difficulty that child has in producing that speech, the motor skills required, the acoustic characteristics, and visual characteristics of the speech the child is trying to produce.

What differentiates the skilled teacher from the less skilled one is the degree to which the teacher knows the child's skills, the number of techniques or strategies that are at her disposal (tricks up her sleeve), how competent she is in selecting an effective strategy for a particular correction and how well that strategy fits that child.

The hardest part of teaching deaf children is making decisions about what to correct and how best to convey to the child what needs correcting. In the beginning stages when the child is just developing speech, the decisions are easier. When the child is working on small units such as syllables or words, the task is relatively easy; but when the child begins producing connected language the task becomes increasingly difficult. The more the child talks the more there is to correct, the harder it is to know what to correct, and the harder it is to convey to the child what it is that is being corrected.

When teaching we have only a split second to decide what to correct and how to correct it. Each time we make a correction we do not consciously make a decision about what size unit to present, what language to use, or which sensory modality will be most effective. These considerations become second nature as we gain experience. When in doubt about what to do next, we rely on the response of the child to guide us in our instructional strategies.

We consider ourselves successful in teaching most deaf children to speak. As you could see on the tapes there is wide variability in the levels of intelligibility and in the ease with which children learn to talk. We still have a long way to go. However, perhaps as experts in various fields join forces, we will eventually find a way to teach faster and better so that more deaf children can acquire the kind of spoken language that you and I take for granted.

Chapter 3

RESIDUAL HEARING AND THE PROBLEM OF CARRY-OVER IN THE SPEECH OF THE DEAF

ARTHUR BOOTHROYD

City University of New York, New York

In speech work with deaf children, long-term carry-over, like education, may be defined as that which is left when you have forgotten what you were taught. Carry-over begins when the probability of occurrence of a speech-motor behavior increases in contexts other than that in which it was learned. It may be considered complete when the behaviors are so automatic that they retain an error probability close to zero for all contexts and all time.

There are several contexts for carry-over. A feature contrast such as consonant voicing, for example, may be taught in one place of articulation and then appear, or at least be more easily taught, in another; a specific vowel or consonant may be used in phonetic contexts other than those in which it was taught; a word may be produced correctly in an unlimited variety of sentences; and speech behavior learned by imitation may be used for communication. The variable of time must also be included. We are concerned not only with generalization, but also with retention.

One often hears the complaint that carry-over is a major problem in speech training and remediation. I would suggest that it is not just *a* problem, but *the* problem. Without carry-over, instruction is valueless. The master speech teacher who can impress a group of observers with the speedy and seemingly effortless production of a new and difficult consonant from a deaf child knows only too well that the real task of speech training has barely begun. Those of us who have invested research effort on the design of instrumental aids for speech training have had to deal with the realization that we were assisting only with the simpler, and preliminary, stages of speech instruction while contributing little or nothing to the more complex issues of generalization and retention.

In this paper I shall review the process of speech acquisition and the conditions that facilitate carry-over. Discussion of the primary role of audition will lead to a review of experimental data supporting the notion that residual hearing can play its natural role for many features and many deaf children. The paper will conclude with suggestions for ways of ensuring that carry-over occurs when the sense of hearing is too badly damaged to play its part.

SPEECH PROCESSES

All motor activities are sensorimotor activities. That is to say they require reliable, immediate, and error-sensitive feedback if they are to be effected properly. In the case of speech, we can distinguish at least three channels of sensory feedback—auditory, orosensory, and proprioceptive. In addition, there are probably several levels of internal neural feedback. These feedback channels form a set of nested loops as illustrated in Figure 1.

Unless the planning and control centers are to be totally confused, there must exist coherence among the feedback channels. That is to say, each must be providing equivalent information. Because the neural patterns themselves are not identical, the system must learn to translate from the language of each channel to the language of the others. Such learning is made possible by the existence of coherence at the physiological and acoustical levels. That is to say certain muscular activities produce certain movements which, in turn, produce certain sound patterns. One task of the developing child is to discover these coherences and thereby establish, at a neurological level, the necessary associations among the feedback channels, and between them and the planning and command instructions.

The imitation of motor-acoustic speech patterns requires, initially, that the input channel also be one of the feedback

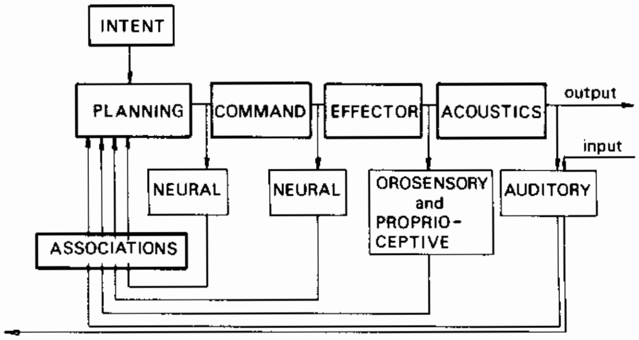


FIGURE 1. Speech, like all motor activities, requires that the planning and control centers receive feedback. Several feedback loops are involved—external sensory, internal sensory, and internal neural. One stage in the acquisition of control is the establishment of coherence among the many feedback channels. Imitation requires that one of the feedback modalities serve also as the input modality. Under normal circumstances, the shared modality for learning speech is hearing.

channels. Given this condition, the existence of previously learned sensorimotor associations permits immediate translation of input patterns into motor patterns. The need for exact correspondence between input and one component of feedback, at least during the early stages of learning, is very important and easily overlooked. The child must, at some point, experience his own speech, and the speech of others, through the same modality. This requirement does not persist however. If a second input channel is paired with the first so that the associations between the two can be learned, then the point may be reached at which the second input modality can provide the input for imitation even though it has no direct counterpart in the feedback system. For example, a totally deaf child may be able to correctly imitate a word or speech sound presented by lipreading, even though many of the key movements and positions are invisible. He can do so, however, only if, at an earlier stage of training, he was provided with sensory access to those invisible movements through a channel that was also available for feedback.

The requirement for correspondence between input and one component of feedback is normally, and obviously, met by hearing. Not only does the sense of hearing provide complete access to all significant speech movements, from breathing to articulation, but it does so during all waking hours. Moreover, the feedback is immediate and inherently sensitive. It has been argued by Stevens (1972) that the development of a phonology is guided by the need to choose motor targets in which there is the least acoustical change for a given change of position. Because acoustical changes can only be detected by hearing, it follows that the sense of hearing is the most suitable feedback modality for error-detection and correction.

The alternative input and feedback modalities presently available to the totally deaf child pale by comparison with normal hearing. Lipreading provides very limited access to speech movements and, except when a mirror is present, relies on learned, or innate, associations between facial movements and their visible appearance. Tactile exploration of the face and mouth also provides limited access and has the added disadvantage of infrequent availability. Descriptive input and feedback are slow and depend on the availability, objectivity and reliability of a teacher. Finally, our best efforts at the design of instrumental speech perception and speech training aids (Levitt, Pickett, & Houde, 1980) are still plagued by unresolved problems of signal processing, sensory coding, and limited wearability. It is indeed a tribute to the skill, insight, and perseverance of pre-electronic speech teachers that their efforts often led to reasonably intelligible speech.

The issue under consideration, however, is not the mechanisms by which speech skills are first acquired but their generalization and retention. What exactly are the conditions that increase the likelihood of carry-over?

The first of these conditions is frequent repetition. The more often the child goes through the correct movement patterns, the more firmly do they become embedded in memory; the more likely they are to be retained over time; the more probable it is that some key feature will generalize to other phonetic contexts; and the more probable it is that the patterns will appear in communicative speech.

The issue of repetition in speech instruction is a difficult one. Repetition, drill, and rote learning are not exactly popular as educational strategies at the present time. Moreover, the time typically devoted to speech instruction does not allow for much supervised repetition, and without supervision, the child is likely to spend more time repeating incorrect movement patterns than correct ones. This last problem can become chronic if the instruction is remedial rather than developmental. Some years ago, during a study of computer-assisted speech training (Boothroyd, Archambault, Adams, & Storm, 1975; Nickerson, Kalikow, & Stevens, 1976), I did a little time and motion study of remedial speech work. During a daily 20-min period of individual instruction (surely a luxury in most programs for deaf children), I noted that some 5 min were typically lost in coming, going, and getting started. During the remaining 15 min, the instructor spoke about twice as often as the student, leaving perhaps 5 min for the student's production of motor-acoustic speech patterns. Furthermore, because this was remedial instruction, the student tended to produce 2 or 3 incorrect patterns for every correct one, allocating perhaps 2 min for the correct ones. While quite a lot of repetition can occur in 2 min, it is not likely to lead to carry-over when it is competing with the more frequently used, and more firmly entrenched, incorrect motor patterns.

How can hearing help in repetition? It can be hypothesized that, under normal circumstances, listening to speech is equivalent to producing speech. Once the associations among input, feedback, planning, and command are firmly established, input patterns can be translated into their motor equivalents, at least at a neurological level, so that both speaking and listening contribute to the total amount of effective repetition. If this is the case, then the nonauditory child is at a double disadvantage. The time spent watching speech, for example, can only contribute to the effective repetition of those speech movements that are visible, while the invisible movements go unrehearsed.

The second condition facilitating carry-over, especially across communicative contexts, is the establishment of automaticity (Ling, 1976). Automaticity may be said to occur when the control of frequently repeated sets and sequences of movements become delegated to neural "subroutines," placing fewer demands on the limited capacity of higher level control systems and therefore suffering less interference from the tasks of formulating ideas and linguistic structures.

A third condition is the continuous availability of an appropriately sensitive error-detecting feedback system. If we hypothesize the existence of targets or target-ranges for certain features of speech movements, it becomes important that deviations outside those target ranges produce large and noticeable changes in the feedback patterns, while movements within the target range produce little or no change. This concept, as mentioned earlier, has been invoked by Stevens (1972) to explain the evolution of phonological systems. If the acoustical component of the speech process is a primary determinant of phonological structure, it follows that audition must be an ideal modality for error-detection. It seems equally clear that if the child must rely solely on proprioceptive and orosensory feedback, the regions of zero gradient on a graph of speech movement versus feedback pattern are likely to shift. That is to say the child may gravitate to other target ranges in which he feels more secure that a deviation will be signalled by the available feedback. This concept has been evoked by Willemain & Lee (1972) to account for the high fundamental frequency often found in the speech of nonauditory children.

A related issue is the need for a feedback system that will signal the need for change in response to physiological maturation. Because growth and maturation of the speech mechanism affect the relationships of motor activities to acoustic output, hearing again becomes the ideal modality for ensuring long-term retention through adaptation.

The final condition that I shall mention is that of reinforcement. Generalization across contexts and retention over time may be facilitated by rewards for appropriate speech behaviors and punishments for inappropriate ones. Two kinds of reward are of relevance to the present discussion. One is success in communication, the other is teacher approval. The corresponding punishments are failure in communication and teacher disapproval. Both are powerful tools, but they should be used with great caution. There is a danger, for example, that they may be used to reinforce inappropriate speech behaviors. Another danger is that true failure in communication is more likely to produce a change of strategy than an improvement of speech, and the child who senses that his communicative efforts are valued for their acoustic form rather than their content may choose to avoid communication with the offending party. Moreover, when accurate speech is demanded, the associated anxiety may produce physiological responses such as muscular tension, that are not conducive to good speech production. Nevertheless, in an atmosphere of mutual trust and respect, an insightful and sensitive teacher can use communicative and social reinforcers to facilitate the carry-over of new skills known to be within the child's capabilities.

The auditory issue related to reinforcement is not the hearing of the child but the hearing of the teacher. Among the many skills required of a speech teacher is the ability to analyze and evaluate the acoustic output of the child independently of, but simultaneously with, linguistic and communicative content. As Sibley Haycock (1933) stated:

> An inexperienced (though possibly "trained") speech teacher, with a poor phonetic ear—that is to say, with an indifferent, undiscriminatory tone perception,—can hopelessly ruin the natural quality of the voice of a young deaf child within the space of 12 months.

Consideration of the factors involved in speech acquisition, generalization, and retention quickly leads to the conclusion that hearing is so well-suited to the task of providing input and feedback, that it will have to be very severely damaged before it is less effective than the alternatives presently available. There is, moreover, considerable empirical support for this position in the form of severely and profoundly deaf children who have attained high levels of proficiency in spoken language as a result of auditorily based training (e.g., Ling & Milne, 1981). It is not always clear, however, that hearing alone can account for these successes. Many other factors must have contributed, and it is possible that these factors could still operate well in the absence of hearing. Nor is it clear what happens to children who do not succeed in auditorily based programs—or why. Moreover, logic alone tells us that there must be a point at which hearing is so badly damaged that it cannot play a useful role as an input and feedback modality for speech acquisition and retention.

What follows is a summary of research conducted by the author into the auditory capabilities of children with severe and profound hearing losses, and on the links between perception and production in the acquisition and retention of speech skills.

EXPERIMENTAL DATA

Speech Perception as a Function of Degree of Hearing Loss

Figure 2 shows the probability that subjects with various degrees of hearing loss will be able to perceive speech pattern contrasts along selected dimensions. These data are derived by interpolation and extrapolation from the results of experiments on orally trained subjects using forced-choice procedures (Boothroyd, 1984). They show, as might be expected, that increasing hearing loss (as measured by better ear, three-frequency average pure-tone threshold) is associated with decreasing performance on all dimensions but that some dimensions are more susceptible to the effects of sensorineural damage than others.

From these data it may be determined that the levels of hearing loss at which the probability of perception of specific

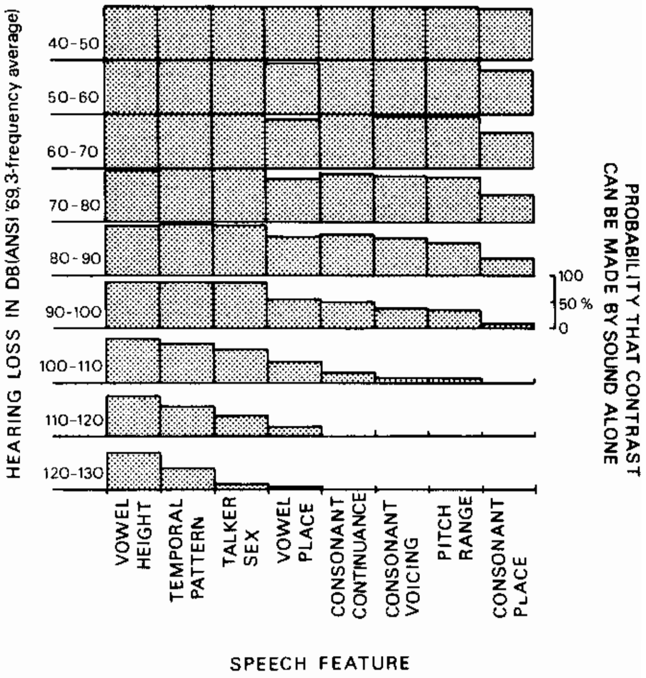


FIGURE 2. The accuracy of perception of speech pattern contrasts along several feature dimensions as a function of degree of sensorineural hearing loss. These data were obtained by extrapolation and interpolation from the results of experiments reported in Boothroyd (1984).

contrasts drops to 50% (after correction for chance) are as follows:

For consonant place:	75 dBHL
For initial consonant continuance:	85 dBHL
For initial consonant voicing:	90 dBHL
For vowel place:	100 dBHL
For talker sex:	105 dBHL
For syllabic pattern:	115 dBHL
For vowel height:	in excess of 115 dBHL

It may further be noted that the contrasts available to the subjects with the most severe hearing losses tend to be those that are perceptible from time/intensity patterns alone.

More recent data using an expanded version of the speech pattern contrast test and a different population are shown in Figure 3. Although the average scores are somewhat lower than in the earlier studies, the findings are very similar, as are those of studies by numerous other researchers (e.g., Erber, 1972; Hack & Erber, 1982; Pickett, Martin, Johnson, Brandsmith, Daniel, Willis, & Otis, 1972; Risberg, 1976; Smith, 1975).

All of these data, however, must be interpreted with caution. We know very little about the nature or causes of the differences of perceptual performance that are observed among subjects with identical hearing losses; little research has been done on the influence of learning on performance; and none of these studies has sought to ensure that the amplification systems delivering the speech patterns were optimally matched to the psychoacoustic characteristics of the subjects. Until these three issues are adequately addressed, the results of experiments on the relationships between speech perception and pure-tone threshold should be taken as lower estimates of auditory potential rather than upper limits. Nevertheless, even these lower estimates show the existence of considerable auditory potential in subjects who only a few decades ago were thought to be beyond the reach of auditorily based training (Hudgins, 1960).

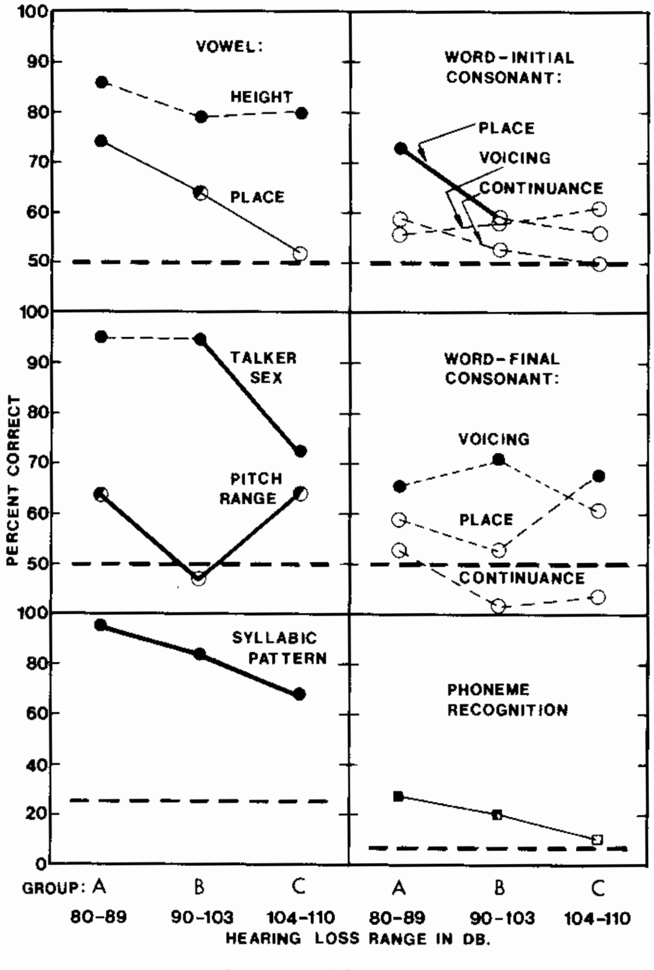


FIGURE 3. Average performance of three groups of hearing-impaired teenagers on 11 subtests of a forced-choice, speech pattern contrast test. Data are also shown for an open-set word recognition test, scored phonemically. Horizontal broken lines show expected chance scores. Filled and half-filled data points are significantly different from chance at the 99% and 95%, levels, respectively. Heavy & light solid lines joined means that they are significantly different at the 99% and 95% levels, respectively.

Speech Production as a Function of Degree of Hearing Loss

Figure 4 shows the relationship between speech intelligibility and better ear, three-frequency average threshold in a group of orally-trained teenage students (Boothroyd, 1984). Speech intelligibility scores are the average percentage of key words recognized by groups of six inexperienced listeners from tape recordings of six read sentences (Magner, 1972). Once again we see a clear association between intelligibility and degree of hearing loss. Furthermore there is a significant

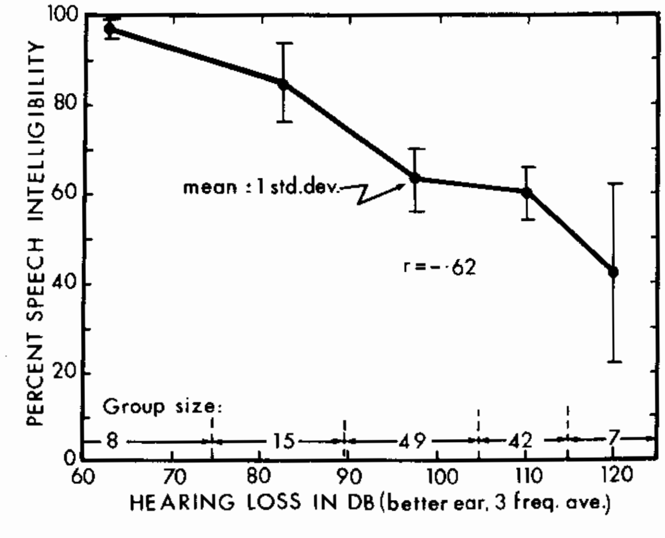


FIGURE 4. Speech intelligibility, measured as the probability of recognition of words in sentence context, as a function of degree of sensorineural hearing loss in 120 orally-trained preteenagers and teenagers. Hearing-loss values are the better ear, three frequency average threshold (Boothroyd, 1984). The difference of mean intelligibility between the 105 to 114 dBHL group and the 115 to 124 dBHL group was statistically significant ($t(46) = 2.14$, $p < .05$).

difference of mean intelligibility between subjects with hearing losses of 105 to 114 dBHL and those with losses in excess of 114 dBHL, suggesting that hearing can contribute something to speech intelligibility for losses as high as 105 to 114 dBHL.

Figure 5, provides evidence of an association between degree of hearing loss and the intelligibility of the phoneme /sh/. The abscissa shows pure-tone threshold at 2kHz and the ordinate shows the probability that /sh/ will be correctly identified by normally hearing listeners in a 2-alternative forced-choice task. The comparison phoneme was/s/ and the stimuli were in sentence or phrase context (e.g., "take a sip" vs "take a ship"). The correlation coefficient of 0.77 reached the 1% level of significance and the regression function suggests that scores do not reach chance levels until a threshold value of 115 dBHL is reached (data from Boothroyd & Gorzycki, 1977). Once again we find evidence that the presence of residual hearing can affect intelligibility, even at the segmental level, for hearing losses as high as 110 dBHL.

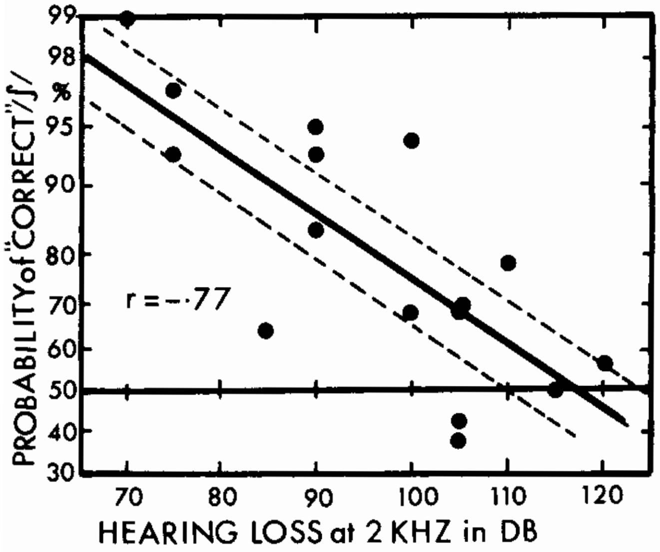


FIGURE 5. Intelligibility of the consonant /sh/ as a function of degree of hearing loss at 2KHz in 16 hearing-impaired subjects.

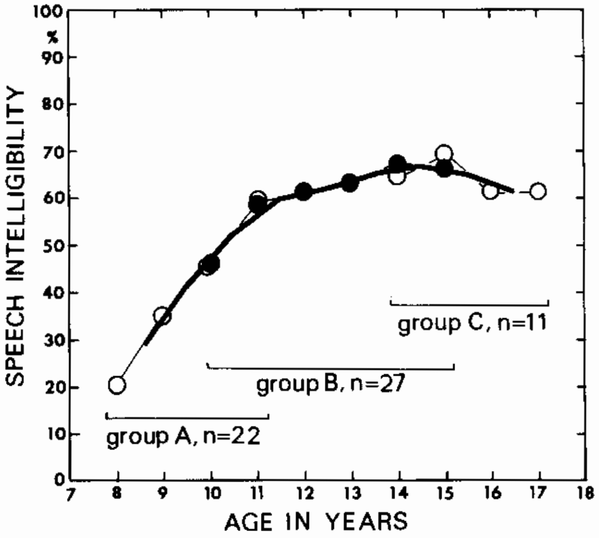


FIGURE 6. Speech intelligibility as a function of age in 3 groups of subjects with severe and profound sensorineural hearing loss. Data for the youngest and oldest groups have been "normalized" so that the mean scores for the two overlapping years are identical.

Speech Production as a Function of Age

Since 1968, the coordinator of the speech department at the Clarke School for the Deaf has measured the intelligibility of students' speech twice a year (Magner, 1972). The test materials are sentences which the students read, and the auditors are student teachers who at the beginning of the school year may be considered relatively inexperienced. The results collected between 1968 and 1978 were examined, and the data were regrouped according to the students' age at the time of testing. We were able to find a group of 22 students for whom intelligibility scores were available between the ages of 8 and 11 years, another 27 for whom scores were available between ages 10 and 15 years, and another 11 for whom scores were available between ages 14 and 17 years (Boothroyd & Lambert, 1980). The average speech intelligibility scores are plotted as a function of age in Figure 6. By extrapolation, it may be inferred that the average intelligibility begins at 0% at age 6½ years. It rises steadily at the rate of approximately 12% per year, reaching 60% at around age 11½ years. Between ages 11½ and 14½ years there is a slow rise at the rate of 2½ percentage points per year to a maximum of 77%. The next 2 years see a decline to about the same score that was obtained at around age 11 or 12 years. These data parallel perfectly those reported by Hudgins (1960).

Figure 7 shows these data broken down further by hearing loss. The pattern is essentially replicated within each hearing-loss group except that subjects with losses from 85 to 95 dB show less positive and negative changes of intelligibility between ages 11½ and 16½ years. This is, in part, due to the

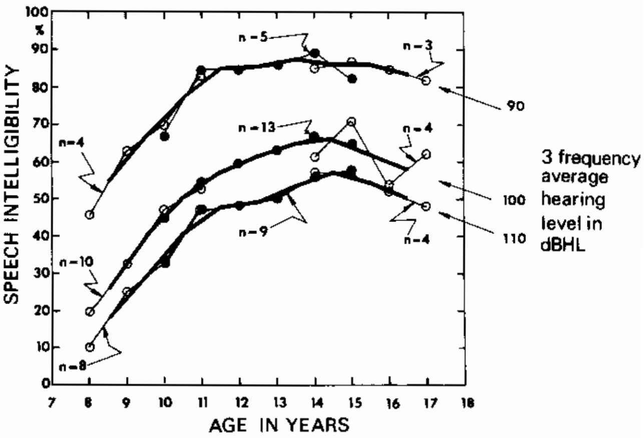


FIGURE 7. Speech intelligibility as a function of both age and degree of hearing loss in 3 groups of subjects with severe and profound hearing losses. Taken at face value, these data indicate that 5 dB of residual hearing are worth 1½ years of instruction in terms of their influence on speech intelligibility.

ceiling effect of percentage scores. If these scores are plotted on a probability scale the three curves are essentially parallel.

Hudgins (1960) hypothesized that the slight deterioration of intelligibility after age 14 years may be due either to the psychological changes accompanying adolescence, or to the expansion of curriculum with correspondingly less emphasis on speech instruction and remediation. Whatever the reason, these data suggest that, at least within the context of a school for the deaf, retention of speech skills cannot be taken for granted. The data also illustrate graphically the benefits of residual hearing. The 90-dBHL group attains a speech intelligibility score of 55% by age 8½ years. The group with a 10-dB greater average hearing loss requires another 2½ years of training and maturation to reach this score, while the group with a 20-dB greater average hearing loss does not reach this score until age 14½ years. If these data are taken at face value, they indicate that a 5 dB difference in residual hearing may be worth as much as 1½ years of speech instruction.

Speech Production as a Function of Speech Perception

In a more recent study, I have examined the speech intelligibility of deaf subjects as a function of their auditory speech perception performance. Receptive and expressive scores were obtained for (a) a forced choice segmental feature contrast test, and (b) a test involving the recognition of monosyllabic words presented in a carrier phrase. Expressive scores were also obtained for a sentence test. Figure 8 shows the relationships between feature-contrast perception and feature-contrast production for the eight features tested. Significant correlations were found between perception and production for vowel height, vowel place, and consonant voicing, both word-initial and word-final. Production of consonant place and continuance contrasts was not significantly correlated with auditory perception.

It will be seen from Figure 8 that significantly suprachance scores on the vowel production tasks were obtained by subjects whose receptive scores were at chance levels. For consonant voicing, however, auditory perception appears to be essential. When we examine the receptive and expressive scores averaged across all eight features, and also the word recognition scores, scored both phonemically and by whole word, we again find a high correlation between perception and production together with significant production performance even in cases of zero auditory perception (Figure 9).

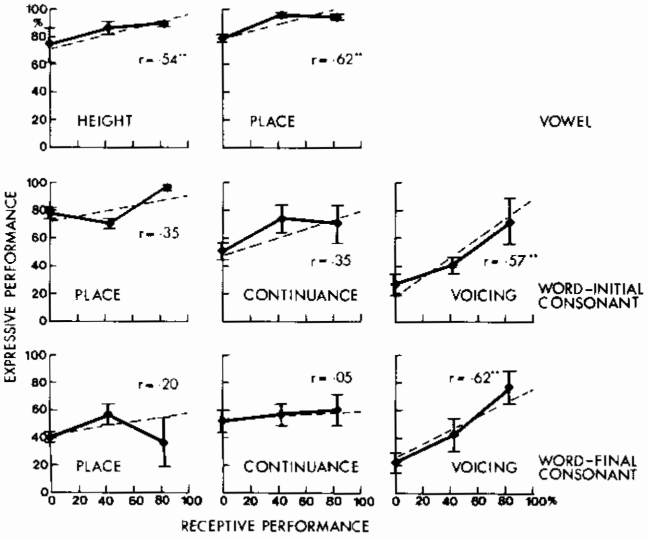


FIGURE 8. Expressive performance as a function of auditory receptive performance on eight segmental subtests of a forced-choice, speech pattern contrast test. Subjects were 20 teenagers with severe and profound sensorineural hearing loss.

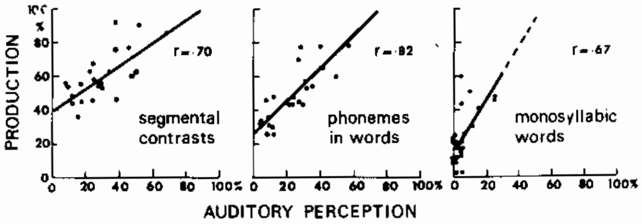


FIGURE 9. Expressive performance as a function of auditory receptive performance on tests requiring the perception of segmental contrasts, the recognition of phonemes within consonant-vowel-consonant (CVC) words, and the recognition of CVC words spoken in a carrier phrase. Subjects were 20 teenagers with severe and profound sensorineural hearing losses.

Influence of Residual Hearing on Carry-over

In an unrelated study, the role of hearing in carry-over revealed itself unexpectedly. Two groups of subjects were trained in pitch-matching tasks using both a visual pitch display and noninstrumental teaching. Performance on pitch control was measured before, during, and immediately after training, and again some weeks after the training was finished. To my surprise, some of the subjects continued to improve after training had stopped. Examination of the data showed that these were the subjects with more residual hearing in the low frequencies, though all of the subjects had "left hand corner" audiograms (Boothroyd, 1973).

CONCLUSIONS

There is much that these data do not say about carry-over and retention. None of these results for example relate to a communicative speech context. Nor were there specific attempts to measure generalization across phonetic and linguistic contexts. In addition they reflect only the performance of particular groups of students taught in particular programs.

Nevertheless, the data indicate a tremendously important role for hearing in the acquisition, generalization, and retention of motor speech skills, even in subjects with losses as high as 105–114 dBHL. At the same time, they reveal the limitations of hearing and provide specific guidance about which features are least likely to be accessible auditorily. The findings also indicate which features are least likely to be accessible by vision and touch and therefore are in the greatest need of improved auditory reception or instrumental sensory assistance.

When hearing cannot play its full role in the generalization and retention of motor speech skills, what steps might be taken to compensate for such deficiencies? I offer the following suggestions—most of them distinguished by a lack of originality:

1. Care should be taken to see that faulty speech behaviors are neither taught, nor allowed to persist. Carry-over is difficult enough when the skills being taught are new. If they must compete with well-rehearsed errors, carry-over becomes almost impossible (Ling, 1976).
2. Maximal opportunity should be provided for repetition, if necessary through drill. To this end, it may be helpful if simple visual or tactile speech displays are used by the student for independent practice (Boothroyd, 1977, 1985a).
3. Teachers should not wait for spontaneous carry-over to communicative contexts, but should pursue it. The concept of speech instruction as something that occurs only at specific times, in specific places, with a specific person, cannot be justified. All teachers should be providing reinforcement, and opportunities for use of appropriate speech skills in all possible contexts.
4. Drills should focus on the production of speech contrasts in addition to specific targets. It is more important that the child consistently divide phonetic space into clearly definable categories than that a few of the categories correspond with those of hearing speakers.
5. Objective methods of evaluation should be used that prevent learning on the part of the teacher from being interpreted as speech improvement on the part of the child. To this end it should be noted that forced-choice procedures of the type referred to earlier, have the advantage of eliminating the experienced-teacher advantage in intelligibility testing (Boothroyd, 1985b).
6. The search should be continued for sensory assistance and substitution devices. Although we are far from realizing the goal of a wearable sensory aid that is a complete substitute for hearing, there is every chance that we can improve carry-over with wearable aids designed to display those speech features that are most in need of display. Such devices should, however, encode the relevant features in such a way as to fill the error-detection role discussed earlier.
7. More research should be conducted on the issue of carry-over itself. Research on techniques of speech instruction has focused largely on the initial acquisition of speech skills. There is a great need for research on the processes of generalization and retention and the factors that might influence them.
8. Finally we need to prepare personnel with the knowledge, skills, and experience commensurate with the demands of the task. No amount of research will be of assistance without the professional personnel to take it from the laboratory into the field. The wholesale abrogation of responsibility for speech teaching by educators of the deaf (Scott, 1983) has left a void which cannot yet be filled by speech-language pathologists. It is to be hoped that we shall return to a condition in which teachers of the deaf, as professionals, recognize their primary role as specialists in the communicative development of deaf children.

REFERENCES

BOOTHROYD, A. (1973). Some experiments on the control of voice in the profoundly deaf using a pitch extractor and storage oscilloscope display. *IEEE Transactions on Audio- and Electroacoustics, 21,* 274–278.

BOOTHROYD, A. (1977). *Development of small speech training aids: Third progress report.* (Sensory Aids Research Project, Report No. 27). Northampton, MA: The Clarke School for the Deaf.

BOOTHROYD, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech and Hearing Research, 27,* 134–144.

BOOTHROYD, A. (1985a, May). A wearable tactile intonation display for the deaf. *IECC Transactions on Acoustics, Speech, and Signal Processing, 33,* 111-117.

BOOTHROYD, A. (1985b). Evaluation of speech production of the hearing impaired: Some benefits of forced-choice testing. *Journal of Speech and Hearing Research, 28,* 185-196.

BOOTHROYD, A., ARCHAMBAULT, P., ADAMS, R., & STORM, R. (1975). Use of a computer based system of speech training aids for deaf persons. *Volta Review, 77,* 178–192.

BOOTHROYD, A., & GORZYCKI, P. (1977). *Sibilant articulation in hearing-impaired children.* (Sensory Aids Research Project Report No. 29a). Northampton, MA: The Clarke School for the Deaf.

BOOTHROYD, A., & LAMBERT, L. (1980). *Influence of age and hearing loss on speech intelligibility.* (Sensory Aids Research Project Report No. 33.2). Northampton, MA: The Clarke School for the Deaf.

ERBER, N. P. (1972). Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. *Journal of Speech and Hearing Research, 15,* 413–422.

HACK, Z. C., & ERBER, N. P. (1982). Auditory, visual, and auditory-visual perception of vowels by hearing-impaired children. *Journal of Speech and Hearing Research, 25,* 100–107.

HAYCOCK, G. S. (1933). *The teaching of speech.* Washington, DC: Alexander Graham Bell Association for the Deaf.

HUDGINS, C. V. (1960). The development of communication skills among profoundly deaf children in an auditory training program. In A. W. G. Ewing (Ed.), *The modern educational treatment of deafness.* Washington, DC: A. G. Bell Association for the Deaf.

LEVITT, H., PICKETT, J. M., & HOUDE, R. (1980). *Sensory aids for the hearing impaired.* New York: IEEE Press.

LING, D. (1976). *Speech and the hearing-impaired child: Theory and practice.* Washington, DC: A. G. Bell Association for the Deaf.

LING, D., & MILNE, M. (1981). The development of speech in hearing-impaired children. In F. H. Bess, B. A. Freeman, & J. S. Sinclair (Eds.), *Amplification in education.* Washington, DC: A. G. Bell Association for the Deaf.

MAGNER, M. (1972). *A speech intelligibility test for deaf children.* Northampton, MA: The Clarke School for the Deaf.

NICKERSON, R. S., KALIKOW, D. N., & STEVENS, K. N. (1976). Computer-aided speech training for the deaf. *Journal of Speech and Hearing Disorders, 41,* 120–132.

PICKETT, J. M., MARTIN, E. S., JOHNSON, D., BRANDSMITH, S., DANIL, Z., WILLIS, D., & OTIS, W. (1972). On patterns of speech feature reception by deaf listeners. In G. Fant (Ed.), *Speech communication ability and profound deafness.* Washington, DC: A. G. Bell Association for the Deaf.

RISBERG, A. (1976). Diagnostic rhyme test for speech audiometry with severely hard of hearing and profoundly deaf children. *Speech Transmission Laboratory Quarterly Program and Status Report.* (Nos. 2 & 3, pp. 40–58). Stockholm: Karolinska Technical Institute.

SCOTT, P. L. (1983). Have competencies needed by teachers of the hearing impaired changed in 25 years? *Exceptional Children, 50,* 48–53.

SMITH, C. R. (1975). Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research, 18,* 795–811.

STEVENS, K. N. (1972). The quantal nature of speech: Evidence from articulatory acoustic data. In E. R. David & P. B. Denes, (Eds.), *Human communication–A unified view* (pp. 51–66). McGraw-Hill.

WILLEMAIN, T. R., & LEE, F. F. (1972). Tactile pitch displays for the deaf. *IEEE Transactions on Audio- and Electroacoustics. 20,* 9–16.

Chapter 4

PLANNING AND PRODUCTION OF SPEECH: AN OVERVIEW

Peter F. MacNeilage
University of Texas, Austin, TX

Michael G. Studdert-Kennedy
City University of New York, NY
and
Haskins Laboratories
New Haven, CT

Bjorn Lindblom
University of Stockholm

This paper is a progress report on part of a biologically-based approach to the understanding of spoken language presently in preparation (Lindblom, MacNeilage, & Studdert-Kennedy, 1984a). Our aim is to view language in the perspective of the evolution of adaptive functions. Our strategy is to begin by focusing on aspects of language which are closest to the transmission process, namely, the production and perception of the sound pattern of language; in linguistic terms, the phonological level. Our rationale is partly practical—these more peripheral aspects are closer to direct observation than more central aspects associated with meaning and thought. But it is also partly theoretical—we believe the constraints of the transmission process have played a crucial role in molding the form of language functions as a whole.

The topic of this paper is the serial organization of language output—the process whereby some intention, which is itself not serially organized, is converted into a rule governed sequence of linguistic symbols. If we look directly at the result of this serial organization process—if we look at the movements of the speech apparatus and their acoustical resultants—we immediately encounter the central paradox of speech research. We find that the parts of this output that are supposed to signal the string of consonants and vowels given us by the linguist are neither context-free, nor marked off discretely from segment to segment. The context-sensitive representation of a given consonant or vowel in the transmission process is termed the *Invariance problem*. The absence of obvious temporal boundaries to the representation of segments is termed the *Segmentation problem* (see Lindblom, 1982). Collectively these two problems constitute what we can call the *Nonisomorphism Paradox*.

Context-sensitivity seems to result primarily from the overlap of gestures required for adjacent phonemes. This overlap is termed *coarticulation* (Kent & Minifie, 1977). An early attempt to account for these coarticulatory effects was the Motor Command Hypothesis (Cooper, Liberman, Harris, & Grubb, 1958) according to which the control system did send out an invariant command for a given segment in all of its contexts, but context effects were produced simply by temporal overlap in the effects of adjacent commands on the mechanically sluggish articulatory system. This hypothesis was laid to rest by electromyographic studies that showed that the instructions sent to the muscles for instances of a given vowel or consonant varied with the identity of both the preceding and following segments (MacNeilage, 1970). Thus, context sensitivity was shown to be, at least in part, built into the control process. With respect to the segmentation problem, a number of attempts were made to show that while coarticulation patterns in continuous speech did not delimit intersegment boundaries they might indicate the presence of higher order boundaries such as syllable boundaries, word boundaries or morpheme boundaries. This search for evidence for units other than the segment in the transmission process has been for the most part unsuccessful, just as the search for evidence for segments has been (Kent & Minifie, 1977). Thus we cannot resolve the nonisomorphism paradox by denying that the signal lacks invariance or discreteness, either for the segment or other units. Can we resolve the paradox by denying the existence of discrete units underlying the production process? Confining our attention to the segment for the moment (i.e., the consonant or vowel) the answer is emphatically no. Errors involving single segments in an otherwise correct output sequence happen often enough in normal subjects to suggest that the segment is an independent unit in the control process. The most well-known example is the reversal or spoonerism, for example:

fish and tackle → fash and tickle

But segments are also omitted, added, shifted in location, and substituted for other segments. When these things happen, the resultant output of the erroneous form seems to be correct at the transmission level, suggesting that the error occurs at an underlying level, and that processes directly concerned with transmission, particularly with coarticulation, can operate independently of more underlying levels (MacNeilage, 1982).

Another source of evidence for the independence of surface and underlying levels of organization comes from a number of studies done recently in which subjects attempt to produce speech after bite blocks of various sizes are placed between the teeth (e.g., Lindblom, Lubker, & Gay, 1979). It has been shown that normal subjects have a remarkable ability to adapt to these blocks virtually immediately, even though such adjustment probably involves generating unique control signals for each segment.

Thus the peripheral stage of speech transmission can perhaps best be described as "an elegantly controlled variability of response to the demand for a relatively constant end" (MacNeilage, 1970, p. 184). However, it is important to note that this phenomenon itself, often termed *motor equivalence,* is not specific to speech or even to human action. Instead it is characteristic of goal-oriented action in general and can be seen in operation well down the phylogenetic scale. For example, Fentress (1983) has noted that if restrictions are placed on limb movements of mice while grooming, movements are reorganized so that grooming successfully occurs. But two aspects of the speech action are probably specific to humans—first, the rate at which independent acts are carried out. We speak at a rate of about 14 segments per second. The fastest rate we are aware of in other primates doing roughly comparable tasks is a rate of about 7½ acts per second in baboons opening a puzzle box (Trevarthen, 1978), but that was after practice on a single sequence and not for sequences in general. Second, the principles of segment organization at the underlying level are probably unique to human vocal output in ways that we will now discuss. Again phonological errors of normal subjects are the best source of information about the organization of underlying levels. These errors show definite patterns, both in terms of the relative involvement of various units and in terms of the places to which units can migrate when they are misplaced. The single segment is by far the most popular unit of error. Collections of errors show it to participate liberally in all five classes of error mentioned earlier (reversals, shifts, additions, omissions, and substitutions). On the other hand whole syllables rarely move around as units and syllable reversals are virtually unknown. Also, as Shattuck-Hufnagel and Klatt (1979) have argued, there is very little evidence that the distinctive feature is a separable unit in errors of serial organization.

Migration patterns for segments in errors are very narrowly defined. For example, in reversal errors it is extremely rare for segments to move to a different position in the syllable than the one they came from. Vowels and consonants never reverse. Consonants which were supposed to precede or follow a vowel in the correct syllable seldom reverse positions. The best conception of what is actually happening at this stage of production has been developed by Shattuck-Hufnagel (1979). She visualizes the process as one of first scanning for information about segments in representations of words from a mental dictionary, and then copying the information into a second representation more directly related to output. She interprets the positional restrictions on segment migration to mean that segmental representations must be copied into slots marked for syllable structure. Additional evidence for a separation between a content specification and a specification for serial organization comes from observations of Shattuck-Hufnagel on omission errors. She notes that a number of omission errors such as "Dr. ___inclair has emphasized" (Sinclair) give the impression that a null consonantal segment is initially represented for the beginning of the word "emphasized," and it is this null segment that is copied into the beginning of the word "Sinclair" instead of the initial "s." If even a representation of the absence of a content element can be moved around to a corresponding structural location, then it would seem that content and serial structure must be independently specified. We would like to describe the form of operation suggested for the phonological level by Shattuck-Hufnagel as a Frame/Content mode of organization. Segmental content elements are copied into syllabic frames.

Language is widely described as having a dualistic structure. At the phonological level, segments are concatenated into patterns. So, for example, one gets the words "tack," "cat" and "act" from the same three segments. But, in addition, there is a morphological or meaning level at which meaning units are concatenated into various forms. So, for example, one gets "John hit Mary" or "Mary hit John," or, according to a slightly different principle, "blind venetian" and "venetian blind" (Miller, 1965). Evidence from language errors suggests that the morphological level of language output may also have a frame/content mode of organization, analogous to the one found in phonology. Garrett (1975) has noted a large class of reversal errors in which what could be called a grammatical frame remains in its correct order while the stem forms of content words (nouns, verbs, adjectives, and some adverbs) are apparently inserted in the wrong positions. An example is: McGovern favors busting pushers ➧ McGovern favors pushing busters. Note here that the grammatical bound morphemes "ing" and "ers" remain in correct position while the content word stems "push" and "bust" exchange. Garrett also notes a class of apparent reversals of adjacent forms (e.g., "little beads of blood" ➧ "beads of little blood") which he believes are also best interpreted as a misplacement of content elements in their syntactic frames.

It is not possible to do justice to the arguments of Shattuck-Hufnagel and Garrett in this context. But their work persuades us of the importance of the frame/content mode of organization at both the phonological and morphological levels. Part of the importance of this mode of organization for spoken language derives from the likelihood that many other complex output processes do not possess it. In typing, for example, reversal errors usually involve adjacent letters. A consideration of the letters involved in reversal errors (MacNeilage, 1984a) reveals no sign of the prohibition of reversals between consonants and vowels seen in speech errors. We do not know of error data for musical performance but the intuition of musicians we have talked to is that they do not tend to exchange elements with common positions in a musical structure (e.g., first beats in the bar). The equivalent of the phonological level of spoken language in the sign language of the deaf seems to be sets of four sign attributes that form something like a monosyllabic word. The attributes are (a) handshape, (b) location, (c) orientation, and (d) movement. Although we know of no collection of sign reversal errors, there appears to be no serial structure restriction equivalent to that in the spoken syllable that would constrain the positions of elements in reversal errors. On the other hand, errors might reveal a frame/content mode of organization at the morphological level. At this level syntactic morphemes are typically signalled by superimposing movements on concurrent signs for lexical stems. It would be at least possible for syntactic information to stay at its correct location in the utterance while signs for lexical stems migrated around.

From our biological standpoint it is now necessary to consider how a frame/content mode of organization might have evolved. Perhaps the first question to ask is: Evolved from what? The best estimate of the status of vocal communication systems at the time when hominids first diverged from an ancestral line common to great apes comes from considering present great ape vocal communication in the natural state. (Incidentally, evidence from molecular biology suggests that the divergence may have occurred as recently as 7 million years ago (Pilbeam, 1984). Evidence suggests that rather than having a dualistic system with concatenation rules at two levels, great apes have a limited number of cries, perhaps not exceeding 30 (Dingwall, 1979), with no combination rules either within cries or between them. One way to pose the question of the evolution of serial organization of language is to ask how did we get from a communication system like that of the great apes to the human one? A plausible scenario for the phonological level has been provided by Hockett and Ascher (1964). They suggested that the principle of sound concatenation may have been forced by the inability of the transmission system (production and perception) to keep distinct the increasing number of holistic signals needed to keep pace with an increasing message capacity. In another paper we have considered issues related to the formation of sound systems under constraints of this type (Lindblom, MacNeilage, & Studdert-Kennedy, 1984b). Although that may have been the selectional pressure, the question remains as to how our predecessors were able to respond with a concatenation strategy of the frame/content type. Our answer is that the form of the response was possible because we had already evolved a mode of organization that could be adapted for the purpose. That mode of organization may have been the one required for bimanual coordination. In particular, we have in mind the mode of organization whereby an object is held in the nonpreferred hand, serving as the frame, and manipulated by the preferred hand (providing content elements). Note, however, that this hypothesis does not commit us to the view that speech evolved from a manual gestural language. We are claiming that only a mode of organization of bimanual function was adapted for language use.

In addition to the obvious analogy, there are other reasons for suggesting bimanual coordination as a precursor to phonological organization. First, from our standpoint, attempting as we are to derive linguistic phenomena partly from motor constraints, bimanual coordination is an obvious candidate for a precursor, because, with the possible exception of speech, it is the complex serial voluntary action that man does best. In addition it is, without exception, the complex serial voluntary action that our nearest primate relatives do best. Second, the coincidence of the control of the preferred hand and of language in the same hemisphere in most humans, particularly language production, is consistent with the hypothesis. Note, in this context, that it is probably more appropriate to regard right hand preference as part of a specialization for *bimanual* coordination than as simply a unilateral specialization. There is evidence that left hemisphere damage affects the function of both hands whereas the same cannot be said for the right hemisphere (Kimura, 1979). Thirdly, the hypothesis that bimanual coordination is a precursor to phonological organization has the merit that it is consistent with the evolutionary principle of conservation of adaptive functions (Jerison, 1973). The principle states that functional adaptations tend to be conserved once they have evolved. Thus it is not necessary to fly in the face of this well-accepted tenet of evolutionary biology, as Chomsky (1968) and others do when they claim that language evolved de novo. The same consistency with the principle of conservation of adaptive functions can be claimed for the additional hypothesis that frame/content organization in morphology, evolved from frame content organization in phonology. (We think most people would accept the speculation that we had phonological organization before we acquired syntax.)

Pulling together the threads of this discussion, the frame/content hypothesis states that there is a three-stage sequence in the evolution of serial organization of language: (a) bimanual coordination, (b) phonological organization, and (c) morphological organization. In this development, evolution proceeds as it usually does, as a tinker, adapting available material to new needs (Jacob, 1977) rather than an engineer, making a new structure from parts specially designed from scratch, just for that purpose.

If we now consider current knowledge of brain-behavior relations, there seem to be two particular problems for the frame/content hypothesis. The first, is the possibility that some individuals control language and the preferred hand from different hemispheres. The second, is the claim that nonhuman primates do not have hand preferences like those of man.

The possibility that language and preferred hand control are in different hemispheres seems to pose a problem for any hypothesis that says that language had an evolutionary precursor in manual specialization. This would appear to be true in any case in which language control is in the hemisphere ipsilateral to the preferred hand, because direct control of the hand is contralateral. It is estimated that this is true of 1% of right handers and 60-70% of left handers (Corballis & Beale, 1983). As left handers constitute about 10% of the population, we are talking about 7-8% of the population. This population needs careful scrutiny. One possibility, that does seem to occur in some instances, is that the control information for skilled voluntary actions originates in a center ipsilateral to the hand, but is then sent across the corpus callosum to then be transmitted contralaterally in the usual way (e.g., Heilman, Coyle, Gonyea, & Geschwind, 1973). In addition, left handers as a population are thought to be slightly more likely to have some early medical problems that could affect localization of cerebral functions and dissociation of language and preferred hand control could occur for this reason. Another problem is that unless clinical studies give equally detailed information on language function, manual function, and lesion site it is difficult to establish instances of dissociation of language and hand control. We intend to look very carefully at this body of literature to see whether it poses problems for the frame/content hypothesis or not.

The second apparent problem—the problem of handedness in nonhuman primates—we have looked into (MacNeilage, Studdert-Kennedy, & Lindblom, 1984). The following paragraphs summarize the conclusions of our review. Efficient bimanual coordination may have first evolved in old world monkeys together with the truly opposable thumb (Napier, 1962). This development probably occurred several million years ago. Great apes are also quite capable of efficient bimanual coordination. If right-hand preference is associated with the evolution of bimanual coordination, as we believe it is, then these primates should show at least some trend toward right-handedness. But it is the virtually unanimous conclusion of a number of recent reviewers and participants in several recent published symposia that other primates are quite unlike humans in hand preference. The consensus seems to be that if preferences are shown, they are about equally frequent for left and right hands, but very task specific and often unstable. Secondly, there is considered to be virtually no evidence to sustain a conclusion that either old world monkeys or great apes have hemispheric specialization of function. To put it bluntly, the likelihood that various nonhuman primates have been capable of bimanual coordination for several million years without the evolution of either patterns of hand preference, or hemispheric specialization, does not appear favorable for the frame/content hypothesis. However, our review of this literature suggests that these negative conclusions are mistaken, or, at the very least premature (MacNeilage et al., 1984).

Almost all the work on primate handedness has been done on old world monkeys, so we will restrict the present discussion to this group. Most of the negative conclusions either come directly from the work of J. M. Warren (e.g., Warren & Nonneman, 1976) or are based on his work. Unfortunately there are three major problems with this work. First, he has typically used relatively young monkeys (under 2 years of age) and other studies suggest that hand preferences may not be fully developed in monkeys at this age. Consequently, it is not too surprising that he finds these monkeys to be inconsistent in hand preference from task to task and on repetitions of the same task. Second, he has used a relatively narrow range of tasks in terms of the necessity for the use of both hands and in terms of the level of complexity of the required manipulative movements. Third, and perhaps most important, his criterion for human-like handedness is that monkeys perform each individual act in multiact tasks with the same hand. In other words his criterion ignores the tendency toward bimanual coordination in humans performing multi-act tasks, in which different hands are favored for different acts. Then when he finds that monkeys also use different hands for different acts in multi-act tasks he declares them to be unlike humans.

Warren's use of the criterion that one hand must be used for all acts may have prevented him and perhaps others from noting that there is a definite pattern in what hand monkeys use for what act, that is like the pattern shown by humans in some respects, but not in others. This pattern is most clearly shown in a paper by Beck and Barton (1972) who have studied a far wider range of tasks than anyone else—17 different tasks. Like humans, the 10 monkeys in this study show a definite preference for the right hand for aspects of tasks that involve complex manipulation. The extreme example of this is shown in a task that required two embedded hasps to be opened by insertion of a single finger under each hasp [2IHC]. The median preference level for right hand actions for these two acts was 96.5%. But in general, all of the 9 manipulative movements studied showed an overall right hand preference.

The other part of the pattern is a tendency to favor the left hand to pick up the reward—in our terms a left hand preference for incentive grasping. The extreme example was a median left hand preference of 100% for incentive grasping in one of the embedded hasp tasks. In general, 14 of the incentive capture movements in the 17 tasks showed an overall left hand preference.

Two other types of task also reveal this trend towards a left hand preference for incentive capture movements. Ettlinger and his colleagues have found it in each of four studies of discrimination tasks in which the monkey obtains food by uncovering a foodwell under the correct one of two simultaneously presented stimuli (Ettlinger, 1961; Ettlinger & Moffett, 1964; Gautrin & Ettlinger, 1970; Milner, 1969). Secondly, three independent field studies have shown about a 2:1 preference of left over right hands in Japanese macaques in tasks that involve picking up food thrown on the ground (Itani, 1957; Itani, Tokuda, Furuya, Kano, & Shin, 1963; Tokuda, 1969). In a fourth field study a 2:1 preference ratio has been shown for the hand used to catch food in a group of smart monkeys who developed one-handed food-catching skills (Kawai, 1967).

The number of monkeys showing a left hand preference for incentive capture increases with age in free ranging monkeys. But in a number of experiments involving manipulation, an increasing preference for right hand even for the incentive capture act is observed with practice. These results lead us to the hypothesis that the predominant natural pattern of handedness in old world monkeys is a dichotomous one—a left hand preference for incentive capture movements and a right hand preference for fine manipulation. In nature, the variety of circumstances associated with manual grasping of food suggests that the left hand preference is for movements that require visual guidance in space because each movement is a relatively novel one for the animal. The left hand preference for incentive capture may also be observed in experiments because the expectation of food or the visual stimulation associated with food may put the animal into a left hand response mode even after the spatial contingencies of the situation have ceased to be novel to it.

On the other hand, under natural circumstances the right hand may more typically be used to manipulate an object already placed by the animal in a particular relatively stereotyped non-novel position with respect to the hand, often with the left hand. The serial effects of increasing right hand preferences observed in experimental tasks may result from some animals increasingly assimilating the stereotypy of the situation and moving to a mode of response which is more suited to stereotyped situations—a right hand response.

The right hand preference we propose for monkeys is obviously analogous to the one found in humans, and we suspect it is accompanied by a similar left hemisphere specialization. At first glance there seems to be no human analog to the left hand preference. But some studies have shown a left hand advantage for right handers in tasks that appear to involve a spatial component (e.g., Kimura & Vanderwolf, 1970). A French group (Guiard, Diaz, & Beaubaton, 1983) has shown that right handers are more accurate with their left hand in tasks that involve a single rapid movement to a visual target. In addition, Hampson and Kimura (1984) have shown a left hand preference in assembling blocks according to nonverbal principles, coexisting with a right hand preference for block assembly following verbal principles. Although it is difficult to see exactly what these tasks have in common, further consideration may suggest some relation between the role of the left hand in monkeys and humans. Perhaps humans and old world monkeys are separated by an evolutionary progression in which the importance of bimanual coordination, with its usually associated right hand preference, has so increased that it has preempted any propensity for left hand use in unimanual tasks under most normal circumstances.

We must now confront the consensus in the literature that old world monkeys do not have hemispheric specialization related to handedness (e.g., Warren, 1980). We believe this conclusion to be premature. Perhaps most importantly, our reanalysis of the monkey handedness literature leads us to suspect that the criteria for determining handedness in the relevant studies were usually inappropriate. Two additional problems with this work have been noted by Charles Hamilton (Hamilton, 1977; Hamilton & Vermeire, 1982). First there have been relatively few studies of tasks of a type that would reveal hemispheric specialization in man. Therefore, there is no reason to expect these tasks to be associated with hemispheric specialization in monkeys. Second, workers in this field find themselves in the uncomfortable position of trying to prove the null hypothesis. This is made especially difficult by a practical constraint that leads investigators to study only a small number of monkeys per experimental condition. Nevertheless in spite of all these problems there have been a few positive findings (e.g., Hamilton & Vermeire, 1982). We contend that with a better theory of handedness, an appropriate choice of task, and a large enough number of subjects, significant hand-hemisphere relations could be found.

Our conclusion that nonhuman primates may indeed possess significant hand preferences is presently specific to old world monkeys. In our opinion there is at present insufficient evidence to conclude one way or another about other taxa. However, as great apes are more closely related to man than are old world monkeys, more careful study may reveal that they too have hand preferences that are worthy of interest. This conclusion is encouraged by the fact that these animals are capable of efficient bimanual coordination, and by their possession of human-like patterns of hemispheric structural asymmetries which cannot be attributed to possession of human-like language (e.g., LeMay, Billig, & Geschwind, 1982).

We believe the frame/content hypothesis has a number of important implications. The emphasis on the importance of bimanual coordination may help to focus more attention on the evolution of bimanual coordination in other primates and the possibility that evolution of brain specialization may parallel evolution of this capacity across the entire primate order. The hypothesis leads to the prediction of universal patterns of speech errors at both phonological and morphological levels, although precise expectations for languages that differ markedly from English (e.g., agglutinative languages or languages with relatively free word order) remain to be developed. Some implications of the frame/content hypothesis for the acquisition of language are presented elsewhere (MacNeilage, 1984b). One important implication is that although there may be a natural propensity for a frame/content mode of organization, infants have to develop it from an initial mode in which frame and content are not separately available at either the phonological or the morphological levels. Thus certain well-known and apparently regressive discontinuities in phonological and morphological development, such as loss of progressive phonological idioms, and use of incorrect regular plural forms for previously correct irregular plurals (e.g., went ➧ goed or wented) might be taken as evidence for shifts towards a frame/content mode.

The claims that both left and right hemispheric specializations for manual functions may have already evolved in old world monkeys has the implication that human specializations may be superimposed on them. A different but well-known approach to human hemispheric specialization is to argue that the human hemispheres have species-specific specialization for *meta functions,* (i.e., generalized capacities that facilitate certain functions). Well known meta functions proposed for the left hemisphere are analytic and serial capacities and for the right, synthetic (holistic) and parallel processing capacities. We would assert that natural selection acts on functions not meta functions and consequently it is better to think in terms of a relatively specific functional adaptation resulting in a capacity for a variety of behaviors than to think of meta functions arising de novo in humans.

Some well-known effects of brain injury on language function are placed in an interesting perspective by the frame-content hypothesis. One finding is that patients who have lost the left hemisphere early in life typically have more problems with syntax and certain aspects of phonology such as rhyming than with semantic or lexical aspects of language (e.g., Dennis & Whitaker, 1977). This is consistent with the implication that the right hemisphere has a disadvantage in representing structural aspects of language independently of content elements (i.e., it lacks a propensity for frame/content organization). In addition, the two major syndromes resulting from damage to the left hemisphere, Broca's and Wernicke's aphasia, can readily be given a general characterization in frame/content terms. The agrammatism of Broca's aphasia can be characterized as a frame disorder at the morphological level, while the lexical choice problems, segmental paraphasias, and neologisms of Wernicke's aphasia suggest a content disorder at both morphological and phonological levels.

To sum up: The first problem we encounter in looking at speaking is the nonisomorphism paradox—the lack of a straightforward relation between underlying context-free linguistic units and context sensitive surface representations of these units. The elaborate patterns of surface adjustments that we observe are made at an extremely versatile motor control stage, beyond the stage of organization of message units. This output stage probably does not involve actions that are in principle different from those seen in goal-seeking behavior well down the phylogenetic scale. What we believe is new about language production is the high rate of output of different elements and the mode of organization of these elements. We suggest that a frame/content mode of organization exists at both levels of the dualistic system: the sound level, where vowel and consonant content elements are inserted into syllable frames; and the meaning level, where content word stems are inserted into syntactic frames. We suggest that the morphological mode may have evolved from the phonological mode, and the phonological mode may have evolved from an analogous mode of organization for bimanual coordination. In total then, we propose a three-stage evolution of language functions which is consistent with the principle of conservation of adaptive functions—serial organization of language arose by capitalizing on an existing adaptation rather than arising de novo in humans.

One issue that this hypothesis brings into sharp focus is the status of the evidence that language and the preferred hand can be controlled by different hemispheres in subjects with neurologically normal histories. Another issue is handedness in nonhuman primates. A reexamination of handedness studies of old world monkeys shows the tendency towards right hand preference for manipulation that would be expected from the hypothesis, and also a left hand preference, which may be a precursor to the right hemisphere specialization for spatial functions in man. The frame/content hypothesis has implications for a number of areas of inquiry, including nonhuman primate evolution, hemispheric specialization, internal organization of the left hemisphere in man, language acquisition, and cross language studies of errors. Finally, we hope that one beneficial function of the frame/content hypothesis, whether it is right or not, is to draw more attention to a biological approach to language, which might lead to a more unified view of language evolution, language development, language pathology, and normal language function than exists at present.

ACKNOWLEDGMENT

Preparation of this paper was supported by Grant No. NS 15336-05 from the Department of Health and Human Services, Public Health Service.

REFERENCES

Beck, C. H. M., & Barton, R. L. (1972). Deviation and laterality of hand preference in monkeys. *Cortex, 8*, 339–363.

Chomsky, N. (1968). *Language and mind.* New York: Harcourt Brace Jovanovich.

Cooper, F. S., Liberman, A. M., Harris, K. S. & Grubb, P. M. (1958). Some input output relations observed in experiments on the perception of speech. *Proceedings of the Second International Congress of Cybernetics,* Namur, Belgium.

Corballis, M. C., & Beale, I. T. (1983). *The ambivalent mind.* Chicago: Nelson Hall.

Dennis, M., & Whitaker, H. H. (1977). Hemispheric equipotentiality and language acquisition. In S. J. Segalowitz & F. A. Gruber (Eds.), *Language development and neurological theory.* New York: Academic Press.

Dingwall, W. D. (1979). The evolution of human communication systems. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics,* Vol. 4, New York: Academic Press.

Ettlinger, G. (1961). Lateral preferences in monkeys. *Behavior, 17*, 275–287.

Ettlinger, G., & Moffett, A. (1964). Lateral preferences in the monkey. *Nature, 204*, 606.

Fentress, J. C. (1983). Hierarchical motor control. In M. G. Studdert-Kennedy (Ed.), *The psychobiology of language* (pp. 40–61). Cambridge: MIT Press.

Garrett, M. (1975). The analysis of sentence production. In G. Bower (Ed.), *The psychology of learning and motivation,* Vol. 9. New York: Academic Press.

Gautrin, D., & Ettlinger, G. (1970). Lateral preferences in the monkey. *Cortex, 6*, 287–292.

Guiard, Y., Diaz, G., & Beaubaton, D. (1983). Left-hand advantage in right handers for spatial constant error: Preliminary evidence in a unimanual ballistic aimed movement. *Neuropsychologia, 21*, 111–115.

Hamilton, C. R. (1977). Investigations of perceptual and mnemonic lateralization in monkeys. In S. Harnad, R. W. Doty, J. Jaynes, L. Goldstein, & G. Krauthamer (Ed.), *Lateralization in the nervous system* (pp. 45–62). New York: Academic Press.

Hamilton, C. R., & Vermeire, B. A. (1982). Hemispheric differences in split-brain monkeys learning sequential comparisons. *Neuropsychologia, 20*, 691–698.

Hampson, E., & Kimura, D. (1984). Hand movement asymmetries during verbal and nonverbal tasks. *Canadian Journal of Psychology, 38*, 102–125.

Heilman, K. M., Coyle, J. M., Gonyea, E. F., & Geschwind, N. (1973). Apraxia and agraphia in a left hander. *Brain, 96*, 21–28.

Hockett, C. F., & Ascher, R. (1964). The human revolution. *Current Anthropology, 5*, 135–147.

Itani, J. (1957). Personality of Japanese monkeys. *Iden, 11*, 29–33.

Itani, J., Tokuda, K., Furuya, Y., Kano, K., & Shin, Y. (1963). The social construction of natural troops of Japanese monkeys in Takasakiyama. *Primates, 4*, 1–42.

Jacob, F. (1977). Evolution and tinkering. *Science, 196*, 1161–1166.

Jerison, H. J. (1973). *Evolution of the brain and intelligence.* New York: Academic Press.

Kawai, M. (1967). Catching behavior observed in the Koshima troop—a case of newly acquired behavior. *Primates, 8*, 181–186.

Kent, R. D. & Minifie, F. D. (1977). Coarticulation in recent speech production models. *Journal of Phonetics, 5*, 115–134.

Kimura, D. (1979). Neuromotor mechanisms in the evolution of human communication. In H. D. Steklis & M. J. Raleigh (Eds.), *Neurobiology of social communication in primates* (pp. 1–95). New York: Academic Press.

Kimura, D., & Vanderwolf, C. (1970). The relation between hand preference and the performance of individual finger movement by left and right hands. *Brain, 93*, 769–774.

LeMay, M., Billig, M. S., & Geschwind, N. (1982). Asymmetries of the brains and skulls of nonhuman primates. In E. Armstrong & D. Falk (Eds.), *Primate brain evolution: Methods and concepts.* New York: Plenum.

Lindblom, B. (1982). The interdisciplinary challenge of speech motor control. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control.* Oxford: Pergamon Press.

Lindblom, B., Lubker, J., & Gay, T. (1979). Formant frequencies of some fixed mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics, 7*, 147–161.

Lindblom, B., MacNeilage, P. F., & Studdert-Kennedy, M. G. (1984a). *Biological bases of spoken language.* New York: Academic Press.

Lindblom, B., MacNeilage, P. F., & Studdert-Kennedy, M. G. (1984b). Self organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie & Ö. Dahl (Eds.), *Explanations of linguistic universals.* Mouton: The Hague.

MacNeilage, P. F. (1970). Motor control of serial ordering of speech. *Psychological Review, 77*, 182–196.

MacNeilage, P. F. (1982). Speech production mechanisms in aphasia. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 43–60). Oxford: Pergamon Press.

MACNEILAGE, P. F. (in press-a). A comparison of serial ordering errors in speech and typing errors. In V. A. Fromkin (Ed.), *Linguistic phonetics.* New York: Academic Press.

MACNEILAGE, P. F. (in press-b). Handedness and the beginnings of speech. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors to infant speech.* London: MacMillan.

MACNEILAGE, P. F., STUDDERT-KENNEDY, M. G., & LINDBLOM, B. (in press). Primate handedness reconsidered. *Behavioral and Brain Sciences.*

MILLER, G. A. (1965). Some preliminaries to psycholinguistics. *American Psychologist, 20,* 15–20.

MILNER, A. D. (1969). Distribution of hand preferences in monkeys. *Neuropsychologia,* 7, 375–377.

NAPIER, J. R. (1962). The evolution of the hand. *Scientific American, 217,* 56–62.

PILBEAM, D. (1984). The descent of hominoids and hominids. *Scientific American, 20,* 84–96.

SHATTUCK-HUFNAGEL, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Lawrence Erlbaum Associates.

SHATTUCK-HUFNAGEL, S. R., & KLATT, D. H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior, 18,* 41–55.

TOKUDA, K. (1969). On the handedness of Japanese monkeys. *Primates, 10,* 41–46.

TREVARTHEN, C. (1978). Manipulative strategies of baboons and the origins of cerebral asymmetry. In M. Kinsbourne (Ed.), *Asymmetrical function of the brain* (pp. 329–391). Cambridge: Cambridge University Press.

WARREN, J. M. (1980). Handedness and laterality in humans and other animals. *Physiological Psychology, 8,* 351–359.

WARREN, J. M., & NONNEMAN, A. J. (1976). The search for cerebral dominance in monkeys. *Origins and evolution of language and speech. Annals of the New York Academy of Sciences* (pp. 732–744). New York: New York Academy of Sciences.

Chapter 5

ATTENTION, MOTOR CONTROL, AND AUTOMATICITY

George E. Stelmach
Barry G. Hughes

University of Wisconsin, Madison

Although speech can be considered a quite specialized form of motor behavior, it shares many characteristics with other actions, such as typewriting, handwriting, or playing a musical instrument. Many motor behaviors, including speech, involve multiple structures, require advance planning for proper execution, occur accurately and in real time, and often proceed without the conscious strategic control of all aspects of production. Speech and handwriting both reveal what Hebb (1949) referred to as motor equivalence (see also Stelmach & Diggles, 1982), the phenomenon whereby a single motor goal (i.e., the production of a particular utterance) can be achieved via quite variable specific muscular activity (e.g., with and without an object clenched between the teeth). Speech, along with other motor behaviors, are also subject to occasional "slips" (see Neumann, 1984; Norman 1981; Reason, 1977) where, for one reason or another, intended acts are carried out inappropriately, or perhaps not at all. Further, speech and other motor behaviors are apt, in certain circumstances, to be considered automatic. Because of such similarities, those concerned with the planning and production of speech should be interested in the relevant research literature on automation from other areas of motor behavior.

In this paper we address the issue of motor skill automation, focusing particularly on how it is characterized and defined in the psychological literature. To discuss the putative automaticity of any behavior, or its constituent processes, requires that consideration also be given to several other psychological phenomena such as consciousness, attention, intention, processing capacity, and motor control. All of these various constructs have, at different times and to different degrees, been associated with automaticity.

We discuss automaticity as it has been associated with two characteristics of human information processing and to determine the degree to which these characteristics might provide a convergent theory of automaticity. One is from the perspective of the characteristic functional structure and capacity of the human information processing system. We discuss the empirical evidence for processing limitations and the possibilities of the concurrent processing of different types of information (i.e., in attempts to listen to two information channels at the same time). The second perspective from which automaticity can be viewed concerns the role of intention in processing information and acting upon it, particularly, how intention might bear on the control of processing and performance. In reviewing these perspectives, we note that attention-demanding and automatic motor acts are viewed as complementary, and the complementarity assigned to attention and automation is often rather unidirectional. Attention models often invoke automaticity to account for nonattention-demanding tasks. A serious account of automaticity if it is to produce at all, can not simply be the default complement of attention-demanding acts.

Toward Operational Definitions

Automaticity essentially refers to processes usually involved in the perception of stimuli and production of responses which, for one reason or another, do not conform to widely accepted characteristics of cognitively dominated processes. That is, they are processes which are performed quite independently of, among other things, a person's current intentions, beliefs, or allocation of processing resources. An act might also be considered automatic if it is produced in the absence of a conscious awareness that it is taking place. Neumann(1984) following Posner (1978), has recently offered the following as the primary criteria of automaticity: that automatic tasks require no processing capacity, that they take place independently or beyond the control of the actor's intentions, and that they are not available to consciousness or introspection. The cognitive-motor literature contains substantial bodies of data on each of these aspects of automaticity. As Neumann (1984) noted, however, theorists do not always weigh these criteria equally in determination of whether or not an act is automatic. The consciousness criterion, for example, has only recently been promoted as a functionally useful distinction, due primarily to the work of Posner and his colleagues (e.g., Posner & Snyder, 1975).

LaBerge (1975) and Shiffrin, Damais, and Schneider (1981) have been primarily concerned with the capacity and avoidability criteria in their recent theoretical reviews of the nature of perceptual and cognitive processing. It is not yet clear whether this particular emphasis is warranted in considerations of cognitive-motor automaticity. Further, interpretations of necessity and sufficiency need to be considered. And as might well be expected in any area which considers automaticity in terms of intentions, consciousness, processing limitations and/or control divergent criteria, and simple binary distinctions are not likely to produce widely affirmed theoretical solutions. We will document some of the alternative positions in detail below. For now it should be recognized that there have been various other criteria proposed to distinguish, but not necessarily define the characteristics of automatic acts—rapidity and stereotypy of performance, for example, and that some theorists have come to view automaticity as an ideal or limiting instance of performance, a level which, if attainable at all, essentially lies beyond the domain of most information processing theories.

Structural and Functional Processing Limitations

The parallel development of information theory and computer sciences during the 1950s and the period that followed began to suggest to the psychologist a large number of theoretical and conceptual ideas about, and attractive metaphors for, cognitive processing. Through Information Theory it was possible to see the human as a limited transducer of information. For example, Miller (1956) found that the immediate span for unrelated alphanumeric characters, as well as categories or "chunks" of information, was only about seven (plus or minus two) items. It was also well-known that people have very severe limitations in dual task performance irrespective of whether these tasks are basically perceptual or motor in nature (see, in particular, Broadbent, 1958; Welford, 1952).

Following Broadbent's (1958) *Perception and Communication*, a rapid development of new experimental techniques permitted attempts to trace the information flow through the perceptual-cognitive-motor system. Much of the subsequent experimentation on information processing was directed to one of two lines of inquiry. The first related to which stages of processing were rate- or capacity-limited, and occupied much of information processing theory's first 15 years. The data now indicate that processing limitations can be shown to occur at many stages in a processing sequence: Broadbent (1958) himself held that it was early on, at the perceptual level, that a bottleneck took place; Keele (1973) demonstrated that a bottleneck could occur at a later response selection stage. It was the apparently protean nature of bottlenecks that eventually forced reconsideration of the idea that the problem was structural and not more functional and context-dependent.

The second set of questions—more characteristic of research in recent years—focused on how limited functions were controlled in order to prevent processing "overloads" or "bottlenecks." The original conception of attention was in terms of a limited-capacity central processor that was protected by a narrow information channel early in the processing sequence system (i.e., in the perceptual system). Although Broadbent's conception was shaped by the limited capacity processor notion, this proposed bottleneck in the perceptual system was not capacity-limited; it was based instead on structural limitations in single channel input storage, and "buffer" systems.

The structural models were soon challenged by a theoretical orientation, initiated by Moray (1967) and subsequently developed by Kahneman (1973), that was rooted in the assumption of a limited general processing capacity pool had to be competed for, tapped, and shared among the multitude of ongoing mental processes. Rather than using the structural bottleneck metaphor proposed by Broadbent (1958), a new, if equally mechanical, metaphor was presented: What Allport (1980a) has characterized as the limited power supply. According to the capacity models, processing "interference is nonspecific, and depends only on the (combined) demands of both the tasks" (Kahneman, 1973, p. 11). Clear demonstrations of decrements in processing performance with increases in processing load, in either single or dual-task conditions, at almost any "stage" from perceptual through memorial to motor processes, were taken to mean that there were not mere bottlenecks that either did or did not hinder processing traffic, but that there was only so much in the way of a functional processing resource that could be allocated. As Allport (1980) described this metaphor: "Once the supply is fully loaded, any more watts consumed by one part of the system means less for all the rest, regardless of what they want it for" (p. 116). In other words, Kahneman (1973) held that there was a finite general capacity which could not be overtapped; whenever it was fully engaged by processes whose demands summed to more than this capacity, performance would fall off. The nature of automaticity in these models was quite simple—Automaticity could be defined as any process which did not require processing capacity. It was, in other words, the complement of an attention-demanding process. Thus, whenever there occurred dual-task performance in which there was no mutual interference, either or both tasks were candidates for automaticity (Allport, 1980a).

There are now recognized to be a variety of reasons why this model is inadequate. Experimentation with dual-task paradigms suggests that mutual interference increases only whenever two or more tasks share common input (Allport, 1970; Allport, Antonis, & Reynolds, 1972; McLeod, 1977; Treisman & Davies, 1973) or output modalities. Second, other studies have shown that there is no mutual interference as measured by comparison with a single task performance (Allport et al., 1972; Shiffrin & Gardner, 1972). The third line of evidence comes from correlational time-sharing paradigms which show little or no correlation in time-sharing performance in a series of task combinations (Hawkins & Olbrich-Rodriguez, 1980). Together, this evidence points toward highly context-specific interference effects for which general capacity models cannot account. Allport (1980a) has been perhaps the strongest critic of these models, particularly because they can not adequately provide details of the mechanisms permitting dual-task performance, and because they do not adequately offer independent criteria for attention-demanding and automatic processes.

There are, in addition, some compelling data that call into question both the concept of an undifferentiated reservoir of processing capacity and with it the position that automatic processing is necessarily capacity-free. Some examples may serve as an illustration. It has been shown that when subjects are required to shift their visual attention at certain unpredictable moments while simultaneously performing other cognitive tasks the latency of the visual shift in attention is dependent on the type of concurrent task. Neumann (1984), for example, had subjects process visual information in one location and respond to a sudden visual target elsewhere. The interference, delays in attention shifting, in this task was quite substantial. Jonides (1981), on the other hand, found no such interference when the ongoing task was memorial in nature.

With a different approach, Posner and Boies (1971) had subjects manually respond to visually presented letters which came one second apart. The subjects' task was to press the right index finger if the letters were the same, the right middle finger if they were different. In addition, subjects had a secondary task: to press with the left index finger whenever they heard an auditory signal (termed a probe). The authors reported that if the probe occurred just prior to or coincident with the presentation of the second letter the latency in the left index finger's button press was very much delayed. At other onset times, the probe resulted in normal latencies. McLeod (1977) replicated this experiment but found that if he changed the secondary task so that it no longer required another finger press, but a vocal response instead, the interference effect disappeared. Data such as these, because they indicate that what is apparently general capacity-demanding in one instance need not be so in another, have served to cast some doubt as to the viability of the criterion that automatic processes are those which do not suffer or cause interference (Allport, 1980a; Neumann, in press; Stelmach & Hughes, 1983; Stelmach & Larish, 1980).

Automaticity As Obligatory Processing

Obligatory (mandatory or involuntary) processing is another criterion for automatic processing. However, it is a criterion quite distinct from the capacity-free criterion discussed above. First and foremost, this criterion bears on the nature of the control of processing; an automatic act, in this sense, is supposedly one which takes place independently of the subject's intention, task strategy or other cognitive state. It is, as Logan (1979, p. 205) noted, "controlled by the stimuli in the task environment," not by the subject. There are, of course, a number of senses in which processing can be said to be obligatory or beyond the preventability of intentions. Dismissing from consideration for the moment the apparent automaticity of reflexes, there are two different senses in which processing might be considered automatic from the standpoint of intentionality: one, the processes may occur without any intention on the part of the subject, or they may take place on the basis of some intention (such as a specific goal) but not conform to them (Neumann, 1984). The crucial question here, as for the issue of capacity demands, is whether there is evidence that such processing, in either sense, is indeed obligatory. The empirical test of automaticity is more direct in the case of this criterion: in essence, if it can be demonstrated that subjects attempt to avoid such processing, but fail, then it would seem that a firmer operational account of automaticity might emerge. Conversely, if it can be shown that even putatively obligatory processes are influenced by task demands, this criterion of automaticity might require major modification.

There are two senses in which processing could be considered unintentional: one, the sense in which a particular processing sequence is set in motion by a certain intention but which fails to result in the production of an act satisfying that intent; and two, that the processing cannot be halted once initiated (Neumann, 1984). Speech errors are widely held to be separate along these lines. Norman (1981) has discussed those verbal slips which can be distinguished along these lines. The Stroop phenomenon (e.g., Dyer, 1973) is one of the most frequently cited instances of "obligatory" processing. Prototypically, it is found that if a subject is presented with a series of letters printed in a certain color ink, his/her reaction latency in naming the color of the letters increases when the letters form a word, and increases even more when the word formed is that of a different color. The same pattern of latency changes has also been found in series of variant tasks involving other modalities, and stimulus characteristics. For example, Navon (1977) found similar Stroop-like interference effects when subjects were to identify the local letter features of a global pattern of a different letter.

Attempts to determine whether such interference results from obligatory processing or whether it can be altered by particular attention or task demands have tended to suggest that this criterion of automaticity also fails to generate unequivocal support. That the interference effects vary in a Stroop task depending on which aspects of the stimulus the subjects are attending is taken as such counterevidence (see Neumann, 1984). For example, the strongest case for obligatory processing occurs when the color to be named is presented in a different color ink (Posner & Snyder, 1975). However, interference can be systematically reduced by separating the spatial characteristics of the interfering stimulus dimensions, and with practice it may even diminish altogether.

What also emerges from consideration of capacity-free and obligatory processing is that it is often difficult to reconcile the two sets of data with respect to automaticity. For example, while the Stroop effect has been widely viewed as evidence of obligatory processing (and therefore as automatic by one criterion) it also demonstrates that its processing is not interference-free.

There is now a solid basis of evidence that a task can benefit by being primed by an earlier task. Priming, it is thought, reduces the number of processing steps involved in recognition. This point has been made recently (Navon & Gopher, 1979) on the basis of a substantial amount of data (e.g., Beller, 1971; Ells & Gotts, 1977; Kadesh, Riese, & Anisfeld, 1976; LaBerge, Van Gelder & Yellott, 1970; Meyer & Schvaneveldt, 1976; Pomerantz, Sager, & Stoever, 1977). However, the mere time savings induced by having to perform certain tasks only once rather than twice (or not at all), does not, without more detailed analysis of the particular processes hypothesized to be affected, explain automaticity with any parsimony. To offer such data as evidence of automatic processing merely widens the already broad range of criteria that can be applied to automaticity without offering any means of explaining its existence or development.

With the exception of several limited views of motor automaticity (e.g., Keele, 1973; Welford, 1976), most of the research on automatic information processing has come from research on the development of highly skilled but peculiarly cognitive acts (such as reading). Surprisingly, little research has been directed at the processes that underlie motor skill automaticity. For the most part, motor-skill automaticity has been equated with learning and practice; the overlearned highly practiced act is considered "automatic," or, to put it another way (Klein, 1978), practice not only makes perfect, it makes automatic (Norman, 1976). However, there are some who view automaticity as developing through a regimen quite distinct from that used in learning. The practice that occurs after criterion performance has been reached is often termed *overlearning*, a label LaBerge (1975) views as misleading "because it implies that a kind of strengthening is being applied to a process which, in most important respects, has already been completed" (p. 51). According to LaBerge, automaticity involves processes distinct from those involved in learning. To LaBerge (1975) and LaBerge and Samuels (1974) automatization is tantamount to the gradual elimination of attention, and automatic processing may occur only when attention is directed elsewhere. Take the example of reading: Before subjects have learned to read, their performance is characterized by slow, serial processing of individual letters. As learning progresses, the time it takes to serially perform these processes decreases. The key to automaticity in reading, however, is a fundamental shift in the nature of the processing, towards much faster, parallel processing of letters (i.e., processing of words) which access memorial representations without the involvement of attention. The appearance of automaticity may reflect a more sophisticated, more organized processing hierarchy.

According to Logan (1979), automatization might involve "a search for the combination of abilities . . . with the least investment of capacity" (pp. 204-205). This view seems to imply a fixed resource capacity of which attention has some form of executive control, and it contrasts with other views, which hold that out of numerous task-specific capacities, none of which is uniquely associated with attention, a nonexecutive control can emerge—a heterarchial control system perhaps—where different capacities can assume control in different situations (Allport et al., 1972; Treisman, 1969).

Insofar as motor behavior is concerned, the unitary, executive-based control idea has been too well criticized to avoid or even delay the search for viable alternative control mechanisms (Kelso, 1981; Stelmach & Diggles, 1980 for recent reviews). Abbs (1982), for example, has claimed that it has become increasingly difficult to view the neuromotor execution of speech as a series of specified descending motor commands coming from some executive locus, reflecting, in some direct manner, an underlying matrix of phonetic features. Rather, it appears that patterns of speech muscle activity may depend upon movement-to-movement peripheral conditions and adaptive modification of less specific descending commands. An executive basis for attention has also lost a lot of influence in the cognitive realm, and the continued search for, and development of, more parsimonious models is clearly warranted. Regardless of the particular experimental issue at hand—be it process automaticity or skill automaticity as products, or automatization as a process—it seems increasingly clear that some attempt needs to be made to clarify the issue of whether attention is at all a requirement in accounting for automaticity.

ATTENTION AND CONTROL OF ACTION

The previous sections, and other papers, (e.g., Allport, 1980a; Stelmach & Hughes, 1983; Stelmach & Larish, 1980) have documented the limitations of both the structural limitation and unitary capacity ideas, at least insofar as they have been tested by experimental data. These data are now forcing consideration of models that combine known structural and capacity limitations, into a number of "structure-specific reservoirs of processing resources" (Wickens, 1980, p. 241). A trend that may possibly accompany the development of multiple resource models is the shift away from any explanatory reliance on just one locus of control. When single processing channels or unitary capacity dictating processing control is dispersed, so too seems to be the locus of control.

The contemporary view that attentional processes, in some conscious way, oversee or control information processing can be traced to the work of Klein (1976, 1978); LaBerge (1975); Posner (1978) and Posner and Snyder (1975). Klein (1976) terms attention "a brain mechanism of limited capacity" (p. 271), or more specifically "an executive mechanism which is actively used to select the inputs, mental operations, and response processes necessary for the accomplishment of perceptual, cognitive, and motor skills" (p. 272). Although there have been frequent fulminations against proposing or searching for any one mechanism of attention (Neisser, 1976) and despite a decrease in the practice lately, there has remained a tendency to view attention as the unitary source of control.

Any account of attention as executive control, whether directed at limb or speech control, runs into numerous theoretical and experimental problems. We mentioned that general capacity pool notions have difficulty accounting for evidence of dual-task noninterference and of improved performance in a dual-task performance, so let us concentrate here specifically on theoretical problems and ambiguities implicit in attention-as-executive control views. The first is that in general capacity processing models it is often assumed that one of two states exists in the processing of information: on one hand there is the processing that is "activated under control of, and through attention by the subject" and on the other we have performance that is "activated without the necessity of active control or attention by the subject" (Shiffrin & Schneider, 1977, p. 156). The first type is "controlled," the second is "automatic," and apparently only one of these two forms can be applied to any performance or process at any one time.

Such a distinction seems to require qualification on at least two points in order for an operational distinction to become clear. First of all, given the complementarity of controlled and automatic processing (Shiffrin & Schneider, 1977), one should seek to determine whether automatization is involved or whether there is some distinct change in control, a unique and independent process, that permits the automatic nature to be manifest, or whether it is some combination of both. Although there have been numerous distinctions made between controlled and automatic processes, (e.g., slower versus faster, more serial than parallel, more and less easily disturbed, etc.) how the shift between one mode and the other is brought about is never clear. This could involve an attention withdrawal but it could also be a process unique to learning; and it could also be an occurrence of both.

The dilemma can be exemplified another way: When the performance of one of two tasks is seen to deteriorate when they are performed together it is practically impossible to tell whether this is due to one task demanding more of a limited supply of resources or whether it is because it is being offered less (Navon & Gopher, 1979). The literature is not clear on this distinction, nor is there a readily apparent means to determine the cause and effect, although such a test would be of major consequence in any theory of attention or automaticity.

This brings us to the second of the qualifications that seem requisite for a coherent account of the controlled/automatic distinction: It does not seem sufficient to merely propose a dichotomy without also hinting at either the mechanisms involved or the contexts and criteria under which such a distinction could be empirically assessed. A major objective, therefore, must lie in determining the mechanism or principle behind such a dichotomy. Otherwise, as Allport (1980a) recognized, the dichotomy offers only "an anodyne heuristic" which does not promote further questioning and examination. As Allport (1980a) recently put it:

> Whenever two simultaneous task-demands conflict with one another, we have a comfortable, ready-made explanation: both must be competing for limited, general-purpose capacity (for 'attention'). When a pair of tasks is found that *can* be performed at the same time, independently . . . that too is easily explained. One or both of the tasks . . . must be "automatic." (p. 121).

Before dealing in more detail with the repercussions for automaticity of a shift in theoretical forms from concerns about when and under what conditions automaticity finally develops, to concerns about the instances where action is normally "unattended," it is opportune that we recognize another significant trend to which this model and other accounts (e.g., Stelmach & Hughes, 1983; Stelmach & Larish, 1980) have pointed. This is that the heretofore traditional context in which automation has been framed—processing stages—is now being reconsidered in terms of tasks-as-wholes. The idea that certain processes (or small groups of processes) could either be attention-demanding or automatic is, as we have attempted to document, turning out to be an ineffectual explanation of the nature of automation, if only because automaticity has been seen as nothing but the complement of attention-demanding processes. By the reasoning that the sequential series of processes are each either demanding (of structure or resource) or automatic, it would seem reasonable to assume that for an act as a whole to be considered nondemanding, each component process—from "perceptual" through "motor"—must be automatic. This is an extremely stringent requirement, and one that, to our knowledge, has never been demonstrably achieved. Yet, the literature (and everyday conversation) abounds with examples of "automatic" acts.

Although we have not, to this point, emphasized the theoretical parallels between attention, automaticity and learning, Norman and Shallice (1980) suggest the possibility that motor learning may reflect not so much an acquisition of skilled subroutines, as a shift in the necessary control of the system's many "degrees of freedom (Stelmach & Diggles, 1980). Bernstein (1967) characterized the degrees of freedom problem" as the theorist's problem of explaining how the many degrees of freedom of the body in terms of muscles, joints, and limb segments are regulated in the course of a motor act. Suppose that a subject learns a complex series of finger movements with the right hand, perhaps a rhythmic sequence of some sort. This sequence is stereotypical from trial to trial and, at the height of learning, could be performed "automatically;" that is, for example, without any detriment in the presence of a concurrent motor act. By the traditional view, a motor subroutine has been acquired, and represented centrally. This view does not adequately explain why there is not a perfect (or even moderately successful) transfer of this "subroutine" to another "output device" (e.g., the left hand). When transferred to the other hand, the primary task rapidly loses its "automaticity," yet the central representation is (presumably) one and the same. Possibly it is the control of the degrees of freedom of the limb that has changed in some way in the transfer condition. These degrees of freedom have not been adequately controlled, and when learning occurs, it seems quite possible that control of, or more accurately, a reduction of, the degrees of freedom is taking place.

There must be some way in which many degrees of freedom can be automatically regulated through the control of a few. In more general terms, it might be suggested that the system responsible for action organizes itself in order to make the action(s) more simple, or more neuroeconomically efficient. With practice—and this practice would need to be, at some nontrivial level, consistent (as Shiffrin & Schneider, 1977, recommended)—reduction in the degrees of freedom of the output devices occurs. This is not to deny that a different level of acquisition of action patterns is occurring; it is suggestive of the way automaticity of action represents a control and not just an attention problem, however. This approach, introduced in some detail by Fowler and Turvey (1978), maintains that skill learning represents the control of numerous action subsystems to the point where perhaps only the intent to achieve a goal is sufficient at any "high level" of the control system; the details of specification, subject as they are to peculiar context contingencies, are left to "lower levels." In terms of attention, the initiation of action may require attention (in the experimental sense of intent or will) with the actual details of execution taking on an automatic appearance.

Automatization, in these briefly outlined terms, may thus reflect in part, the subject's discovery of an optimum self-organizational strategy, in terms suggested by Navon and Gopher (1979), Norman and Shallice (1980), and Fowler and Turvey (1978). Emerging models of attention are not only better able to account for the data that have long plagued purely structural or general capacity models, but more importantly, they have the potential to become important rallying points around which information processing theorists can finally come to grips with a dynamic movement (Stelmach & Diggles, 1980, 1982).

CONCLUSION

Much too often areas of science each having a strong relation to each other, function quite independently without much cross fertilization, yet when one delves into both literatures it is often enlightening to see their similarities in issues, methods, and concepts. In this paper, we discussed the various criteria of automaticity and its long-time associates attention and intention, in the hope that those interested in the planning and production of speech will become more aware of how such concepts have been studied in other motor behaviors and what the contemporary issues are. In this process, we documented the complexities involved in establishing the defining attributes of task automaticity and argued, as others have, that automaticity cannot be adequately understood solely on the basis of understanding attention or processing resources. It was also noted that some of the research in automaticity has come from research on the development of highly skilled acts and has been equated with learning and practice. For too long automaticity of processing has been used as a characterization that deals with those processes that do not seem to fit the description of being either capacity-demanding, intentional, or controlled.

We also pointed out that when commonly used criteria for automaticity are applied to processing stages assumed to occur in motor behaviors such as speech, it appears that most acts in total do not unequivocally qualify as automatic. It remains to be determined what functional automatism means in a control context. It may mean abandoning the unitary, executive-based control idea in favor of alternative control models. Although there are many distinctions between controlled and automatic processes, how the transition develops has remained quite mysterious. We speculated that automatization may reflect, in part, a discovery of an optimum self-organization strategy for perception, cognition and action, and suggested that this collective perspective is an important basic position. Although the study of motor skill automaticity must nevertheless be directed towards a viable theory that includes and corresponds with contemporary theories of both attention and control of motor behavior, it seems imperative that automatism cannot continue to be considered two ways—one for attention, one for control—if a common account of its characteristics is desired.

REFERENCES

Abbs, J. A. (1982). Speech motor system perspective on nervous system control variables. *The Behavioral and Brain Sciences, 5,* 541–542.

Allport, D. A. (1980a). Attention and performance. In G. L. Claxton (Ed.), *New directions in cognitive psychology.* London: Routledge.

Allport, D. A. (1980b). Patterns and actions: Cognitive mechanisms are content-specific. In G. L. Claxton (Ed.), *New directions in cognitive psychology.* London: Routledge.

Allport, D. A. (1970). Parallel encoding within and between elementary stimulus dimensions. *Perception and Psychophysics, 10,* 104–108.

Allport, D. A. Antonis, B., & Reynolds, P., (1972). On the division of attention: A disproof of the single-channel hypothesis. *Quarterly Journal of Experimental Psychology, 24,* 225–235.

Beller, H. K. (1971). Priming: Effects of advance information on matching. *Journal of Experimental Psychology, 87,* 176–182.

Bernstein, O. W. (1967). *The co-ordination and regulation of movements.* Oxford, England: Pergamon Press.

Broadbent, D. E. (1958). *Perception and communication:* London: Pergamon.

Dyer, F. N. (1973). The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes. *Memory and Cognition, 1,* 106–120.

Ells, J. G., & Gotts, G. H. (1977). Serial RT as a function of the nature of repeated events. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 234–242.

Fowler, C. A., & Turvey, M. T. (1978). Skill acquisition: An event approach with special reference to searching for the optimum of a function of several variables. In G. E. Stelmach (Ed.), *Information processing in motor control and learning.* New York: Academic.

Hawkins, H. L. & Olbrich-Rodriguez, E. (1980). *Is time-sharing a general ability?* Unpublished manuscript.

Hebb, D. O. (1949). *The organization of behavior.* New York: John Wiley & Sons.

Jonides, J. (1981). *Voluntary versus automatic control over the mind's eye's movements.* Tech. Rep. No. 15, University of Michigan-University of Chicago.

Kadesh, I., Riese, M., & Anisfeld, M. (1976). Dichotic listening in the study of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 15,* 213–225.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping timesharing II. Auditory secondary task. *Acta Psychologica, 40,* 343–362.

Keele, S. W. (1973). *Attention and human performance.* Pacific Palisades, CA: Goodyear.

Kelso, J. A. S. (1981). Contrasting perspectives on order and regulation in movement. In J. Long and A. Baddeley (Eds.), *Attention and performance IX.* Hillsdale, NJ: Erlbaum.

Klein, R. M. (1976). Attention and movement. In G. E. Stelmach (Ed.), *Motor control: Issues and trends.* New York: Academic.

Klein, R. M. (1978). Automatic and strategic processes in skilled performance. In G. C. Roberts & K. M. Newell (Eds.), *Psychology of motor behavior and sport.* Champaign, IL: Human Kinetics.

LaBerge, D. (1975). Acquistion of automatic processing in perceptual and associative learning. In P. M. A. Rabbitt and S. Dornic (Eds.), *Attention and performance V.* London: Academic.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293–323.

LaBerge, D., & Van Gelder, P., & Yellott, J. (1970). A cueing technique in CRT. *Perception and Psychophysics, 7,* 57–62.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 189–207.

McLeod, P. (1977). A dual-task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology, 29,* 651–667.

Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science, 192,* 27–33.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica, 27,* 84–92.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review, 86,* 214–255.

Neisser, U. (1976). *Cognition and reality.* San Francisco: Freeman.

Neumann, O. (1984). On controlled and automatic processes. In W. Prinz & A. Sanders (Eds.), *Cognition and motor processes* (pp. 225–295). New York: Springer.

Norman, D. A. (1976). *Memory and attention* (2nd ed.), New York: Wiley.

Norman, D. A. (1981). Categorization of actions slips. *Psychological Review, 88,* 1–15.

Norman, D. A. & Shallice, T. (1980). *Attention to action: Willed and automatic control of behavior.* CHIP Tech. Rep. No. 99.

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 422–435.

POSNER, M. I. (1978). Mental chronometry and the problem of consciousness. In R. Klein & P. Jusczyk (Eds.), *Structure of thought: Essays in honor of D. O. Hebb.* Hillsdale, NJ: Erlbaum.

POSNER, M. I., & BOIES, S. J. (1971). Components of attention. *Psychological Review, 78,* 391–408.

POSNER, M. I., & SNYDER, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The loyal symposium.* Potomac, MD: Erlbaum.

REASON, J. (1977). Actions not as planned: The price of automation. In G. Underwood & R. Stevens (Eds.), *Aspects of conscious.* New York: Academic Press.

SHIFFRIN, R. M., & GARDNER, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology, 93,* 72–82.

SHIFFRIN, R. M., DAMAIS, S. T., & SCHNEIDER, W. (1981). "Characteristics of automatism." In J. Long & A. Baddely (Eds.), *Attention and performance IX.* Hillsdale, NJ: Laurence Erlbaum.

SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

STELMACH, G. E., & DIGGLES, V. A. (1982)..Control theories in motor behavior. *Acta Psychologica, 50,* 83–105.

STELMACH, G. E., & DIGGLES, V. A. (1982). *The Behavioral and Brain Sciences, 5,* 566.

STELMACH, G. E., & HUGHES, B. G. (1983). Does motor skill automation require a theory of attention? In R. A. Magill (Ed.), *Memory and control of action (pp. 67–91). Amsterdam: North Holland.*

STELMACH, G. E., & LARISH, D. D. (1980). A new perspective on motor skill automation. *Research Quarterly, 51,* 141–157.

TREISMAN, A. M. (1969). Strategies and models of selective attention. *Psychological Review, 76,* 282–299.

TREISMAN, A. M., & DAVIES,, A. (1973). Divided attention to ear and eye. In S. Kornblum (Ed.), *Attention and performance IV.* New York: Academic.

WELFORD, A. T. (1976). *Skilled performance: Perceptual and motor skills,* Glenview, IL: Scott, Foresman.

WELFORD, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance: A review and a theory. *British Journal of Psychology, 43,* 2–19.

WICKENS, C. D. (1980). The structure of processing resources. In R. Nickerson (Ed.), *Attention and performance VIII.* New York: Erlbaum.

Chapter 6

DEVELOPING AND DISORDERED SPEECH: STRATEGIES FOR ORGANIZATION

RAY D. KENT

University of Wisconsin, Madison

> It must have come by inspiration. A thousand, nay, a million children could not invent a language. While the organs are pliable, there is not understanding enough to form a language; by the time that there is understanding enough, the organs are become stiff . . . (Boswell, 1951, p. 1,044).

These comments on the origin of language may, with minor wording changes, express the basic issue of this conference. How is it that children discover a language; to harness their articulators to the demands of linguistic structure and social communication; to speak in essentially the same way as other children but yet to differ from them in so many subtle features of speech?

The two fundamental questions that confront us are (a) What is the organization or representation of speech information in the brain of the speaker/listener? and (b) How does this organization come about? It seems that traditional answers to these two questions preserve their separateness. On the one hand, theories and data are offered to describe the organization of speech, and on the other hand, theories and data are offered to explain how this organization comes into being in both normally developing children and in children with communication disorders. A primary thesis of this paper is that the process of organization is not necessarily distinct from the product of organization. Indeed, systems that are termed self-organizing or auto-regulatory (Guntern, 1982; Prigogine & Stengers, 1984) are understood in terms of a capacity to create and transform order. I propose that speech acquisition can be understood as a self-organizing process. But before I explain this idea further, it is helpful to review briefly alternative theories on the acquisition of speech and language.

THEORETICAL BACKGROUND: UNIVERSALIST ACQUISITION THEORIES

Macken and Ferguson (1981) described three universalist acquisition theories that have dominated research on child phonology for the last 15 years. These theories are summarized below so that points of comparison and contrast with a self-organizing system view of phonological development can be offered.

1. Jakobson's structuralist theory (Jakobson, 1941). Basically this theory assumes a "universal hierarchy of structural laws that determine the inventory of phonemic systems and the relative frequency, combinatorial distribution, and assimilatory power of particular phonemes" (Macken & Ferguson, 1981, p. 112). Examples of predictions based on this theory are: (a) order of acquisition of minimal inventories of vowels and consonants (e.g., the order /p/ before /t/, /t/ before /m/, and before /n/); (b) order of emergence of sound classes (e.g., stops before nasals, nasals before fricatives, and fricatives before liquids); and (c) order of acquisition based on features or combinations of features (e.g., front consonants are acquired before back consonants).

2. Universalist/nativist theory. This theory derives in large part from the work of Chomsky and has been applied to developmental phonology by Smith in particular. Smith (1973, 1975) identified four universal tendencies: (a) toward consonant and vowel harmony, (b) toward cluster reduction, (c) toward systematic simplification, largely through deletion and substitution, and (d) grammatical simplification.

3. Natural phonology. This theory, identified primarily with Stampe (Stampe, 1969, Donegan & Stampe, 1979), explains phonological acquisition with respect to universal innate natural processes. Macken and Ferguson (1981) wrote about this theory as follows:

> Stampe's basic assumption is that the phonological system of a language is the residue of a universal system of processes, governed by forces implicit in human articulation and perception. During acquisition, those processes not applicable to the particular language being learned are constrained by the mechanisms of suppression, limitation, and ordering. (pp. 112-113)

Examples of processes are fronting of consonants, reduction of consonant clusters, deletion of unstressed syllables, stopping of consonants, and final consonant devoicing.

Each of these theories has been criticized on both empirical and theoretical grounds (Ferguson & Garnica, 1975; Kiparsky & Menn, 1977; Macken & Ferguson, 1981; Menn, 1981; Olmsted, 1971). The major criticisms have been that predicted orders of acquisition have been violated, that acquisition is not adequately explained as a process of successive acquisition of phonetic oppositions, that the theories ignore cognitive developments, and that the theories generally predict a gradual convergence on, or progress toward, an adult system when in fact phonological acquisition is characterized by both regression and overgeneralization.

These criticisms and reconsiderations of the process of phonological development have led to a different kind of theory, the principles of which are reviewed by Macken and Ferguson:

> These data document the existence of significant, widespread individual differences between children acquiring the same language and show that the acquisition process, in certain key respects, is not a linear progression of unfolding abilities—as assumed by the universalist model. The emerging model recognizes several types of learning (e.g., accretion and tuning) but emphasizes the cognitive aspects of acquisition. In this view, the child is an active seeker and user of linguistic information who forms hypotheses on the basis of input data, tests and revises these hypotheses, and constructs more complex systems (or grammars) out of earlier, simpler ones. The shift is away from a deterministic linguistic model toward a flexible model that accommodates variation in development by acknowledging the active role of the child, the diversity of input, and the variety of solutions possible. (p. 114)

However, some aspects of the "older" theories may still have explanatory value. Menn (1981) observed, "But the vision shared by Stampe (1969) and Jakobson (1941) has a core that still is accepted completely by the problem-solving theory: the inarticulate child gradually learning to overcome some built-in tendencies of her articulatory apparatus as she attempts to say what she hears" (pp. 134-135).

Speech Acquisition as a Self-Organizing Process

My contention is that the problem-solving or interactionist theories of phonologic development are a type of auto-organization (or self organization). Basically, self-organizing systems are systems that have the capacity to generate, maintain, or transform order (Guntern, 1982; Prigogine & Stengers, 1984). Speech and language development in the young child can be understood as a self-organizing process constrained by the child's physical and biosocial environments, the genetic program, and ontogenetic learning. This theory incorporates the desirable features of older theories (particularly the considerations of biological maturation which determine an infant's capacity to hear or produce sounds) but also embraces the concept of the child as an active, cognitive agent in learning speech and language. The self-organizing theory accommodates regression and overgeneralization as consequences of maintaining and transforming order in the face of increasingly complex structures of information to be represented.

Self-organization is discussed in almost every contemporary field of science (Jantsch, 1979) as well as the history of science itself (Nalimov, 1981). Self-organizing processes may determine many aspects of development, beginning with embryogenesis. Couly (1982) speculated that vertebrate embryo development is controlled by morphological dynamics resulting from flotation in a liquid medium, giving rise to pentameric symmetry. The concept of self-organization also has influenced the field of ethology. For example, Fentress (1981) remarked that the "dynamics of interactive and self-organizing systems in behavior suggest that similar properties of order may operate in both the short and long (development) time frames" (p. 365, emphasis in original). The theory of self-organizing systems may help to explain the development of infants at risk. Woodson (1983) raised the possibility that "failure of measures of newborn behavior to predict developmental outcome may reflect the role of newborn behavior in the resolution of perinatal complications which would otherwise have resulted in subsequent deficits." Thus, correlation studies may fail to identify adaptive features of behavior that can resolve pathologic conditions. In this perspective, behavioral change is the appropriate focus of developmental study.

Some ideas highly similar to concepts of auto-organization have appeared in the literature on speech and language development. Slobin (1973), in discussing the cognitive prerequisites for the development of grammar, stated that "the child brings certain operating principles to bear on the task of learning to speak, regardless of the peculiarities of the particular language he is exposed to" (p. 176). These operating principles can be considered as "self-instruction" or "general heuristics" for organizing and storing language. Elbers (1982) concluded from a phonetic analysis of babbling that development of babbling is governed largely by two "operating principles," variation and combination, that qualify as general heuristics which have not been learned. Elbers proposes a cognitive continuity theory "which views repetitive babbling as a continuous and largely self-directed process of exploration, during which the infant uses certain operating principles for constructing his own springboard to speech" (p. 61).

Operating principles also are mentioned by Branigan (1976) in accounting for early word formation in a child studied between the ages of 16–21 months. Branigan identified a "universal operating principle" according to which a child concentrates on one aspect of a domain at a time (e.g., producing open monosyllables). This operating principle worked hand in hand with a "universal constraint" which restricted the selection of adjacent segments so that they would be maximally distinct.

An interesting property of auto-organizing systems is that an "initial kick" (directed vector) can be amplified to produce new order, structure and organization (Maruyama, 1960). Thus, in Guntern's formulation of morphogeneration, order can be created out of a random process (as when a condensation nucleus in saturated humid air creates a cloud) or an organized process (as when soldiers fall into well-ordered columns at the command of an officer). Thus, it is plausible that a young child might produce ordered information structures concerning erroneous conclusions or random events, although these structures may not be long-lived. Evidence that children learn their own errors during language development has been reported by Butler, Platt, and MacWhinney (1983). Such "erroneous auto-input" presumably comes about as children discover certain incorrect solutions to language structure problems. These errors may dominate over correct forms that the child hears in adult language, until such time as the errors must be reconciled with increasing competence in language structure. Morphogeneration, as an order-creating process, is complemented by the processes of morphotransformation (in which existing order is transformed into an order of higher, lesser, or equal complexity) and morpholysis (in which existing order is destroyed). All three of these processes can contribute to the auto-organization of language. One kind of order may be maintained even as another is transformed and still another is destroyed. Thus, the overall developmental process is continuous but some behavioral structures may disappear or assume altered form.

For example, Labov and Labov's (1978) account of growth and decay of phonetic development in their daughter's early word productions may be an instance of the transformation of order. Although phonetic records may indicate patterns of creation and dissolution "we have seen words appear and disappear, phonetic elements integrated into words and then abandoned, phonological contrasts constructed and then neglected" (Labov & Labov, 1978, p. 843), the child must preserve (transform) order or else convergence on a language system would never happen. Order from an apparently abandoned stage of phonetic development must be carried over to the next stage (albeit in an altered form) or each new attempt at a phonetic system would begin with nothing. Thus, the "progressive idiom" (Moskowitz, 1973) is an illustration of the transformation of order in which an apparently precocious utterance may undergo deterioration relative to the adult model as the child's phonologic system changes. Although the progressive idiom may appear to regress (e.g., Hildegard's change from [prɪtɪ] to [bɪdɪ] for the word *pretty*; Leopold, 1947), this change is simply one adjustment in a transformation of order.

Regressive forms and overgeneralizations present questions regarding the management of speech and language disorders in children: How is a clinician or language teacher to know when an ostensible reversal in a child's language is actually a sign of progress? If regression and overgeneralization are a matter of course in the normal development of language, then should we not also expect these features to be a part of language development in children with language impairments? If so, how can the clinician or teacher recognize these features as signs of progress rather than as failures of intervention? Of course, not all regressions or overgeneralizations are necessarily desirable, but how can the desirable ones be distinguished from those that are not?

Thus, it appears that phonologic development does not proceed by a process of accretion, in which elements are gradually added to a fixed phonologic base, but rather by a process more akin to hypothesis testing, in which an entire system may be overthrown or revised when it cannot meet the requirements of linguistic growth. Moreover, it seems that the phonemic contrasts favored by most linguists are not suitable for the description of early phonologic development. It has been proposed that in early stages of speech development, the units of phonologic contrast are not phonemes, but syllables (Moskowitz, 1973) or words (Ferguson, 1977; Leonard et al., 1980). Several reports indicate that early words are not strictly phonemically principled but rather are quite fluid in their phonetic structure (Ferguson & Farwell, 1975; Leonard et al., 1980; Roberts, 1979). I have proposed elsewhere (Kent, 1981a) that,

> motor control [or phonologic organization] that is adapted to the production of phonetic segments in a variety of phonetic contexts in different words perhaps comes about as the child discards the principle of preparing word-sized motor sequences for each word in his or her lexicon. That is, the child is forced to a segmental (phonetic) motor organization through sheer force of economy and manageability. (p. 179)

As a segmental awareness develops, the child enters a phase in which he/she is almost religious about phonetic segments; that is, the child's phonetic patterns are not as highly coarticulated as those of the adult (Kent, 1981b, 1983; Kuehn & Tomblin, 1974; Thompson & Hixon, 1979). After concatenative facility is established, this segmental religiosity gives way to the rapid, highly overlapped articulatory pattern typical of adult speech. Incidentally, we (Kent, Lippmann, & Osberger, 1980) have observed in the speech of some deaf adolescents a restriction on coarticulation similar to that seen in normal-hearing young children. This result is depicted in Figure 1, which shows formant patterns for the word *box*. Notice that particularly for the slower talkers represented at the right, F2 raising (anticipation of the velar consonant) is not pronounced during the vowel segment.

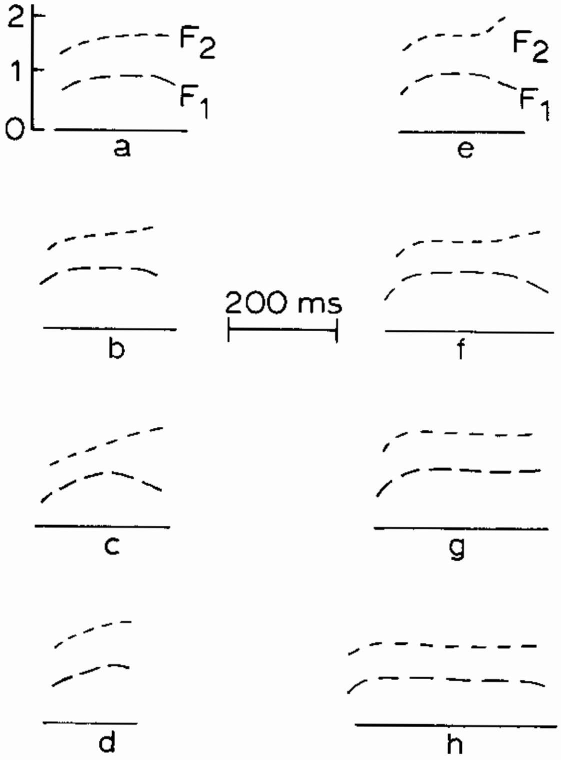


FIGURE 1. *F*1 and *F*2 formant-frequency trajectories for the word *box* produced by eight hearing-impaired children. Patterns *a-d* approximate the results expected for normal-hearing adults: *F*2 frequency increases nearly continuously during the syllable nucleus, apparently reflecting the anticipatory tongue raising for [*k*]. Patterns *e-h* are more typical of young children in that the *F*2 frequency stabilizes during the syllable nucleus. Patterns *a-d* reflect stronger coarticulatory effects than do patterns *e-h*. Frequency in kHz is scaled on the ordinate.

Furthermore, reorganization as a developmental principle is not restricted to the phonology. Cromer's (1983) research showed that lexical acquisition is not a gradual process of addition or differentiation but rather a reorganizational process. The acquisition of language by a child is not so much a matter of adding new pages of information to a loose-leaf binder as it is a rewriting and overwriting to produce an expanded text.

HYPOTHESIZED PROCESSES IN SPEECH ACQUISITION

The self-organizing theory can be given a more concrete form in terms of hypothesized processes that play a role in early speech development. Evidence of organizational processes that may lead to speech development can be seen well before the infant produces well-formed sound patterns that are recognized as early words. Elbers (1982) reported that an infant's babbling progresses through four major stages, beginning with single babbles (such as syllable /ba/ produced in isolation), proceeding to repetitions of a basic sound pattern (so-called reduplicated babbling; /ba ba ba ba/), leading next to alternations or concatenations of contrasting sound patterns (/ba ba da da di de/), and culminating in a stage of jargon babbling, which is a complex mix of the earlier forms. During the same developmental period, the infant begins to babble with patterns of rhythm and stress similar to those of adult language.

Several observers have noted that intonation is one of the earliest aspects of speech to develop in the infant. In reviewing the evidence, Cruttenden (1982, p. 116) concluded that (a) "The use of some of the forms of intonation is present from a very early age (almost certainly during the babbling stage)" and (b) "A very restricted use of intonation to convey some of the meanings conveyed by intonation in the adult language is also present from an early age (for *some* children as early as late babbling)."

Data reported by Stern, Spieker, & MacKain. (1982) indicate that "the intonational envelope itself is among the earliest and most basic units of interpersonal signaling in the auditory domain"(p. 734). Analyses of the utterances of six mothers to their 2-, 4- and 6-month-old infants showed that two intonation contours were context specific for all mothers. The first of these, rising contours, was observed when the mother was attempting to establish eye contact with an infant who was not attending visually. Secondly, sinusoidal and bell-shaped contours were observed when the mother was attempting to maintain, rather than re-establish, the infant's positive affect and gaze (The major types of intonation contours observed by Stern et al. are illustrated in Figure 2).

Somewhat different intonational devices were described by Bruner (1981) for two mother-infant dyads. When the child was not attending (i.e., not in eye contact with the mother), she used "a highly routinized attentional vocative," which took the form of a sharp rise-fall (triangular) contour or a square-wave (held-rise-held-fall-held) contour. After these intonational devices had been established as a means of securing the infant's attention (by about 5 months of age) they were expanded to include other intonation contours and a variety of verbal content. The management of attention, once the mother had established it, was accomplished differently. For example, one mother introduced objects to her already attending infant by presenting the object and vocalizing with a sinusoidal intonation.

Bruner (1981) cites J. Lyons' (1977) description of these attentional vocatives as "undifferentiated deictics" that indicate to the infant that there is something to attend to in the context or environment. Bruner also cites work by Ryan (1976) in this regard, commenting in particular on the rapidity with which these devices come to play a role in alerting the child to a possible shift in attention. Ryan observed that mothers of 12-month-old infants were especially likely to use a rising intonation to their infants when they were attending to a toy other than the one the infant was playing with. The infant, in turn, was likely to respond to this intonation by attending to the toy the mother was holding.

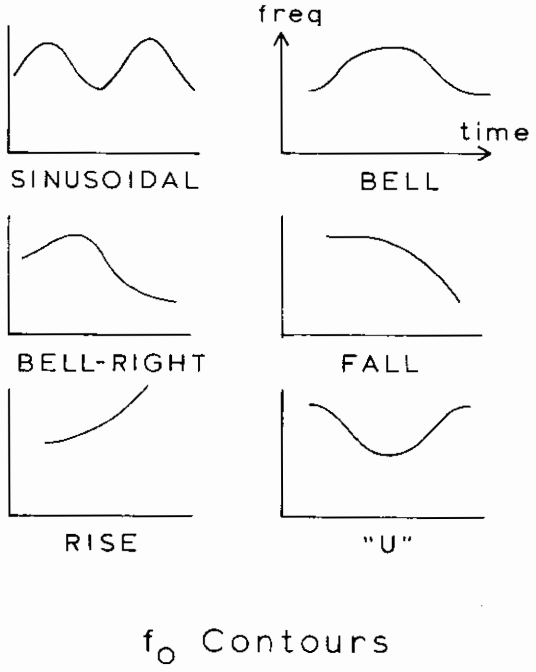


FIGURE 2. Intonation (fundamental frequency) contours typically observed for mothers' vocalizations to infants. After Stern et al. (1982).

A particularly interesting question about the development of intonation is the extent to which the infant imitates intonational patterns of the caregiver and vice-versa. Papousek and Papousek (1982) examined vocal imitations in mother-infant dialogues for infant ages of 2, 3, 4, and 7 months. In mother-infant sequences, pitch matching showed a steady decline, matching of melodic contour first fell and then rose in percentage occurrence, matching of temporal structure showed a late increase, and matching of phonetic structure had a variable pattern. Pitch matching was the predominant form of matching, achieving a high of almost 90% of all matching episodes at 7 months. Similarly, pitch matching was the most frequently occurring match in infant-mother sequences, accounting for about 85% of all matches at 2 months and almost 50% at 7 months. Mothers' matching of their infants' melodic contour and temporal structure showed an overall increase but was accelerated in the period between 4 and 7 months.

Another basic organizational process that contributes to early sound patterns is one of rhythmic pattern. This process may underlie the phenomenon of reduplicated babbling, a form of babbling in which a basic pattern is repeated within a string. Actually, rhythmic or cyclic patterning is a nearly universal characteristic of living systems. Delcomyn (1980) has reviewed evidence for a periodic organization of many phenomena in living systems. Cyclicity becomes an especially important force in organizing natural systems in that cyclic coupling or entrainment can occur within different parts of the same organism or between different organisms. Mechanisms of rhythmic organization have been proposed to explain vastly different phenomena in living systems, ranging from gastrointestinal rhythms (Bardakjian & Sarna, 1981) to the organization of leg movement in insect locomotion (Gallistel, 1980; Wilson, 1966), to the interactions between mothers and their infants (Jaffe, 1977).

As mentioned earlier, strong evidence for the rhythmic organization of early vocal patterns is the fairly regular appearance of a sound pattern called reduplicated babbling, which refers to reiteration of a basic syllable pattern (Figure 3). Reduplication is important not only to describe characteristics of babbling but it applies as well to sound patterns in early word formation, such as a child's production of /gaga/ for *doggy*. But reduplication is one of the very few phonological patterns that is unique to first-language acquisition; that is, reduplication is not observed in the acquisition of a second or later language. Moreover, although reduplication frequently has been considered as a special kind of phonological phenomenon, it appears that reduplicated babbling is coincident with a more generalized rhythmicity of behavior, or what Thelen (1979, 1981a, 1981b) has termed *rhythmic stereotypies*. These rhythmic movements of the limbs, fingers, and torso are very common in the first year of life and appear to reach their peak frequency at about 6–7 months of age (Figure 4), which is about the same time that reduplicated babbling is generally acknowledged to begin. Thus, it has been suggested that reduplicated babbling is part of a general behavioral process in which movements are organized in cyclic patterns (Kent, (1984). It is pertinent here to note reports that deaf babies who learn to communicate by manual signs pass through a "babbling" stage (Siple, 1978). One wonders if this babbling stage of the infant signer really is distinct from some of the repetitive hand and finger movements typical of infants in the first year of life.

The importance of repetitive babbling to speech development is unclear. However, Elbers (1982) concluded that this vocal pattern is a starting point for first words and a springboard for the gradual approximation to adult words. She commented that when jargon babbling and first words appeared, repetitive babbling coexisted with them and "seemed to be developing further according to its own principles" (p. 60). Repetitive babbling may be important to phonetic development insofar as it provides opportunities for the emergence of intra-utterance sound contrasts as well as early experience with a kind of prosodic shaping. Repetitive babbling yields both vocal productivity and sound contrast within a prosodic grouping established by the rhythmic pattern. Elbers saw in repetitive babbling the need to propose a cognitive continuity theory of prespeech vocalizations. For that matter, repetition, although not necessarily periodic repetition, is a primary means to accomplishments in the sensorimotor period of Piaget's theory of cognitive development. Repetition underlies the circular reactions observed during this period, ranging from primary circular reactions in which a repetitious be-

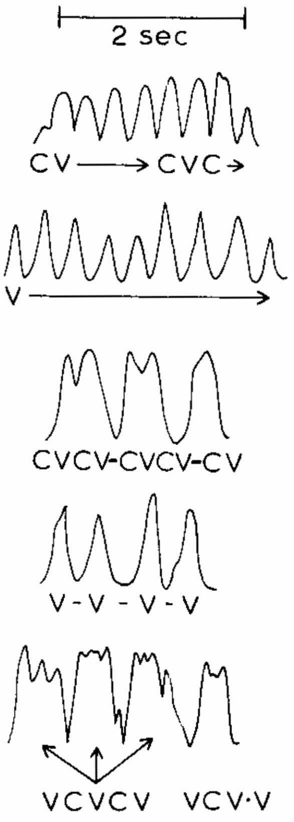


FIGURE 3. Examples of intensity envelopes for strings of reduplicated babble. Syllabic structure is indicated below each trace (*C* = consonant and *V* = vowel). A basic periodicity tends to be preserved in each reduplicated string.

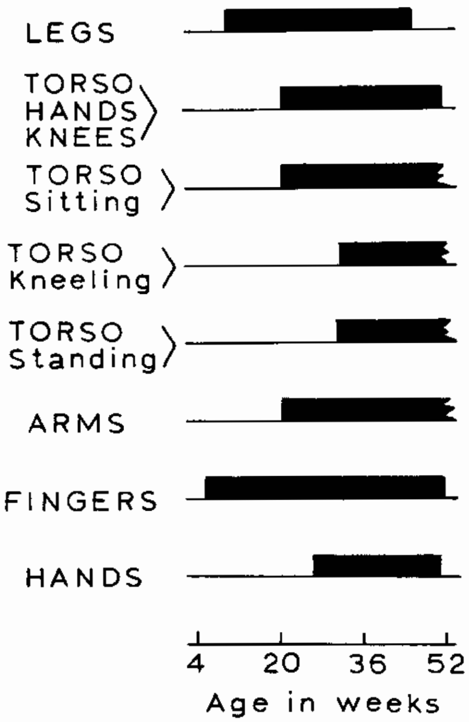


FIGURE 4. Peak periods of rhythmic stereotypies observed for various body movements of human infants. For example, rhythmic arm movements tend to occur principally between 20–52 weeks. (After Thelen, 1981a). (Reprinted with permission from *The American Journal of Physiology*, 1984, 246.)

havior is focused on the child's own body, to secondary reactions in which actions are repeated on external objects, to tertiary reactions in which new events or objects are thoroughly examined by repetitious manipulation.

We have been studying these processes in the vocalizations of infants and young children. Given the topic of this symposium, it is pertinent to compare features of vocal development in normal-hearing and hearing-impaired infants. It is fairly well established that even deaf infants babble (Mavilya, 1969) but it is less certain that the acoustic-phonetic content of babbling is the same for deaf and normal-hearing infants. We are following the vocal development for twin boys, one of whom has normal hearing and the other of whom has profound hearing impairment (Goldschmidt, Kent, Netsell, & Osberger, 1984). In the first 2 years, both boys vocalized and in doing so uttered a predominance of front vowels and front consonants. They differed especially in the production of fricatives and trills. The normal-hearing infant exercised these sounds frequently and with long durations. But they were virtually absent in the babbling of the hearing-impaired boy. It should be noted that the acoustic spectra of infant trills and fricatives are characterized by noise energy at exceptionally high frequencies—to 14 KHz and above. Our own data (Bauer & Kent, in press) on these sounds are illustrated in Figure 5, which shows the primary noise energy for fricatives produced at five different places of articulation. Obviously, an infant with any degree of high-frequency hearing loss is at a serious disadvantage in hearing himself or herself produce fricatives.

However, the twin with hearing impairment did not lack in the production of stop consonants, front vowels, and (eventually) reduplicated babble. Examples of his babble are given below.

[gæ ga gæ]
[ba baɪ ʌm]
[ba baɪ bi bi]

Apparently, the tendency toward rhythmically organized vocalizations was not compromised by the hearing loss.

Principles of self-organization also may underlie some general aspects of phonological and lexical acquisition. Stemberger (1982) proposed two major hypotheses about the nature of phonological segments in the lexicon. The Feature-Segment hypothesis asserts that segments are bundles of features. The Indivisible-Segment holds that features are derived from a segmental lexicon and that segments in the lexicon cannot be analyzed. A major distinction between the two hypotheses is that "putting archisegments and fine phonetic detail into the lexicon" (p. 235) amounts to a simplification under the Feature-Segment hypothesis but a complication under the Indivisible-Segment hypothesis. Stemberger concluded from speech error data from adults that different types of errors are more accurately predicted by the Indivisible-Segment hypothesis.

In evaluating the two hypotheses against developmental data on child language, Stemberger again chose the Indivisible-Segment hypothesis. This hypothesis is consistent with the strategy that early words will not be placed in a segmental storage but rather will be stored as feature matrices, "probably without any division into segments" (p. 256). Segmentation presumably arises as the child recognizes recurrent similarities in a growing vocabulary. Eventually, the child establishes a segmental lexicon to avoid storage of feature matrices in many different locations within the lexicon. In contrast, the Feature-Segment hypothesis maintains that early words will be stored as feature bundles associated with a segmental composition. Thus, even the earliest productions would require a copying mechanism that derives features from a foregoing segmental analysis.

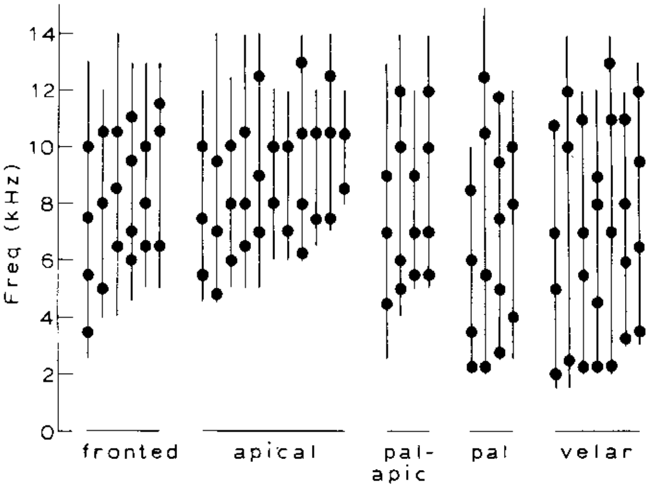


FIGURE 5. Primary frequency regions of noise energy for infant fricatives produced at five places of articulation. Each vertical line is a result for one sound. Filled circles indicate spectral peaks within the noise region.

Evidence to support the Indivisible-Segment hypothesis is mainly in the form of data showing that phonetic productions may get worse (i.e., deteriorate with respect to the adult model) as acquisition proceeds. Stemberger points to two major types of evidence: regressions and idioms. He notes that phonetic regressions typically appear at the time when the child has acquired about 50 words. At this point, words that fairly closely approximate the adult phonetic form may undergo an apparent regression, as in the case of Hildegard's replacement of /prɪtɪ/ by /bɪdɪ/ (Leopold, 1947). Stemberger explains regression as reflecting the child's establishment of a segmental lexicon. No longer can the child retrieve feature matrices directly from the lexicon, for he/she must instead access the segmental lexicon. When this segmental lexicon is first put into effect, the child has little facility in its use and is therefore highly susceptible to errors. Phonological idioms are phonetic forms that conform more closely to the adult model than do the majority of the child's productions. These islands of phonetic propriety stand in contrast to the more abundant childlike productions. Stemberger explains idioms by assuming that these items are still lexicalized as feature matrices while most other productions are derived from the segmental lexicon. In other words, two lexical representations may co-exist: one based on the segmental lexicon and another based on a direct feature-matrix lexicon.

An important consequence of Stemberger's proposals for child language acquisition is that they lead to the view that many phenomena in phonological development can be ascribed to errors rather than to rules (e.g., the simplification rules favored in many accounts of phonological development). He remarks: "The orthodox view of child language completely ignores the possibility of error, leading to the paradoxical conclusion that young children, who have as yet acquired a language only very poorly, make no errors in performance, while adults, who have acquired the language well, make frequent errors" (p. 257). Stemberger goes on to observe that many of the "processes" in phonological development, such as syllable and segment deletions and consonant harmony, are similar to adult performance errors. Moreover, the fact that a child will improve phonetic accuracy upon repeated attempts at an utterance also accords with error data for both normal and aphasic adults.

Interestingly the idea that developmental sound patterns reflect errors rather than rules or processes, however this term is to be defined, is consonant with recent investigations of confusions in auditory recognition and memory. Aitchison and Chiat (1981) examined the recall errors of children between 4 and 9 years in a task of word learning. The patterns of recall error strongly resembled phonological processes described in the literature on children's speech development. For example, common errors included (a) high frequency of occurrence of consonant-vowel syllable structure, (b) omission of unstressed syllables, (c) consonant harmony, (d) substitution of one consonant by another of the same natural class, usually differing by one feature, and (e) stress placement on vowels. The authors proposed two possibly compensatory explanations. First, they suggested that many phonological processes in developing speech are attributable to faulty recall; (i.e., new words exceed a child's memorial capability so that some features are not stored immediately). The second explanation is that recall is determined by perceptual salience, such that features with the greatest perceptual potency also are most potent in auditory memory.

Some relevant data on these explanations have been reported by Macari (1978), who tested the predictions of "natural processes" in phonology (Stampe, 1973) against data for consonant confusions in three studies of adult listeners (Miller & Nicely, 1955; Wang & Bilger, 1973; Wickelgren, 1966). Generally, the natural process predictions were compatible with the confusion data. Macari interpreted his results to support Stampe's claim that the processes are central mental operations rather than peripheral physical constraints. An alternative explanation is that the processes are determined primarily by perceptual or memorial factors and their appearance in production is merely a reflection of their influence on auditory analysis and/or auditory memory. Although a conclusion as to the relative contributions of auditory perception and auditory memory is premature, it is pertinent that Wickelgren's (1966) study evaluated errors in short-term memory and controlled for errors in perception. Because the natural process predictions generally held for Wickelgren's data, it might be concluded that the error patterns resulted more from memorial factors than auditory perceptual factors.

Whatever the relative importance of perceptual vs. memorial factors might be in explaining errors and progress in phonological acquisition, it is important to recognize that the infant is an active pattern seeker. Studdert-Kennedy (1981) described this process of discovery as follows:

> The infant does not simply imitate, matching a particular utterance to a particular type of situation. Rather, it searches out contrasts among components of its own repertoire and uses them to signal contrasts in its desires, experience, or behavior. Often the contrasts, in both signal and message, are entirely novel and without counterpart in the adult system. (p. 553)

Aspects of Neural Organization

Although it often is supposed that neural representations of behavior are fixed patterns of neural connectivity, some recent evidence points to a very different perspective. Merzenich, Kass, Wall, Nelson, Sur, & Felleman (1983) reported that the somatosensory representation of a monkey's hand changes as the result of temporary denervation, bandaging, or amputation of fingers. For example, following amputation of a finger, neurons originally involved in somatosensory representation of that finger are effectively reassigned to non-amputated fingers. Thus, neural connectivities may be dynamically optimized to promote efficient control of behavior. However, see Dykes (1983, pp. 85-87) for a different view of these studies.

A similar idea was advanced by Ojemann (1983) in conjunction with cortical stimulation performed during craniotomy. He concluded that language representation in the brain changes as a function of language proficiency. Rosenbek, Kent, and LaPointe (1984) proposed that,

> it might be better to think of functional centers not as invariant connections of neurons but as connectivity patterns that are continually subject to reassignment and revision. Thus, an important property of the brain would be its ability to reformulate neuronal connections to achieve optimal or efficient control of function. In other words, the brain would assume auto-regulatory (self-organizing) properties in which the process of organization is not distinct from the product of organization.

All of this becomes relevant to the problem of teaching the deaf to talk when this problem is cast in terms of how we can best encourage the process of self organization. The task may not be to teach rules or other formalized structures, but rather to attempt to present bodies of information that are sufficient to launch the process of self organization. We sometimes like to think that we "teach" rules by giving the right kinds of examples to illustrate the rule. What we actually may be doing is to provide the suitable raw material to stimulate self-organizing processes.

REFERENCES

AITCHISON, J., & CHIAT, S. (1981). Natural phonology or natural memory? The interaction between phonological processes and recall mechanisms. *Language and Speech*, *24*, 311–326.

BARDAKJIAN, B. L., & SARNA, S. K. (1981). Mathematical investigation of populations of coupled synthesized relaxation oscillators representing biological rhythms. *IEEE Transactions on Biomedical Engineering*, *BME-28*, 10–15.

BAUER, H. L., & KENT, R. D. (1982). *Acoustic analyses of closant sounds in infants' comfort-state utterances.* Paper presented to the 103rd Meeting of the Acoustical Society of America, Chicago, IL.

BOSWELL, J. (1951). *The life of Samuel Johnson* (p. 1,044). New York: The Modern Library.

Branigan, G. (1976). Syllabic structure and the acquisition of consonants: The great conspiracy in word formation. *Journal of Psycholinguistic Research, 5,* 117–133.

Bruner, J. (1981). The social context of language acquisition. *Language and Communication, 1,* 155–178.

Butler Platt, C., & MacWhinney, B. (1983). Error assimilation as a mechanism in language learning. *Journal of Child Language, 10,* 401–414.

Couly, G. (1982). New concepts of the biology of cephalic development. Observations and hypotheses. *Annales de Genetique, 25,* 201–206.

Cromer, R. F. (1983). A longitudinal study of the acquisition of word knowledge: Evidence against gradual learning. *British Journal of Developmental Psychology, 1,* 307–316.

Cruttenden, A. (1982). How long does intonation acquisition take? *Papers and Reports on Child Language Development, 21,* Department of Linguistics, Stanford University.

Delcomyn, F. (1980). Neural basis of rhythmic behavior in animals. *Science, 210,* 492–498.

Donegan, P. J., & Stampe, D. (1979). The study of natural phonology. In D. A. Dinnsen (Ed.), *Current approaches to phonological theory* (pp. 126–173). Bloomington, IN: Indiana University Press.

Dykes, R. W. (1983). Parallel processing of somatosensory information: A theory. *Brain Research Reviews, 6,* 47–115.

Elbers, L. (1982). Operating principles in repetitive babbling: A cognitive continuity approach. *Cognition, 12,* 45–63.

Fentress, J. C. (1981). Order in ontogeny: Relational dynamics. In K. Immelman, G. W. Barlow, L. Petrinovich, & M. Main. *Behavioral development: The Bielefeld interdisciplinary project* (pp. 338–371). Cambridge, UK: Cambridge University Press.

Ferguson, C. A. (1977). Learning to pronounce: The earliest stages of phonological development in the child. In F. D. Minifie & L. L. Lloyd (Eds.), *Communication and early cognitive abilities: Early behavioral assessment.* Baltimore: University Park Press.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51,* 419–439.

Ferguson, C. A., & Garnica, O. K. (1975). Theories of phonological development. In E. H. Lenneberg & E. Lenneberg (Eds.), *Foundations of language development* (Vol. 2) (pp. 153–180). New York: Academic Press.

Gallistel, C. R. (1980). *The organization of action: A new synthesis.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Goldschmidt, C., Kent, R., Netsell, R., & Osberger, M. J. (1984, November). *Effects of hearing loss on vocal development: A twin study.* Paper presented at the American Speech-Language-Hearing Association, San Francisco, CA.

Guntern, G. (1982). Auto-organization in human systems. *Behavioral Science, 27,* 323–337.

Jaffe, J. (1977). Markovian communication rhythms: Their biological significance. In M. Lewis & L. A. Rosenblum (Eds.), *Interaction, conversation, and the development of language.* New York: Wiley.

Jakobson, R. (1941). *Child language, aphasia and phonological universals* (A. R. Keller, Trans.) The Hague: Mouton.

Jantsch, E. (1979). *Die selbstorganisation des universmus.* Munchen: Hansen Verlag.

Kent, R. D. (1981a). Sensorimotor aspects of speech development. In R. N. Aslin, J. R. Alberts, & M. R. Peterson (Eds.), *The development of perception: Psycho-biological perspectives* (pp. 161–189). New York: Academic Press.

Kent, R. D. (1981b). Articulatory-acoustic perspectives on speech development. In R. E. Stark (Ed.), *Language behavior in infancy and early childhood* (pp. 105–126). New York: Elsevier/North Holland.

Kent, R. D. (1983). The segmental organization of speech. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 57–89). New York: Springer-Verlag.

Kent, R. D. (1984). Psychobiology of speech development: Coemergence of language and a movement system. *American Journal of Physiology, 246,* R888-R894.

Kent, R. D., Lippmann, R., Osberger, M. J. (1980). Acoustic characteristics of the speech of hearing-impaired adolescents. Unpublished report.

Kiparsky, P., & Menn, L. (1977). On the acquisition of phonology. In J. Macnamara (Ed.), *Language learning and thought* (pp. 47–78). New York: Academic Press.

Kuehn, D. P., & Tomblin, J. B. (1974). *The use of cineradiographic techniques for the study of articulation disorders.* Paper presented at the Annual Convention of the American Speech and Hearing Association, Las Vegas, NV.

Labov, W., & Labov, T. (1978). The phonetics of *cat* and *mama. Language, 4,* 816–852.

Leonard, L., Newhoff, M., & Mesalam, L. (1980). Individual differences in early child phonology. *Journal of Applied Psycholinguistics, 1,* 7–30.

Leopold, W. F. (1947). *Speech development of a bilingual child: A linguist's record, Vol. 2: Sound learning in the first two years.* Evanston, IL: Northwestern University Press.

Macari, N. J. (1978). Some psychophysical evidence for natural phonological processes. *Journal of Phonetics, 6,* 103–126.

Macken, M. A., & Ferguson, C. A. (1981). Phonological universals in language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition. Annals of the New York Academy of Sciences, 379,* 110–129.

Maruyama, M. (1960). Morphogenesis and morphostasis. *Methods, 12,* 251–296.

Mavilya, M. (1972). Spontaneous vocalization and babbling in hearing impaired infants. In G. Fant (Ed.), *International symposium on speech communication ability and profound deafness.* Washington, DC: Alexander Graham Bell Association for the Deaf.

Menn, L. (1981). Theories of phonological development. In H. Winitz (Ed.), *Native language and foreign language acquisition. Annals of the New York Academy of Sciences, 379,* 130–137.

Merzenich, M. M., Kaas, J. H., Wall, J. T., Nelson, R. J., Sur, M., & Felleman, D. J. (1983). Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience, 8,* 33–55.

Miller, G. A., & Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27,* 338–352.

Moskowitz, B. A. (1973). Acquisition of phonology and syntax. In K. J. J. Hintakka, J. M. E. Moravcsik & P. Suppes (Eds.), *Approaches to natural language* (pp. 48–84). Dordrecht: Netherlands: Reidel Publishing.

Nalimov, V. V. (1981). *Faces of science.* Philadelphia: ISI Press.

Ojemann, G. A. (1983). Brain organization for language from the perspective of electrical stimulation mapping. *The Behavioral and Brain Sciences, 6,* 189–230.

Olmsted, D. L. (1971). *Out of the mouth of babes.* The Hague: Mouton.

Papousek, H., & Papousek, M. (1982, March). *Vocal imitations in mother-infant dialogues.* Paper presented to the International Conference on Infant Studies, Austin, TX.

Prigogine, I., & Stengers, I. (1984). *Order out of chaos.* New York: Bantam Books.

Roberts, E. W. (1979). Strategy relationships in the acquisition of /CV/, /CVCV/, /CVC/ and other structures: A case study. *Phonetica, 36,* 130–143.

Rosenbek, J. C., Kent, R. D., & LaPointe, L. L. (1984). Apraxia of speech: An overview and some perspectives. In J. C. Rosenbek, M. R. McNeil, & A. E. Aronson (Eds.), *Apraxia of speech: Physiology, acoustics, perception, treatment* (pp. 1–72). San Diego: College-Hill Press.

Ryan, M. L. (1978). Contour in context. In R. N. Campbell & P. T. Smith (Eds.), *Recent advances in the psychology of language.* New York: Plenum Press. (pp. 237–251).

Ryan, M. L. (1978). *Contour in context.* Paper presented at the Psychology Language Conference, Stirling, Scotland, 1976.

Siple, P. (Ed.). (1978). *Understanding language through sign language research.* New York: Academic Press.

Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development,* (pp 175–208). New York: Holt Rinehart & Winston.

Smith, N. V. (1973). *The acquisition of phonology.* Cambridge, UK: Cambridge University Press.

SMITH, N. V. (1975). Universal tendencies in the child's acquisition of phonology. In N. O'Connor (Ed.), *Language, cognitive deficits and retardation* (pp. 47–65). London: Butterworths.

STAMPE, D. (1969). *The acquisition of phonetic representation* (pp. 443–454). Papers from the Fifth Regional Meeting of the Chicago Linguistic Society.

STAMPE, D. (1973). *A dissertation on natural phonology*. Unpublished doctoral dissertation, University of Chicago, IL.

STEMBERGER, J. P. (1982). The nature of segments in the lexicon: Evidence from speech errors. *Lingua, 56,* 235–259.

STERN, D. N., SPIEKER, S., & MACKAIN, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology, 18,* 727–735.

STUDDERT-KENNEDY, M. (1981). The beginnings of speech. In K. Immelmann, G. W. Barlow, L. Petrinovich, & M. Main (Eds.), *Behavioral development: The Bielefeld interdisciplinary project* (pp. 533–561). Cambridge, UK: Cambridge University Press.

THELEN, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behavior, 27,* 699–715.

THELEN, E. (1981a). Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology, 17,* 237–257.

THELEN, E. (1981b). Kicking, rocking, and waving: Contextual analysis of rhythmical stereotypies in normal human infants. *Animal Behavior, 29,* 3–11.

THOMPSON, A. E., & HIXON, T. J. (1979). Nasal air flow during normal speech production. *Cleft Palate Journal, 16,* 412–420.

WANG, M. D., & BILGER, R. C. (1973). Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America, 54,* 1248–1266.

WICKELGREN, W. A. (1966). Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Socity of America, 38,* 583–588.

WILSON, D. M. (1966). Insect walking. *Annual Review of Entomology, 11,* 103–122.

WOODSON, R. H. (1983). Newborn behavior in the transition to extra-uterine life. *Infant Behavior and Development, 6,* 136–144.

Chapter 7

SPEECH PRODUCTION AND ACOUSTIC GOALS

KENNETH N. STEVENS

Massachusetts Institute of Technology, Cambridge, MA

When we generate an utterance, we attempt to control the positions and movements of the articulatory structures and the respiratory system so as to achieve a certain pattern of acoustic goals or targets. There are some aspects of these goals that are obligatory and must be achieved with some precision if the utterance is to be understood by a listener, and there are other aspects that are optional or variable, or can be realized with less precision. These latter aspects may be used by a listener to enhance the decoding of an utterance, but are, to some extent, redundant. In this paper we propose to discuss some of the acoustic properties that form these goals, and we shall give a number of examples of situations where some of these acoustic properties are redundant or optional, and where the redundant properties can serve to enhance the ability of a listener to perceive a phonetic distinction.

These concepts of variability, redundancy, and enhancement could have consequences in the way one might approach the training of hearing-impaired children in the acquisition of speech production and speech perception skills. At the end of this paper we shall attempt to indicate what some of these consequences might be.

SOME EXAMPLES OF ACOUSTIC GOALS OR PROPERTIES

A listener-speaker of a language has available an inventory of acoustic properties that can be used to distinguish between sounds and words in the language. There is a limited set of these properties—probably around 20—that are used in the languages of the world to make distinctions between words (Chomsky & Halle, 1968; Jakobson, Fant, & Halle, 1963; Ladefoged, 1971). For each word in the language, a pattern of these properties is specified, and, except for certain situations where some of the properties may be optional or where there are rules for deleting, substituting, or adding new properties, these properties form the goals that a speaker must achieve when producing the word.

This inventory of acoustic properties seems to be constrained according to two general principles. One of these is that the auditory system produces a distinctive type of response when a particular property is present in the sound. The other principle is that the speech production system can be controlled to produce each property in a consistent and repeatable way independent of the context of other properties that are being produced in the same word.

Continuing research efforts are being directed towards establishing the inventory of properties that are used to make distinctions between speech sounds in language. We shall give here a few examples of these properties, particularly those for which there are recent data or perhaps some new insights. Some of these examples are also selected because they may be of particular concern in the speech training of deaf speakers.

One of the basic properties that is used to distinguish sounds in words is determined by whether or not there is a buildup of pressure behind a constriction in the vocal tract. Linguists and phoneticians sometimes use the term *obstruent* to define the class of sounds that are produced with pressure behind the constriction, whereas sounds with no pressure buildup in the vocal tract are *sonorants*. If there is a pressure buildup, then some noise is generated by the rapid flow of air through the constriction. In the case of fricative consonants such as /s/ and /z/, this noise persists throughout the time when there is a constriction. For stop consonants such as /t/ and /d/, which are produced with a complete closure in the vocal tract, noise is produced during a brief interval following the release of the closure. There is no pressure buildup and no noise generation at a constriction for sonorants, which include nasal consonants, vowels, liquids and glides.

It is important to note that the pressure buildup behind the constriction for an obstruent consonant is not the result of conscious pulmonary action to increase the pressure. Rather, it is an automatic consequence of maintaining a more-or-less constant pressure in the lungs and trachea throughout an utterance. Forming a closure or narrow constriction in the vocal tract (with a closed velopharyngeal opening) while maintaining a relatively open glottis results automatically in a buildup of pressure behind the constriction, so that this pressure tends to become equal to the lung pressure. No special force by the system of muscles surrounding the thoracic cavity is required to produce an obstruent consonant. Deaf speakers often do not control the respiratory system to yield a constant subglottal pressure throughout an utterance. Informal observations suggest that some speakers adopt a strategy of building up too high a pressure during an obstruent consonant, particularly a stop consonant. This increased pressure results in a burst at the consonant release that is too strong. On the other hand, deaf speakers often produce an utterance at the lower end of their respiratory range (Whitehead, 1983), and this could result in too low an intraoral pressure, and turbulence noise that is too weak.

Within the class of obstruent consonants, we distinguish two different ways of generating the noise resulting from the rapid flow of air through the constriction. If the air is directed against an obstacle in the vocal tract then there is much more noise energy than if there is no obstacle against which the air impinges. These different mechanisms for generating noise form the basis for the strident-nonstrident distinction. Most fricative consonants in English and in many other languages are strident, presumably because the greater noise intensity within the constricted interval helps to distinguish these consonants from other classes of sounds such as stop consonants.

In English, the consonants /s/ /š/ /z/ /ž/ /č/ /ǰ/ are strident, and the obstacle against which the airstream is directed is the lower incisors. As is well known to those involved in speech training of the deaf, it is difficult to train a speaker with no high-frequency hearing to shape the constriction between tongue blade and palate or alveolar ridge in such a way that the airstream is properly directed toward this obstacle. For the nonstrident consonants /θ/ and /ð/ the tongue blade is placed in such a way that the airstream does not impinge against teeth or lips. The labiodental consonants /f/ and /v/ are usually made by directing the airstream against the upper lip to form a strident sound.

One of the ways of producing vowels with distinctively different acoustic characteristics is to displace the tongue body in a forward or backed position in the mouth. The principal acoustic consequence of a front-back displacement of the tongue body is to raise or lower the frequency of the second formant of the vowel (Delattre, 1951; Fant, 1960; Stevens & House, 1955). The second-formant frequency is high and close to the third formant for front vowels, and is low and close to the first formant for back vowels.

Because this type of tongue-body movement is not readily visible, and since changes in second-formant frequency may not be audible to deaf individuals even if there is some residual low-frequency hearing, it is common for deaf speakers not to create a sufficient range of front-back movement of the tongue body and hence to have a limited frequency range for the second formant. A limited range of tongue-body movement can also lead to improper production of some consonants, since some front-back tongue-body adjustment is needed to position the constriction properly, particularly for consonants produced with the tongue blade. Thus limitations in front-back tongue body movement can lead to severe reduction in intelligibility (Monsen, 1976).

We observe that these three distinctions—sonorant-obstruent, strident-nonstrident, and front-back—all have well-defined acoustic and articulatory correlates. These are just three examples of the twenty-odd properties that appear to be available for making distinctions in language, and for which articulatory and acoustic correlates can be defined. They also illustrate the kinds of problems that deaf speakers may experience in producing sounds with properties that cannot be detected with an impaired auditory system.

Redundancy, Enhancement, and Variability

As we have indicated, the process of speech production involves controlling the movements and configurations of the various articulatory structures, including the respiratory system, so as to achieve a desired pattern of acoustic properties. The phonetic aspects of language are structured, however, in such a way that there are more acoustic properties available than are actually needed to distinguish among the words in a language, assuming a clear and noise-free communication channel. That is, there is redundancy in the speech signal.

Thus, for example, to distinguish between the words *beet* and *bead* (spectrograms of which are shown in Figure 1), the speaker may use several properties: the duration of the vowel preceding the stop closure, the presence or absence of vocal-fold vibration during the initial part of the stop closure, and the intensity and duration of the burst of noise at the release. Any one of these properties may be sufficient to communicate to a listener which of the two words is intended, if the communication channel is free of significant distortion or noise.

Another example is the distinction between the words *beet* and *bit*, also illustrated in Figure 1. The distinction between these vowels is carried by several different properties: the formant frequencies for the vowel /ɪ/ are different from those for /i/, the vowel /ɪ/ is shorter, and the trajectories of the formants are different, with the vowel /i/ being diphthongized toward /y/, and /ɪ/ having an offglide toward the schwa /ə/.

Many examples of this type can be cited. In fact, it is difficult to find a situation where a distinction between a minimally different pair of words is carried by just a single acoustic property. Almost always one finds more than one property playing a role in signalling a distinction between minimal pairs of utterances.

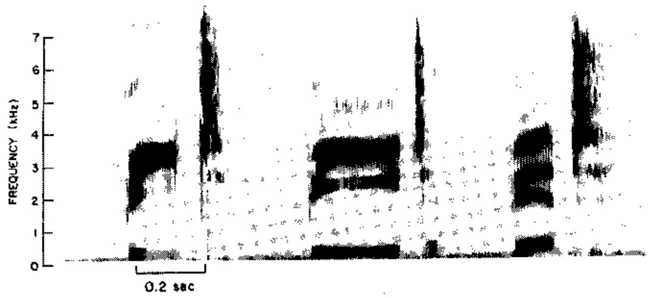


FIGURE 1. Spectrograms of the words *beet* (left), *bead* (middle) and *bit* (right). These spectrograms illustrate that several properties are potentially available for distinguishing voiceless /t/ in *beet* from voiced /d/ in *bead*, and, likewise, several properties are available for distinguishing the vowel /i/ in *beet* from the vowel /ɪ/ in *bit*.

One consequence of this predisposition to introduce redundant properties is that there is an opportunity for variability in the way words are produced. Since several properties may be available to signal a distinction, it is not necessary to produce all of these properties in the sound. Consequently there are many situations in which a speaker may choose to include or to omit certain properties.

Another kind of redundancy arises from the fact that the sequence of sounds that can occur in a word are constrained, and the sequence of words that can occur in an utterance are also constrained by syntactic and semantic considerations. These sources of redundancy introduce still further opportunities for omitting or modifying properties of speech sounds as they occur in words or sentences. For example, in a sentence like "A stitch in time saves nine" it is necessary for a speaker to provide only minimal acoustic information concerning the sounds in the final word, since it is largely predictable from the context.

If a speaker chooses to produce a reduced or modified set of properties to represent an utterance, the listener can decode the utterance in terms of a sequence of words only if he or she has knowledge of the set of alternative properties that can be used to represent the sounds and has sufficient knowledge of the language to indicate what are the constraints on sequences of speech sounds and of words. In other words, knowledge of the language is needed if one is to make use of redundancy in the sound to enhance or to increase the reliability of the decoding of an utterance.

Clear and Conversational Speech

To illustrate the variability that can occur in the acoustic representation of speech sounds, we shall give examples of the same utterances produced in two different modes. In the first mode, the speaker is talking clearly, as though trying to make himself understood to a listener with a hearing impairment or in a situation where there is substantial background noise. The other mode is a conversational mode, which is normally used in everyday interactions when there are no special problems with the listener or the communication channel. We shall point out some of the differences between clear and conversational speech, and we shall attempt to interpret these differences in terms of a concept of optional or redundant features. Study of the differences in intelligibility and in the physical properties of speech in these different modes has been carried out by several investigators (Chen, 1980; Picheny, 1981).

Figure 2 shows spectrograms of two sentences, each produced in both a clear and a conversational mode. The sentences are: "The policeman will send you a nickel," and "Can you multiply these numbers together?"

The first thing we observe about these utterances is that the clearly spoken sentences are about twice as long as the sentences produced in a conversational style. Both vowels and consonants are longer, although the increase in length does not occur uniformly across all types of acoustic segments. For example, the /z/ in *these* or the initial vowel in *nickel* has about the same duration for the two speaking styles, whereas the vowel in *can* is about 200 msec long in the clear speech but is reduced almost to zero duration in the conversational style. The same kind of reduction in duration can be seen in the initial unstressed vowel in *together* and in *policeman*.

There are a number of differences in the way individual speech sounds are produced, apart from modifications in timing. For example, the consonant sequence joining the two words *numbers̲ t̲ogether* in the conversational mode does not show a complete closure for the stop consonant, although a reduction in amplitude of the noise provides evidence for the presence of the stop consonant. The vowel in the word *will* is completely reduced to a syllabic /l̩/ in the conversational mode, whereas it is represented as /ɪ/ in the clear speaking mode. The consonant /t/ in *multiply* is flapped in the conversational mode, and released with aspiration in the clear speech. The consonants joining the words *send you* are produced as a palatal affricate in the conversational mode (i.e., [sɛnǰu]) but as a sequence /d y/ in the clear speaking mode.

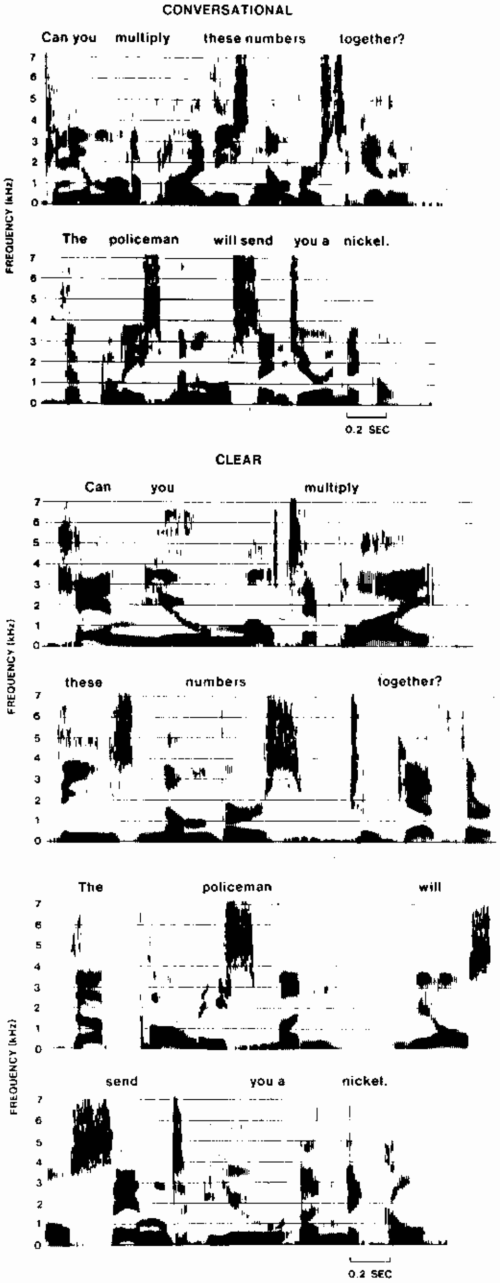


FIGURE 2. Spectrograms of two sentences produced in a conversational style (upper two rows) and in a clearly spoken style (bottom four rows, with each sentence occupying two rows). The sentences are: "Can you multiply two numbers together?" and "The policeman will send you a nickel."

The following article *a* is joined continuously with *you* in the conversational version, but the word boundary is marked with a glottal stop in the clear version. The final diphthong in *multiply* shows clearly the appropriate formant movements in both versions, except that the second formant does not reach as high a frequency in the conversational utterance as it does in the clearly spoken version. In the final vowel in the word *numbers*, the lowering of the third formant frequency, which is the main acoustic characteristic for retroflexion, is more extreme in the clear speech than in the conversational mode. The fricative consonant /ð/ is strongly voiced when it occurs in the conversational version, but is voiceless over much of its duration in the clear version.

Most of these variations in the properties of speech sounds in the two speaking modes occur in unstressed syllables or at word boundaries. Stressed vowels and consonants preceding these vowels are not greatly affected, except possibly in duration.

The examples we have shown illustrate the kinds of redundancy and variability that can occur in the speech signal, depending on the context in which a speech sound occurs and on the mode of speaking. We are suggesting that the variability arises primarily because there is the potential for redundancy in the signal. The sounds in an utterance usually contain more acoustic properties than are really needed by a listener to understand the utterance. This potential for variability can give speakers an opportunity for individual differences and style. The redundancy is, however, needed in situations where the communication channel is impaired. A difficult and important challenge for speech researchers is to determine the rules that govern the selection of properties to use in a particular utterance in a particular situation, and that indicate which properties can be modified and which ones can be omitted.

RELEVANCE TO SPEECH PRODUCTION AND PERCEPTION IN THE DEAF

These remarks have attempted to emphasize that the production of speech involves (a) the ability to produce the requisite acoustic properties in the proper combinations and sequences, and (b) the capability to bring into play certain redundant properties to enhance the distinctions produced by some of the properties or to omit or reduce these properties in situations where they may not be necessary. Can these concepts of redundancy and enhancement provide us with insight into how to work with deaf children in improving their speech reception and speech production skills?

It is well known that one of the most effective ways of improving the speech reception capabilities of the hearing impaired is to speak clearly (cf. Chen, 1980; Picheny, 1981). Speaking clearly does three things: (a) it strengthens or enhances particular required properties in the speech signal; (b) it introduces certain redundant properties, which provide extra cues for the identification of a word; and (c) it lengthens the utterance, and consequently provides the listener with additional time to decode the units contained within the utterance. In working with the hearing impaired, then, it is important to speak clearly so that their understanding of speech is maximized. If small amounts of residual hearing are available, good communication will be enhanced and strengthened if strengthened properties are provided. Data collected by Picheny (1981) have provided quantitative data indicating the gain in intelligibility by hearing-impaired listeners for clear speech as opposed to conversational speech. A sample of Picheny's data, given in Figure 3, shows a significant increase in intelligibility for words in sentences, when clear speech is used. The use of a clear speaking mode is particularly important with individuals whose knowledge of language is still developing. Omission or reduction of some acoustic properties can be particularly damaging for an individual who is developing a basic lexicon, who is not yet conversant with phonotactic rules governing allowed sequences of phonetic elements, and who is still attempting to grasp the essentials of grammar in the language.

On the other hand, hearing-impaired individuals are often exposed to conversational speech in their day-to-day living, and consequently need to become familiar with the types of reductions and other modifications that can occur in conversational speech. Thus it is important at a certain stage to use conversational speech with hearing-impaired children and to point out to them the kinds of differences that can exist between clear speech, which contains many redundant properties, and conversational speech, which omits or modifies many of the properties.

With regard to the training of speech production skills, it is possible that one's approach to training might be positively influenced if one had awareness of the variability in speech and the sources of this variability. For those hearing-impaired individuals whose speech production skills are well devel-

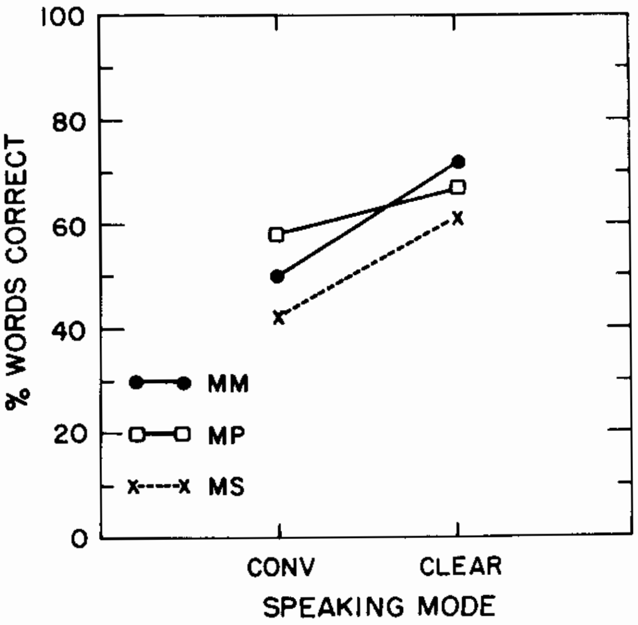


FIGURE 3. Examples of intelligibility scores for conversational and clear speaking modes, for three different speakers. Average data for five hearing-impaired listeners are shown. Scores represent percent content words correct in short nonsense sentences. (Reprinted with permission from Picheny, 1981.)

oped, the naturalness and fluency of their speech might be improved if they knew which properties could be omitted or reduced, and which properties are important to maintain. For any speaker who is not completely fluent, there might be an advantage in knowing what acoustic properties need to be strengthened or maintained in order to maximize the ability of a listener to understand the speech. When asked to repeat an utterance in an everyday communication situation, it would be helpful for the speaker to know how to modify or enhance the utterance so that the second time it is heard it is more likely to be understood.

When a deaf child is in the early stages of speech and language training, he or she should learn what are the primary acoustic properties that need to be reproduced accurately and consistently, and that are not subject to reductions and omissions. Although we are generally aware of what these properties are, more research is needed if we wish to catalogue these attributes systematically in a way that would make them optimally useful to those working with deaf children.

ACKNOWLEDGMENT

Preparation of this paper was supported in part by Grant No. NS-04332 from the National Institute of Neurological and Communicative Disorders and Stroke.

REFERENCES

CHEN, F. R. (1980). *Acoustic characteristics of clear and conversational speech at the segmental level.* Unpublished doctoral dissertation, Massachusetts Institute of Technology.

CHOMSKY, N., & HALLE, M. (1968). *The sound pattern of English.* New York: Harper & Row.

DELATTRE, P. (1951). The physiological interpretation of sound spectrograms. Publications of the Modern Language Association of America, *66*, 864–875.

FANT, G. (1960). *Acoustic theory of speech production.* The Hague: Mouton.

JAKOBSON, R., FANT, C. G. M., & HALLE, M. (1963). *Preliminaries to speech analysis.* Cambridge, MA: M.I.T. Press.

LADEFOGED, P. (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.

MONSEN, R. B. (1976). Normal and reduced phonological space: The production of English vowels by deaf adolescents. *Journal of Phonetics, 4*, 189–198.

PICHENY, M. A. (1981). *Speaking clearly for the hard of hearing.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

STEVENS, K. N., & HOUSE, A. S. (1955). Development of a quantitative description of vowel articulation. *Journal of the Acoustical Society of America, 27*, 484–493.

WHITEHEAD, R. L. (1983). Some respiratory and aerodynamic patterns in the speech of the hearing impaired. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training and personnel preparation*, (pp. 97–116). Baltimore: University Park Press.

Chapter 8

CONTEXT SIMILARITY CONSTRAINTS ON SEGMENTAL SPEECH ERRORS: AN EXPERIMENTAL INVESTIGATION OF THE ROLE OF WORD POSITION AND LEXICAL STRESS

STEFANIE SHATTUCK-HUFNAGEL

Massachusetts Institute of Technology, Cambridge, MA

Despite the progress made in recent years, teaching speech skills to those with speech dysfluencies remains a difficult and sometimes frustrating task, not made any easier by our profound ignorance of the representations and processing mechanisms that underlie the production of normal fluent speech. One source of information about this complex process is the patterns found in errors that arise in spontaneous speech. In the past 15 years, large collections of speech errors have become available and have led to an increased interest in modelling the speech production planning process. A number of models have been proposed that draw on data from those corpora (Baars, 1980; Dell, 1984; Ellis, 1979; Fromkin, 1971; Garrett, 1975, 1984; Leveldt, 1983; MacNeilage, chapter 4; Shattuck-Hufnagel, 1979; Stemberger, 1982). For example, the frequent occurrence of single segment errors supports claims that speech planning involves the representation of individual sublexical elements like phonemic segments, and the commonly-observed feature similarity between pairs of interacting error segments provides clues to the way those units are represented (Fromkin, 1971; MacKay, 1970; Nooteboom, 1967; Shattuck-Hufnagel & Klatt, 1979).

A different kind of similarity also provides information about production planning representations, in this case about the larger structures that govern the processing of individual phonemic segments. This second kind of similarity concerns the contexts of two interacting segments in their respective syllables, morphemes, words, and phrases. For example, several error investigators have noted a position similarity constraint on segmental errors: initial segments tend to interact with other initial segments, and final with final (Boomer & Laver, 1968; Fromkin, 1971; MacKay, 1971; Nooteboom, 1967; Shattuck-Hufnagel, 1983). This observation suggests that larger constituents provide a structural framework for the organization of individual phonemic segments during speech production planning.

The picture is clouded, however, by the fact that lexical stress has also been observed to influence segmental interaction errors, and these two factors (position and stress) often overlap. For example, Boomer and Laver (1968) point out that in their small corpus of errors in English:

> The origin syllable and target syllable of a slip are metrically similar, in that both are salient, or both are weak, with salient-salient pairings predominating.[1] (p. 7)

Leaving the question of target-source context similarity aside, Fromkin (1977) noted that segmental errors occur more frequently in syllables that bear lexical stress than in syllables that do not; in 612 segmental errors of various types from her UCLA corpus, 82% occurred in stressed syllables. At the same time, 73% occurred in word-initial position, leading her to remark:

> This frequency analysis . . . tells little about the . . . role of stress beyond suggesting the preponderance of word-initial stress in English. (p. 19)

Disentangling the effects of position from those of stress is equally difficult in segmental errors from the MIT corpus (Shattuck-Hufnagel, 1983). Most of the errors occur between segments that are both word-initial and prestressed, like "bits and pieces" "pits and bieces." Do these segments interact only because of their intrinsic phonemic similarity, or does contextual similarity play a role as well? If contextual sim-

---

[1] Quite a different view of the role of stress is presented in MacKay (1971). He proposed a model in which segments from stressed syllables, being activated earlier or to a higher level than segments in unstressed syllables, tended to displace segments in earlier unstressed syllables when their level of activation exceeded threshold too early. His proposals are based on a small corpus of within-word exchanges in German. In a similar vein for phrasal stress, Boomer and Laver (1968) note that in their small corpus of mixed types of between-word errors, "In any segmental replacement slip there is a high probability that its origin will be found . . . in the tonic word of the tone-group in which it appears . . ." (p. 8).

These observations raise an interesting question about the difference between (a) a similarity constraint on interacting segments, and (b) a tendency for segments from a "strong" category to replace segments from a different "weak" category. Shattuck-Hufnagel (1979, 1983) has argued that the evidence for similarity constraints on interacting segments is strong and widespread across a number of dimensions for similarity, while the evidence that segments in "strong" categories systematically displace segments in "weak" categories is scattered and at times contradictory. This issue is beyond the scope of the present paper, and will not be dealt with further here.

ilarity plays a role, are consonantal interaction errors influenced by their common word position, or by the shared lexical stress values of their following nuclei?[2] The fact that such a large proportion of words of English carry main stress on the first syllable adds to the difficulty of distinguishing between these two context effects.

The critical cases are errors that occur between words where the initial syllable does not bear stress (i.e., where the influence of stress similarity and word-position similarity would lead to different errors). Some examples of this type in the MIT corpus are shown in Table 1. Although these examples are compatible with the view that a word-position constraint might do most of the work, they are far too scanty to permit the conclusion that there is no effect of lexical stress on segmental interaction errors. Only a set of errors evoked with stimuli that systematically separate these two factors can resolve the question of how best to capture the contextual influences on errors. The experiments described below address this question directly, by comparing elicited error rates for the same set of consonant pairs when they share position versus lexical stress.

TABLE 1. Segmental exchanges that separate word position and lexical stress.

1. math review ↔ *r*ath *m*eview
2. shoulder separation ↔ *s*oulder *sh*eparation
3. cult of personality ↔ *p*ult of *c*ursonality
4. form persuasive ↔ *p*orm *f*ersuasive
5. node of Ranviér ↔ *r*ode of *n*anviér

Experimental examination of contextual effects in general, and of the role of position versus stress in particular, is of interest for several reasons. First, the widespread observation of contextual effects in collections of spontaneous errors supports the claim that larger constituents provide the structure for organizing the planning of individual phonemic segments; experimental confirmation of these effects would support this claim, and raise doubts about any model in which higher levels of structure are not invoked, or have been erased by the time segmental errors occur. Thus the results of such experiments can rule out whole families of possible models.

A second motivation for experimental probing of context effects is the possibility of establishing differences in the way they operate. By pitting position against stress in a way that seldom occurs in spontaneous speech, because of the high proportion of stress-initial words, we can determine whether one or both are at work. If the results support both factors, then we can manipulate the experimental speaking conditions to see whether these changes affect the operation of the two constraints in different ways. If position constraints and stress constraints respond differently to variations in speaking conditions, then it will suggest that the word as a unit, and the lexical stress value of its syllables, are aspects of separate parts of the planning representation. This view can be contrasted with the opposing view that these two factors are characteristics of a single hierarchically-organized representation. Thus the results of experiments on contextual constraints can help to differentiate separate aspects of the planning process.

Finally, by testing the hypothesis that constituent structure constrains errors, and exploring the possibility that different aspects of that structure operate in different ways, we can discover regularities which allow us to make the remaining models more detailed. For example, the 'slots-and-fillers' model proposed in Shattuck-Hufnagel (1979, 1983) includes (a) a set of candidate words (with their phonemic segments), (b) a framework of structural slots defining the utterance, and (c) a mechanism for merging the two representations, but the precise form of these three aspects of the operation is left unspecified. Evidence that points to the word as a unit that guides segmental processing, or to lexical stress as a significant factor at the point where segmental errors occur, will suggest a more detailed specification of the representations and processing mechanisms in that model.

In Experiment 1, tongue twisters are used to address the question of whether the word imposes a position constraint on segmental interaction errors, whether the lexical stress value of the following syllabic nucleus plays a role, or both. Experiment 2 explores what happens to contextual similarity constraints when nonlexical items, rather than existing lexical items, are spoken.

EXPERIMENT 1: WORD POSITION AND LEXICAL STRESS

Stimuli

The error elicitation experiments to be reported here use the technique of presenting a short list of four words to the speaker, who reads the list aloud and then immediately recites it from memory. The words lists are tongue twisters that have been constructed on the model of "she sells sea shells," so that they contain pairs of confusable consonants in alternating patterns that produce many segmental errors. The target segments may appear either (a) in initial position in the word, (b) before the vowel that carries the main stress of the word, (c) both, or (d) neither. Some examples will help make this clear.

In Table 2, four types of tongue twister stimuli are illustrated for each of the consonant pairs r/l, p/f, and d/t. Each stimulus consists of an initial bisyllabic word, a final bisyllabic word, and two monosyllabic words between them. The four types of stimuli can be contrasted in the following way:

1. In stimulus 1, /r/ and /l/ have word position in common but differ in the lexical stress of their following vowels, since /r/ precedes an unstressed vowel and /l/ precedes a stressed one.

2. In stimulus 2, /r/ and /l/ have lexical stress in common but differ in word position, since /r/ is word-medial and /l/ is word-initial.

3. In stimulus 3, /r/ and /l/ have both word position and

---

[2] We say nothing here about interaction errors between vowels, which occur far less frequently than consonant errors and which may obey their own separate constraints. Indeed Sussman (1984) has suggested that vocalic segments are processed separately from consonantal segments, a proposal which would account for the fact that consonants and vowels do not substitute for each other in errors.

TABLE 2. Examples of 4 types of experimental tongue twisters.

Type (1): Word position same		
	r/l	remove lag lick remote
	p/f	parade fad foot parole
	d/t	defy tin talk defend
Type (2): Lexical stress same		
	r/l	morass lag like moraine
	p/f	repeat fad foot repaid
	d/t	foredone tin talk fordoom
Type (3): Both same		
	r/l	rumor lag like roaming
	p/f	peril fad foot parrot
	d/t	duffle tin talk daffy
Type (4): Neither same		
	r/l	moral lag like marry
	p/f	ripple fad foot rapid
	d/t	foredeck tin talk fiddle

lexical stress in common, since both are word-initial and both occur before vowels with main word stress.

4. In stimulus 4, /r/ and /l/ share neither word position nor stress, since /r/ is word-initial and precedes an unstressed nucleus, and /l/ is word-initial and precedes a stressed vowel.

These observations are summarized in Table 2. On the basis of this characterization of the differences among the four types of stimuli, we can make several predictions.

A. If word-position similarity plays a role in constraining segmental interaction errors, then stimulus 3 should evoke more errors than stimulus 2. This is because /r/ and /l/ share lexical stress in both of these stimuli, but differ in word position in stimulus 2. This difference should reduce the number of errors in stimulus 2, if word position similarity is an important determinant of segmental errors. If word position similarity plays no role, types 2 and 3 should provoke similar numbers of errors.

B. If lexical stress similarity plays a role, then stimulus 3 should also evoke more errors than stimulus 1. This is because /l/ and /r/ share word position in both of these stimuli, but differ in lexical stress in stimulus 1. This should reduce the number of errors in stimulus 1 if lexical stress similarity is an important determinant of segmental interactions. If lexical stress similarity plays no role, then types 1 and 3 should evoke similar numbers of errors.

C. If word position is a more powerful influence on segmental errors than lexical stress, then stimulus 1 should provoke more errors than stimulus 2. The reasoning here depends on two differences rather than one. The target segments /r/ and /l/ differ in word position in stimulus 2, which should reduce the number of errors for this stimulus, to the extent that shared word position facilitates segmental errors. In contrast, the two target segments differ in stress in stimulus 1, which should reduce the number of errors for this stimulus if shared stress facilitates segmental errors. Thus, if both contextual constraints are functioning, the relative number of errors in type 1 versus type 2 stimuli will indicate their relative strength.

D. If the simple presence of two confusable segments in the target string is adequate to allow their interaction in an error, then there should be a substantial number of errors for stimulus 4. If on the other hand contextual similarity is necessary in order for interaction errors to occur, then the number of errors elicited by this stimulus should be very small or none.

Note that in all four types of tongue twisters, there is a third initial segment that is not phonemically similar to either member of the target pair. For the r/l stimuli in Table 2, this segment is /m/. This third or "filler" consonant in the bisyllabic words will become of interest later.

Selection of Target Segment Pairs

Twelve pairs of highly-confusable segments were selected from the phonemic confusion matrix that summarizes the 1978 count of single segment interactions in the MIT corpus of errors in spontaneous speech (Shattuck-Hufnagel & Klatt, 1979). The selected pairs were f/p, r/l, b/g, l/y, b/p, m/n, r/w, d/g, p/k, l/n, d/t and ǰ/d. For each pair, four tongue twisters were constructed in the patterns shown in Table 2 above. For some pairs of segments it was possible to construct more than one set of stimuli, so that a total of 24 sets (96 tongue twisters) were used.

Speakers

Twenty speakers were chosen from a pool of normal adult native speakers of English (MIT undergraduates) with no reported history of hearing or speech difficulties. Half were male and half were female. Speakers were paid three dollars per half-hour experimental session.

METHOD

Scoring

The 96 stimuli were divided into two balanced sets of 48 each. Half of the speakers received one set of 48 stimuli in two sessions of 24 each, and the other half received the other set of 48 in two similar sessions. Each half of the speaker population and each set of 48 stimuli was further divided in half, and the order of presentation of subsets of stimuli was counterbalanced across subgroups.

Each twister was typed on a 3 × 5 card for visual presentation to the speaker, who read the stimulus aloud three times and then turned the card over and at once recited it three times from memory, for a total of six utterances. Successive stimulus cards were separated by a card indicating a sentence generation task that was part of a different experiment. This intervening task, which took 15–30s, was quite different in nature from the tongue-twister task, and was included in order to insulate each twister from the effects of the previous one.

The speaker was seated in an acoustically quiet room in front of a microphone connected to a tape recorder. All utterances were tape-recorded and later transcribed and scored for errors of various types. Both transcribing and scoring were done by the investigator, SSH. Each utterance was reviewed at least twice, and utterances containing errors were reviewed up to 15 times to ensure accuracy.

RESULTS

Results for 5,760 utterances of the 96 tongue twisters are shown in Table 3. Data include all of the interaction errors that involved the two target segments: exchanges (counted as two separate errors), anticipatory and perseveratory substitutions, and incomplete errors. Not included in this table are errors that involve segments other than the target pairs (i.e., substitutions of segments not in the stimulus, unexpected interactions between segments in the stimulus, and within-word errors). Also omitted are errors that involve sequences of consecutive segments, like CVC- or -VC, which made up less than 5% of the total errors. How do these results compare with our earlier predictions?

TABLE 3. Results of experiment 1.

Stimulus type	*Number of errors*
Type (1): Word Position —word position same —lexical stress different	121
Type (2): Stress —word position different —lexical stress same	58
Type (3): Both —word position same —lexical stress same	178
Type (4): Neither —word position different —lexical stress different	8

Prediction A: If word position similarity plays a role, then Type 3 stimuli (where the target segments share word position) should evoke more errors than Type 2 stimuli (where the target consonants appear in different word positions.)

This prediction is confirmed: Type 3 stimuli provoke significantly more errors than Type 2 stimuli, $p < .01$. This result makes it difficult to dismiss the word as a significant aspect of the representation that governs segmental processing at the point where individual segmental errors occur. Instead, it suggests a model in which the word plays an active role in the segment planning process.

Prediction B: If lexical stress similarity plays a role, then Type 3 stimuli (where target segments share lexical stress) should also evoke more errors than Type 1 stimuli (where they do not.)

This prediction is also confirmed: Type 3 stimuli evoke significantly more errors than Type 1 stimuli, $p < .01$. This result suggests that lexical stress similarity, like word position similarity, has an influence on segmental errors.

The implication of these results taken together is that the structure of larger constituents does indeed play a role in segmental processing. Apparently, word-sized units as well as the lexical stress patterns of those words are significant aspects of the representation that guides processing, at the point where segmental errors occur.

Prediction C: If word position similarity is a more powerful influence on segmental errors than lexical stress similarity, then Type 1 stimuli (where targets share word position but not stress) should evoke more errors than Type 2 stimuli (where targets share stress but not word position.)

This prediction is confirmed; Type 1 stimuli provoke significantly more errors than Type 2 stimuli, $p < .01$. In other words, shared word position without shared stress provokes more errors than shared stress without shared word position. Apparently, although both kinds of contextual similarity function to constrain segmental errors, word position similarity is more powerful than shared lexical stress.

Prediction D: If both word position and lexical stress impose constraints on segmental interaction errors, then Type 4 stimuli should produce fewer errors than any of the other three types.

This prediction is confirmed; the number of target interaction errors in Type 4 stimuli is extremely small, only 2% of the total errors and .5% of the opportunities for target interactions in Type 4 stimuli. The implication of this result is that the simple occurrence of two confusable segments in the target string is usually insufficient to permit those two segments to interact in an error; some degree of contextual similarity is necessary before an error can occur.

Unexpected Errors

Unexpected evidence that bears on these predictions is also provided by a different set of errors, a set that was not envisioned when the experiment was designed. These errors involve an interaction between the initial target consonant in the monosyllables and the phonemically-dissimilar filler consonant in the bisyllabic words. The pairs of segments involved in these unexpected errors appear in the same four configurations as the original target pairs: (a) they share word position, (b) they share lexical stress, (c) they share both or (d) they share neither. However, the stimulus types that correspond to these configurations are reversed. The new type labels for the unexpected interactions are shown in Table 4.

The results for the smaller number of unexpected errors show a pattern similar to those for the original target pairs, as shown in Table 5, where the earlier results for expected errors are reproduced for comparison. Again, errors occur between segments that share either word position or stress, with word position errors more frequent. Sharing both contextual characteristics leads to even more errors, and sharing neither dimension results in few or zero errors.

TABLE 4. Reversed stimulus categories for unexpected errors.

Sample stimulus	*Type for expected r/l error*	*Type for unexpected m/l error*
remove lag lick remote	Word position	Stress
morass lag lick moraine	Stress	Word position
rumor lag lick roaming	Both	Neither
moral lag lick marry	Neither	Both

TABLE 5. Results for unexpected errors in experiment 1.

Stimulus type	*Expected errors*	*Unexpected errors*
Type (1): Word position	121	24
Type (2): Lexical stress	58	11
Type (3): Both	178	30
Type (4): Neither	8	—

The segment pairs in these unexpected errors were p/g, k/y, b/g, d/p, m/l, r/n, l/p, p/n, l/f, f/t, d/p, b/r, g/l, f/b, ǰ/r, l/d, r/g, b/w, l/w, r/y, p/l, r/f, b/n and m/p. These pairs are less similar in terms of shared features than the original target pairs; using a simple three-way classification of manner, place and voicing, the original 24 pairs (i.e., 12 separate pairs, with multiple use of some pairs) differed by an average of only 1.2 features, while the 24 pairs in this set differ by an average of 2.3 features out of 3. This indicates that contextual similarity is able to pull two target segments into an interaction error even when the two segments are relatively dissimilar. The resemblance in the pattern of results for the four stimulus types for both phonologically similar and phonologically dissimilar pairs suggests that the two factors (word position and lexical stress) are extremely powerful determinants of segmental errors.[3]

Do the Two Contextual Similarity Constraints Interact?

Comparison of the number of errors for the four types of stimuli in this experiment shows that the number of errors in Type 3 (i.e., stimuli where target segments share both word position and lexical stress) is no larger than would be expected if the two factors were operating independently. Type 3 stimuli elicited 178 errors, very close to the 174 errors which are predicted from the error rates for the two factors of word position and stress when they operate separately in stimulus types 1 and 2. The fact that there is no facilitation or interaction when both similarity conditions are satisfied is compatible with the hypothesis that the two constraints arise from separate aspects of the representation used by the speaker during segmental planning. That is, the word may be a structural unit for one aspect of the representation, while the lexical stress patterns are actively represented in another part of the representation. To test this possibility further, we varied the nature of the spoken stimuli to see if the change would affect the two similarity constraints in different ways. The dimension we chose to vary was the lexical status or "wordhood" of the stimulus items.

This choice was motivated by two considerations. First, it was our hypothesis that the word-position similarity between interacting segments might arise during the word-oriented process of accessing the phonemic shapes of lexical items for production. It seemed reasonable that lexical access for production might be the locus of a word-based constraint on errors. If this is so, the effect should disappear for nonwords, since they are not stored in the mental lexicon and therefore presumably do not undergo lexical access. In contrast, we predicted that the lexical stress similarity effect would not disappear for nonword stimuli, since the lack of interaction in Experiment 1 suggested that the two constraints arise from different aspects of the processing representation.

These two lines of reasoning were the genesis of Experiment 2—first, the search for further evidence that the two contextual similarity effects arise separately, and second, a test of the prediction that the word-position similarity constraint should disappear for nonwords. This prediction follows from the claim that the word-position constraint arises during the accessing of lexically-stored information for production.

EXPERIMENT 2: CONSTRAINTS ON ERRORS IN REAL WORDS VERSUS NONEXISTING WORDS

One possible interpretation of the word-position finding is that many segmental errors occur during the process of locating the phonemic shapes of candidate lexical items in the mental lexicon, and retrieving them. It is clear that a word-retrieval process must take place, and it has been hypothesized that word-initial elements might play a privileged role in that process. If the proclivity of word-initial segments to interact in errors arises because of their special role in lexical access, then nonword stimuli should not provoke this same error pattern. In other words, if the word-position constraint is imposed by the nature of the lexical access process, then it should disappear in sets of errors that occur in nonsense items, since we assume that those items cannot undergo lexical retrieval. To test this prediction, a set of nonword stimuli were generated in the following way.

Stimuli

The 16 most error-prone quadruples of tongue twisters were chosen from the original set of 24 used in Experiment 1. For these 16, the Type 3 and Type 4 stimuli were discarded, leaving only the Type 1 (word position same, stress different) and Type 2 (word position different, lexical stress same) stimuli, or 32 twisters made up of real words. The corresponding nonword twisters were formed by denuding these 32 stimuli of their vowels, and inserting new vowels to produce structurally legitimate nonwords with the same consonant patterns. For example, the original stimulus twister "July dog dock Gillette" became "Jelaye deg dack Jelotte." The 32 original real-word stimuli and 32 new nonword stimuli formed the set of 64 stimuli used in Experiment 2.

METHOD

Speakers

The resulting set of 64 twisters was divided into two bal-

---

[3] Several control analyses rule out the possibility of accounting for these results in other ways. For example, the word position effect is significant across subjects ($p \leq .01$), suggesting that it is not due to a few special speakers. Similarly, it is significant across target pairs ($p \leq .01$), suggesting that it is not the result of a large asymmetry in a small number of atypical stimuli.

anced sets of 32 each. Ten new speakers spoke the stimuli in two experimental sessions, with presentation conditions and intervening sentence generation tasks like those in Experiment 1, and the utterances were similarly tape recorded, transcribed and scored.

RESULTS

The results of this realword-nonword comparison pose some problems for a lexical-access account of the original finding that word position is a significant determinant of segmental interaction errors. The word-position similarity effect is replicated for both the original real word and the new nonword stimuli, but this direct comparison of real words with nonwords shows that the effect is significantly stronger for nonwords. As Table 6 shows, the difference between words and nonwords in the errors provoked by Type 1 stimuli is highly significant, while there is no difference for Type 2 stimuli. Rather than eliminating the word position similarity constraint, the use of nonword stimuli causes even more interaction errors between word-initial segments, while the number of errors between prestressed consonant remains the same. This finding holds for the smaller number of unexpected errors as well: the number of word-initial errors is greater for nonwords than for real words, but the number of prestressed errors stays about the same. These observations suggest two points:

TABLE 6. Results for experiment 2, nonword versus real word stimuli.

	Expected errors		*Unexpected errors*	
Stimulus type	*Real words*	*Nonwords*	*Real words*	*Nonwords*
Type (1): Word position	118	230	27	83
Type (2): Stress	85	84	7	4

1. The increased incidence of errors between position-sharing segments for nonwords does not come about because of a general difficulty with these stimuli, since the prestress segments escape any increase in error rate, and
2. We have a second piece of evidence to support the claim that the two contextual constraints on segmental errors, word-position similarity and stress similarity, arise from different aspects of the representation that guides the production processing of individual phonemic segments. Not only do these two factors fail to interact with each other, as shown by the results of Experiment 1, but they are differentially sensitive to lexical status. The word-position constraint is sensitive to the lexical status of the material to be spoken, while the stress constraint is not.

How can these results be accounted for under the claim that the word-position similarity constraint is related in some way to the process of lexical access or transfer of information from the lexicon? It is possible that the speaker begins to formulate a lexical representation of each nonword as the stimulus is presented for the first time. There is some evidence from lexical decision tasks that speakers may treat nonwords that they have seen only once as a kind of lexical item on subsequent exposures (Salasoo, Shiffrin, & Feustel, 1985). If speakers do form a kind of intermediate lexical representation of a nonword on first encountering it, these embryonic representations might be particularly susceptible to the kinds of errors that normally occur during lexical retrieval (i.e. to errors between pairs of word-initial segments). This would account for the striking increase in word-initial segments for the nonword stimuli in this experiment.

On the other hand, it is possible that the results for nonwords indicate that the word-position similarity constraint does not operate during lexical access, but rather during some part of the process that (a) applies to words and nonwords alike, and (b) is more susceptible to error when it deals with nonwords. For example, if candidate words (or nonwords) for an utterance must be transferred into a temporary utterance-specific store that is organized as a list of words, and segments are read out of this word-based list as they are integrated into another framework that encompasses a larger constituent like a phrase, then the word-position similarity constraint on segmental interaction errors might arise during some part of this process. The word-position constraint suggests strongly that the word plays an active role in segmental processing, but this may occur in one of several ways. Unfortunately the results of Experiment 2 do not distinguish among the several possibilities.

DISCUSSION

The results of Experiments 1 and 2 demonstrate that, for these stimuli at least, models of segmental processing that limit their representations to a simple string of target segments must be rejected. Instead, some account must be taken of the powerful contextual constraints that govern segmental errors, like the position of the consonants in their respective words, and the lexical stress value of the following syllabic nuclei. In both expected and unexpected errors, word position appears to be the more powerful of the two contextual constraints. The fact there was no evidence for an interaction between them suggests that these two factors exert their influence separately (i.e., that they may be characteristics of two separate aspects of processing representations). This conclusion is further supported by the finding in Experiment 2 that the word-position factor is sensitive to the lexical status of elements in an utterance, while the stress factor is not.

Implications for a Segmental Processing Model

In terms of the slots-and-fillers model proposed in Shattuck-Hufnagel (1979, 1983), these results offer some support for the model and also suggest several ways in which the characterization of the segmental processing mechanism can be made more specific. First, the experimental confirmation of contextual similarity constraints is compatible with the claim that errors occur by mis-selection between similar segments, rather than by displacement of weak segments by strong ones.

Second, the establishment of both word position and lexical stress as contextual factors constraining segmental errors imposes the requirement that both of these aspects of larger constituent structure be represented during segmental processing. The evidence that they operate separately suggests that the word-position of a consonant and the stress value of its following syllabic nucleus may be represented in different parts of the processing representation. These requirements can be satisfied by positing that the phonemic segments of the candidate words for an utterance are placed in a buffer that is organized in terms of word-sized units, and the lexical stress pattern is imposed on the elements in this buffer at a later processing stage. On this view, the word-position constraint might arise during the transfer of candidate lexical items into the buffer, and the stress constraint during the integration of this word-based representation into a larger framework, perhaps representing the structure of a phrase.

Further Questions

One question that arises concerns the generality of these results for other kinds of stimulus utterances, particularly because differences have been reported between the errors elicited by list utterances and those elicited by utterances that are grammatically well-formed phrases (Shattuck-Hufnagel, 1982). Pilot results indicate that errors in phrasal utterances show the effects of both the word-position and lexical stress constraints; pilot results also show that the same pattern is observed when the errors are elicited in a sentence generation task. The effects observed in Experiments 1 and 2 appear to be robust across a variety of speaking conditions. As these experiments are completed, we hope to learn more about possible differences in contextual constraints, at a more detailed level of analysis, on errors across different speaking conditions.

A second question that is not dealt with here concerns the role of a possible syllable structure constraint. A number of investigators have expressed the position constraint on segmental errors in spontaneous speech in terms of syllable structure, noting that onset consonants interchange with other onset consonants, nuclei with other nuclei and coda consonants with other coda consonants. The word-position constraint established by the experiments reported here makes it necessary to re-examine these claims. For example, in a set of 191 complete and unambiguous between-word consonant exchanges in the MIT corpus, 183 occurred between segments that shared word position (166 initial and 17 final, ignoring inflectional affixes). Only two 2 exchanges occurred between word-medial consonants and 6 between consonants that crossed word position. Because segments that share initial or final word position also share syllable position, it is possible to capture this position similarity constraint in terms of syllable structure. However, given the evidence for the word as a powerful source of contextual similarity between interacting error segments, one would need to see compelling evidence before adding the representation of syllable structure to the requirements for the processing representation at the point where segmental errors occur. Detailed analyses of both spontaneous and experimentally elicited errors presently under way should shed more light on this problem, and provide further information about the processing that underlies the production of normal fluent speech. It is to be hoped that information of this kind will be useful in the search for more effective speech training methods.

ACKNOWLEDGMENTS

This work was supported by Grant No. NS-04332 from the National Institute of Neurological and Communicative Disorders and Stroke. Thanks are due to W. J. M. Leveldt for helpful comments.

REFERENCES

BAARS, B. J. (1980). The competing plans hypothesis: An heuristic viewpoint on the causes of errors in speech. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech*. The Hague: Mouton.

BOOMER, D. S., & LAVER, J. D. M. (1968). Slips of the tongue. *British Journal of Disorders of Communication, 3*, 1–12.

DELL, G. S. (1984). *A spreading activation theory of retrieval in sentence production*. University of Rochester Cognitive Science Technical Report. No. 21.

ELLIS, A. W. (1979). Speech production and short-term memory. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics series Vol. 2, Structures and processes*. London: Elek & Cambridge, MA: MIT Press.

FROMKIN, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47*, 27–52.

FROMKIN, V. A. (1977). Putting the emPHAsis on the wrong syllAble. In L. A. Hyman (Ed.), *Studies in stress and accent* (pp. 15–26). Southern California Occasional Papers in Linguistics No. 4, Dept. of Linguistics, University of California at Los Angeles.

GARRETT, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 133–177). New York: Academic Press.

GARRETT, M. F. (1984). The organization of processing structure for language production: Applications to aphasic speech. In D. Caplan, A. R. Lecours, & A. Smith (Eds.), *Biological perspectives on language*. Cambridge, MA: MIT Press.

LEVELDT, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41–104.

MACKAY, D. G. (1971). Stress pre-entry in motor systems. *American Journal of Psychology, 84*, 35–51.

NOOTEBOOM, S. G. (1967). *Some regularities in phonemic speech errors*. Instituut voor Perceptie Onderzoek Annual Progress Report 2, pp. 65–70.

SALASOO, A., SHIFFRIN, R. M., & FEUSTEL, T. C. (1985). Building permanent memory codes: Codification and repetition effects in word identification. *Journal of Experimental Psychology, 114*, 50–77.

SHATTUCK-HUFNAGEL, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing* (p. 295–342). Hillsdale, NJ: Lawrence Erlbaum.

SHATTUCK-HUFNAGEL, S. (1982). Position of errors in tongue twisters and spontaneous speech: Evidence for two processing mechanisms? Speech group working papers: Research laboratory of electronics, Vol. 1, (pp. 1–8). Cambridge, MA: MIT.

SHATTUCK-HUFNAGEL, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 109–136). New York: Springer-Verlag.

SHATTUCK-HUFNAGEL, S., & KLATT, D. H. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior, 18*, 41–55.

STEMBERGER, J. P. (1982). *The lexicon in a model of language production*. Unpublished doctoral dissertation, University of California, San Diego.

SUSSMAN, H. M. (1984). A neuronal model for syllable representation. *Brain and Language, 22*, 167–177.

Chapter 9

THE ROLE OF PRODUCTION VARIABILITY IN NORMAL AND DEVIANT DEVELOPING SPEECH

KATHERINE S. HARRIS
JUDITH RUBIN-SPITZ
NANCY S. MCGARR

Center for Research in Speech and Hearing Sciences
Graduate School and University Center,
City University of New York
and
Haskins Laboratories
New Haven, CT

The idea of an underlying structure which is given some kind of imperfect surface manifestation is, of course, a rather common one in description of behavioral phenomena in general, and linguistic systems in particular. Following the lead of Jakobson's (1968) famous monograph investigations of child language have been couched in terms of underlying phonological systems, related to a child's phonetic output by rewrite rules, like the rules governing morphophonemic alternations in adult speech. Thus, a child who omits the final /g/ in the word "dog," but will produce the diminutive "doggie" may be described as having an underlying representation which includes the /g/, with a rule which deletes it in syllable-final position.

Many scholars, notably Smith (1973) and Ingram (1976), have asserted that the underlying phonology of normal children at the time of beginning vocabulary development is that of the ambient community. This belief rests in part on old anecdotal evidence that children often can recognize words which they cannot produce, and, in part, on more recent evidence regarding the ability of infants to discriminate differing speech sounds (Eimas, 1982). However, as Studdert-Kennedy (in press) points out "I do not doubt that infants can form auditory categories, but there is no evidence that this capacity is either needed for or brought to bear on early speaking."

Much the same view of the relationship of two levels is often taken of the underlying phonology in functionally misarticulating children. For a history of the use of phonological process analysis within speech-language pathology (see Edwards & Shriberg, 1983). That is, it has often been assumed that the misarticulating child has a normal underlying perceptual process, but obeys rule-governed restrictions in output.

Recently, Elbert, Dinssen, and Weismer (1984) and Maxwell (1979) have suggested that misarticulating children differ among themselves in the relationship of underlying and surface forms. While some children give evidence, either by the presence of morphophonemic alternations (e.g., /dɔ/ but /dɔgɪ/) or by preservation of acoustic differences in output for two forms in which a phone is omitted in transcriptional description, others do not. These authors suggest, therefore, that the nature of a child's phonological structure should be demonstrated on a phone-by-phone basis, rather than assumed.

It is possible to take the more radical position that description of children's early word attempts might be couched in auditory and motoric rather than linguistic terms (Studdert-Kennedy, in press). After all, it is not necessary to assume that the child has internalized phonological categories which conform to the description of adult linguistic behavior (Harris, 1983; Menn, 1980; Menyuk & Menn, 1979). The fact that transcription has been the method of choice for describing children's production has tended to push description toward adult categories. However, Ferguson has presented evidence that early words are learned on a one-by-one basis (Ferguson & Farwell, 1975) and that attempts at an early word are highly variable. While it is extremely difficult to abandon the transcriptional description of words, even transcriptions show that ubiquitous variability is an essential component of the description of the child's categories.

This same variability has been repeatedly shown in instrumental descriptions as characteristic of the speech of children, even when they produce apparently mature forms (Kent, 1976). Eguchi and Hirsh (1969) described the spectral variability of production of vowels in children's speech. While the extent to which their data were affected by measurement error has been the subject of some discussion (Monsen & Engebretson, 1983), there seems to be little doubt about the appropriateness of Eguchi and Hirsh's characterization of the variability phenomenon itself. Similar production variability has been shown to characterize temporal aspects of developing speech production capabilities (see Smith, 1978).

We emerge, then, from the description of normal child phonology with two general principles. First, a phonological inventory description must be supported by production data of some sort which demonstrates the differentiation of units which are presumed to be phonologically distinct. Often, forms distinct in the adult model are collapsed in the child's output, or are differentiated on a basis which is different from the adult. Second, it may be that the description of a child's speech in terms of an underlying phonological structure fails to capture at least the important variability aspect of performance.

When we turn to deaf children, we find that the same kind of phonological structure approach has been used in describing their speech, especially by Monsen (1976, 1983) and by Fisher, King, Parker, and Wright (1983). For hearing-impaired children there is, of course, no question that the representations supporting the phonological structure must be very different from that of the hearing community, since we presume that the sensory information on which such children base any structure and maintain differentiation between items is very different from that for normals. Thus, in Fisher's et al. (1983) description, a single form is produced by deaf children for forms which are differentiated in the adult model, or a given contrast, while preserved, is preserved in phonetically different terms. One of the most interesting points made by Fisher and his colleagues (op. cit.) is that intelligibility for those deaf speakers who maintain a system of deviant contrasts may be reduced by a speech training regime which moves some phones towards the normal model, but removes certain contrasts that are preserved on a deviant basis.

What kind of evidence might be marshalled in support of the point of view that the oral deaf preserve contrasts between phones as normals do? We can examine, carefully and systematically, the variability of production of some class of sounds. A deviant phonology would be indicated by normal production variability, co-occurring with a failure to differentiate pairs of sounds, or an abnormally based distinction.

An indirect form of evidence for the "deviant phonology" hypothesis could be provided by the listener effect, an effect which has been investigated by several researchers at the Central Institute for the Deaf. If deaf speakers differentiate between sounds in production in a way that is different from normal, then teachers, who are experienced in listening to hearing-impaired talkers, might be able to invoke a special listening strategy, based on the use of cues which naïve listeners ignore. For example, if it were true that some deaf speakers systematically substitute fundamental frequency variation for formant variation (Angelocci, Kopp, & Holbrook, 1964), then an experienced listener might simply focus on this characteristic as a way of differentiating vowels (or classes of vowels). The listeners would then show a heavier dependence on F_0 than on spectral characteristics of individual tokens. Alternatively, if deaf speakers simply overlay some abnormal characteristic (Stevens, Nickerson, & Rollins, 1983), such as too high or too low pitch on their speech, experienced listeners might learn to ignore the deviant overlay, and focus on vowel cues. In this case, the pattern of differentiation would be the same for experienced and inexperienced listeners, although experienced listeners would show superior performance.

An essential component of the listener effect is that listeners must be able to identify speakers as deaf. Some time ago, Calvert (1961) demonstrated very convincingly that experienced teachers of the deaf can identify speakers as deaf, but that the teachers' performance depends very heavily on the presence of articulatory movement in the samples judged—that is, the time-dependent deviance of deaf articulatory patterns is detectable, and hence, might serve as the basis of a detection strategy. Moreover, the fact that sustained vowels produced by deaf talkers are less readily identified than vowels produced in context suggests that such identification does not depend on an overlaid characteristic, such as voice quality.

In what follows, we will discuss three studies that bear on the issues above. The first is a doctoral dissertation by Judith Rubin (1984). Obviously, there is a great deal more detail in her study than can be reported here. We will then go on to discuss some physiological work on interarticulator timing in the productions of deaf talkers (McGarr & Gelfer, 1983; McGarr & Harris, 1983; McGarr & Löfqvist, 1982) and also in normal speakers (Harris, Tuller, & Kelso, 1984; Tuller & Kelso, 1984; Tuller, Kelso, & Harris, 1982, 1983).

The object of Rubin's study was first, to make a direct test of the hypothesis that deaf speakers produce vowels with the same variability as normal talkers. Beyond that, she wanted to compare the strategies that experienced and inexperienced listeners use in decoding deaf and normal vowels.

The subjects of her study were six orally trained, severely or profoundly hearing-impaired high school students and two age-matched normals. The speakers were asked to say "You got me the /bvb/" with any of seven test vowels in the vowel slot. Each token was produced 15 times. The results were analyzed acoustically, using an LPC algorithm; F_0, F_1, F_2 and duration were measured.

In the perceptual part of the study, experienced and inexperienced listeners were asked to make two judgments—first, they were asked to identify each vowel token as to whether it was produced by a deaf or a normal talker. Second, they were asked to identify the vowel. Stimuli were presented in three conditions—first, the whole utterance; second, the /bvb/ syllable alone; and third, a short, more-or-less steady state segment gated out of the middle of the /bvb/ syllable. The stimuli were grouped by condition, but not by speaker.

We will first describe the results of the acoustic formant analysis. First, on average, deaf talkers show a reduced range of average F_1 and F_2 values, relative to normals—durations are prolonged as has been previously reported, and fundamental frequency is a little higher on average. (Note that the talkers were preselected to avoid subjects with such severe source problems that LPC analysis would become problematic). However, when we look at individual talkers, comparing mean plots and variability plots, a different and more complicated picture emerges.

While individual differences are not discussed here in detail, some of the speakers showed small variability for the point vowels (/i/, /ɑ/, and /u/), with much greater variability for intermediate vowels such as /ɛ/. Some showed overlap between front and back vowels while some showed a great deal of variability for all vowels. Thus the placement of the average values in F_1-by-F_2 space does not predict the relative variability of the tokens around average values.

This point is illustrated in the average data for two hearing-impaired speakers. Average vowels for the first speaker shown in Figure 1 are more or less appropriately distributed in formant space.

In Figure 2, the ranges of the tokens for the same speaker are shown by adding lines drawn to enclose the points representing all tokens. For this speaker, the three point vowels /i, a, u/ are reasonably well defined; however, intermediate vowels are much more variable.

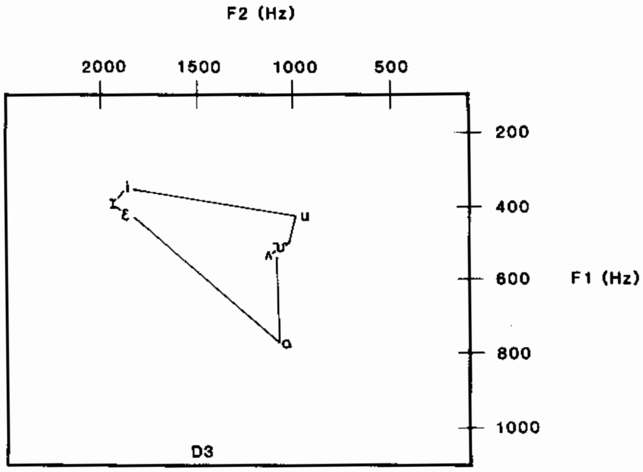


Figure 1. Average vowels for Talker D3.

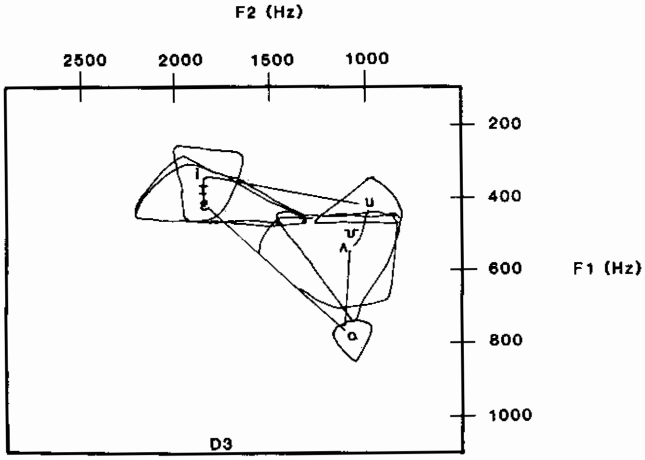


Figure 2. Range of vowels for Talker D3.

Average values for a second deaf speaker are similar to those for the first, as shown in Figure 3, but when we examine the distribution around the average values, as shown in Figure 4, we find a great deal of smear for all vowels. That is, the average values do not give a clear picture of the token-to-token variability.

Figure 5 shows the standard deviations of F_1 and F_2 for the six talkers, while Figure 6 shows standard deviations for the four acoustic measures summarized in a somewhat different fashion. The important point here is that deaf talkers are statistically significantly more variable than normals on every acoustic dimension. Thus, a description of average formant values fails to capture differences between vowel systems.

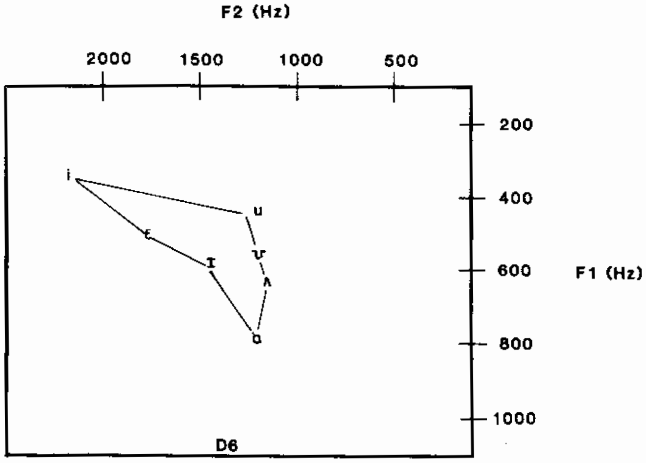


Figure 3. Average vowels for Talker D6.

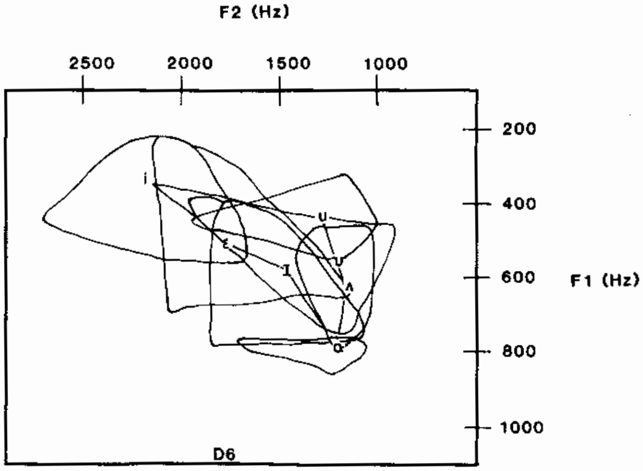


Figure 4. Range of vowels for Talker D6.

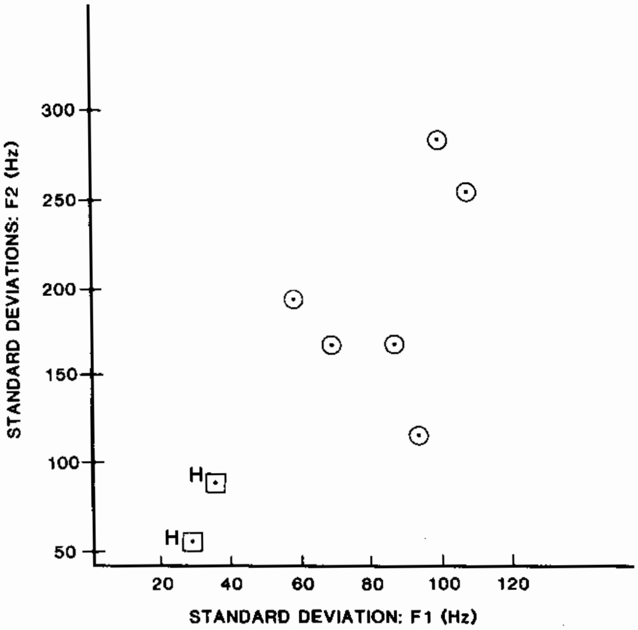


Figure 5. Standard deviations of F_1 vs. F_2 for all subjects.

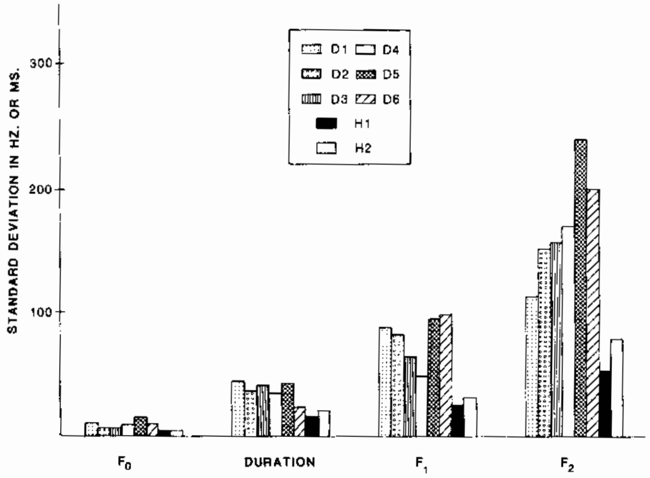


FIGURE 6. Standard deviations of F_0, duration, F_1 and F_2 for all subjects.

There remains the possibility that hearing-impaired talkers were using F_0, or duration, alone or in combination with F_1 and F_2 in their attempt to discriminate between vowels. This possibility was checked by comparing two linear discriminant analyses, to see how many vowel targets can be discriminated using F_0 and duration, which were not discriminated by F_1 and F_2 alone. We find that for the most part, adding F_0 and duration information does not change the number of vowels which can be discriminated statistically, on a talker by talker basis. This provides additional support for Bush's (1981) finding that deaf talkers do not substitute F_0 differentiation for formant differentiation in vowel production.

Finally we turn to the perceptual part of the study. As we discussed above, a strong listener effect would be indirect evidence suggesting that deaf unintelligibility is, in part, due to a systematic, but deviant production strategy.

As Figure 7 shows, there was no statistically significant difference between experienced and inexperienced listeners. The listener effect for vowel identification has been reported by McGarr and Gelfer (1983), but not by Gulian and Hinds (1981). A listener effect for word identification has been found by Mangan (1961), Markides (1970), McGarr (1978), Nickerson (1973), and Thomas (1963).

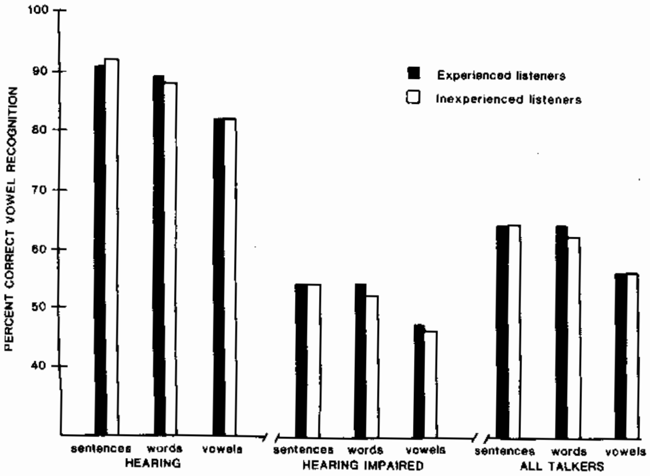


FIGURE 7. Effects of context on vowel recognition by experienced and inexperienced listeners, listening to hearing and hearing impaired talkers.

Let us turn now to an examination of the effects of context. While the effects of context on vowel identification in normals has been the subject of debate in voluminous literature (see Ochiai & Fujimura, 1971; Pisoni, Carrell, & Simnick, 1979; Verbrugge, Strange, Shankweiler, & Edman, 1976), studies have at least suggested that phonetic context aids in recognition. That is the case here. Listeners, whether experienced or inexperienced, were most successful with sentences and syllables and least successful with gated segments excised from the vowel.

Context also was important in the other judgment the listeners made, that is, whether the speaker was deaf or normal. Since there were two normal and six hearing-impaired speakers in the study, d' was used as a measure of the ability of listeners to identify the speakers as hearing or deaf, as shown in Figure 8. Again, the effects of experience were minimal. However, the listeners were increasingly correct in judging the speaker to be deaf as they had more dynamic information. This result qualitatively confirms Calvert's thesis result (1961). However, at a quantitative level, listeners in the present study could be shown to behave statistically slightly above

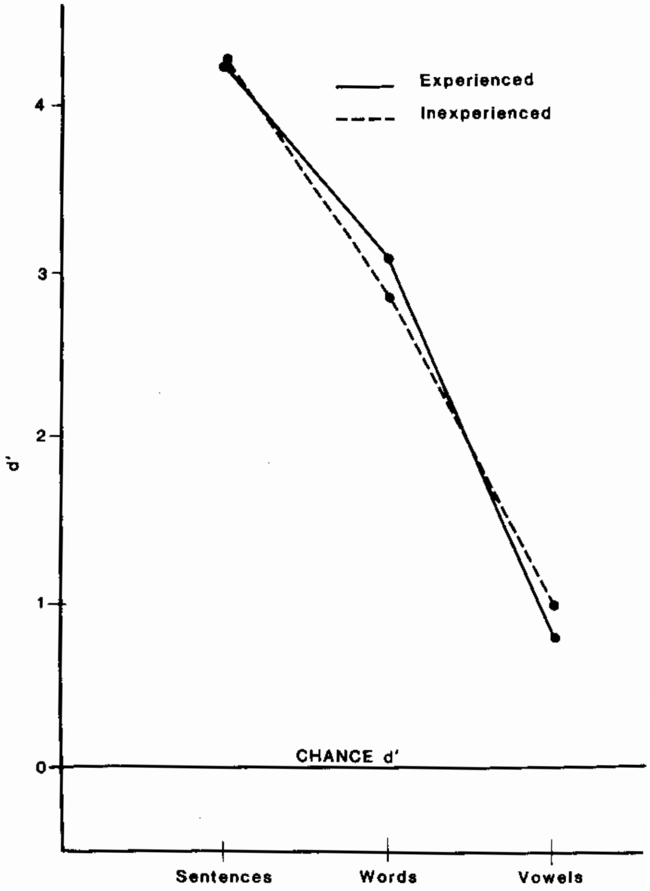


FIGURE 8. d' for vowels in various contexts.

chance levels in judging even isolated vowels. The ability of listeners to judge a vowel correctly was statistically independent of their ability to judge it as produced by a hearing or deaf child, whether the listener was experienced or inexperienced. This result again suggests that there is no special strategy which is effective in decoding deaf vowels.

Still another analysis was made of whether listeners were using conventional information in making vowel identity judgments for deaf talkers. Figures 9 and 10 show the acoustic data for the two individual deaf talkers discussed earlier, with circles around those vowels which are judged correctly at least 70% of the time. The effect of context is to enlarge the "correct vowel" area. Thus, we can speculate that placing a vowel within a consonant transition context allows the listener to be less dependent on precisely appropriate specification of vowel formant information.

Let us summarize these results, and go on to say a bit about production. First, these analyses fail to provide any evidence that deaf speakers were using a substitution strategy in vowel production, or that experienced listeners were better than inexperienced, because of a different way of judging deaf speech. Deaf speakers were more variable than normals, although the pattern of variability was different from talker to talker. One interpretation of the results presented is that it is not appropriate to describe these talkers as presenting a deviant phonology. Indeed, we would argue that a "deviant phonology" description of their production does not capture essential aspects of their performance. The results we have seen for these children suggest that they are behaving, in a more extreme way, like normal children, as Kent (1976) describes them. Performance variability is an essential characteristic of all the speech of children as they learn to talk, and as they attain control of the production apparatus.

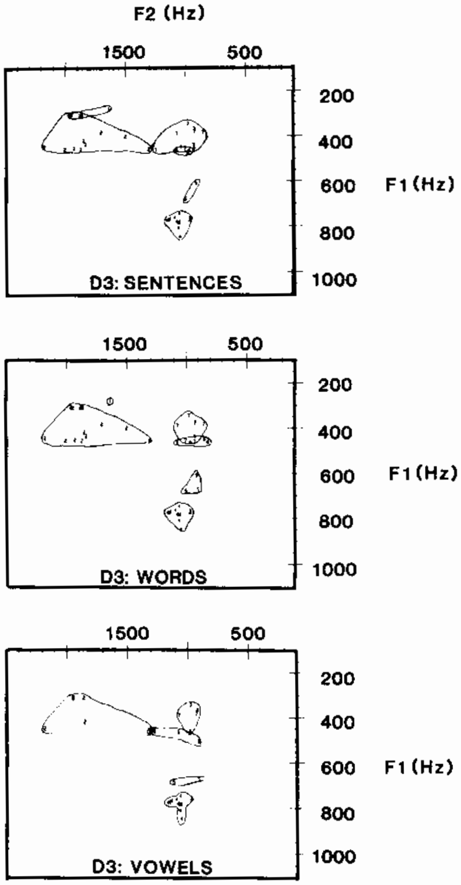


FIGURE 9. $F_1 \times F_2$ plots of vowel tokens perceived correctly in the three experimental contexts, for Speaker D3.

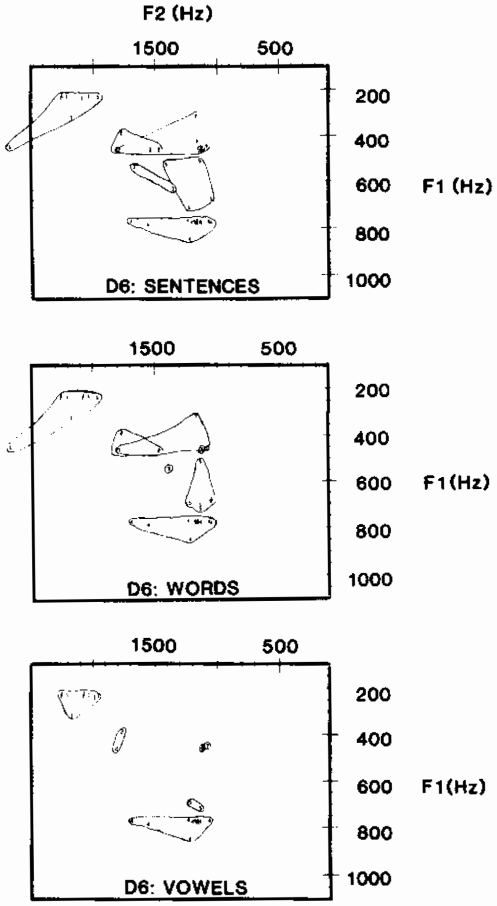


FIGURE 10. $F_1 \times F_2$ plots of vowel tokens perceived correctly in the three experimental contexts, for Speaker D6.

The nature of the articulator routines underlying the variability in acoustic output is unresolved by the study just described. However, we might note that the sequence of upper articulator movements in producing the utterance /bvb/ is fairly simple. The subject closes the lips for the initial and terminal bilabial consonants, and between these two gestures, s/he must produce an appropriate tongue configuration. If these gestures are produced in an inappropriately timed sequence, the acoustic result will be inappropriate, but the consequences of changing the relative timing of the gesture sequence is not directly represented in the acoustic signal.

One of the observations made by Ferguson and Farwell (1975) was that a normal child, in attempting to produce the word "pen," engaged in attempts which were variable precisely because she did not output the required sequence of articulatory gestures in the correct order. We believe that the characteristic variability in deaf speech may arise in part from the same sources (cf. McGarr & Gelfer, 1983; McGarr & Harris, 1983; McGarr & Löfqvist, 1982).

We illustrate this point with data from a tongue-tip coordination study of McGarr and Harris (1983) in which stimuli not unlike Rubin's, (i.e., a bilabial-V-bilabial sequence) were used. Articulatory timing was monitored by electromyographic techniques. When muscle fibers contract, a change in potential is generated in the surrounding medium and these changes in potential can be measured by appropriately placed electrodes. Lip closure (e.g., in bilabial production) is accomplished in part by the contraction of the orbicularis oris muscle, a muscle whose fibers ring the lips. For production of a high vowel such as [i], the tongue body is bunched and raised by contraction of the genioglossus, a muscle whose fibers radiate through the center of the tongue mass. The EMG record then indicates gesture sequencing.

Results are shown for a hearing speaker producing the utterance /əpɑpip/ in Figure 11. These data represent the ensemble average of about 20 repetitions or tokens of each utterance, with each token on the average showing essentially the same pattern of activity (see Harris & McGarr, 1980; McGarr & Harris, 1983). The line-up point, indicated by the vertical line at 0 ms, is the release burst of the second /p/. The data for the orbicularis oris (00) show three well-defined peaks of activity corresponding to the lip gestures for the three /p/ closures in /papip/. The line-up point falls between the second and third peaks. For the genioglossus (GG), there is a peak of activity associated with /i/ but not /a/, because genioglossus is active in raising and bunching the tongue. Peak genioglossus activity occurs approximately at the acoustic line-up. This is not surprising because EMG activity typically precedes the articulatory event to which it is attached by about 50 to 100 ms. Shifting of stress from the first (Figure 11A) to the second vowel (Figure 11B) does not disrupt this temporal relationship.

Figure 12 shows similar data for an oral deaf adult. The EMG pattern for 00 shows, as for the hearing subject, three well-defined peaks of activity. The duration of the peaks is prolonged, however. In Figure 12A, peak GG activity occurs between the second and third orbicularis oris peaks but is late relative to the acoustic event. This pattern was most like normal. In Figure 12B the GG activity was too late. In Figure 12C, activity begins during what should be /a/ production, when the GG should be silent. Thus, the EMG pattern for GG is quite variable from token to token. This variability is reflected in a less well-defined average pattern (see McGarr & Harris, 1983, for more details).

While this evidence is fragmentary, it suggests precisely the sort of production variability we might expect; that is, while the behavior of a visible articulator is more or less normal, activity for one of the muscles associated with tongue movement is variable in its temporal alignment with the ac-

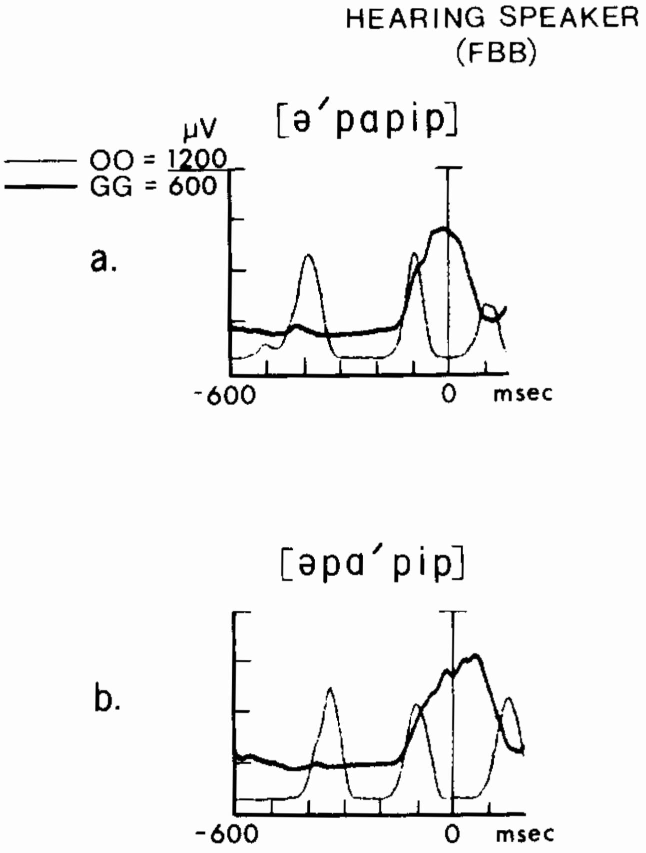


FIGURE 11. Average 00 and GG outputs as a function of time, of simple nonsense utterances, for a normal talker. (Reproduced with permission from McGarr & Harris, 1983.)

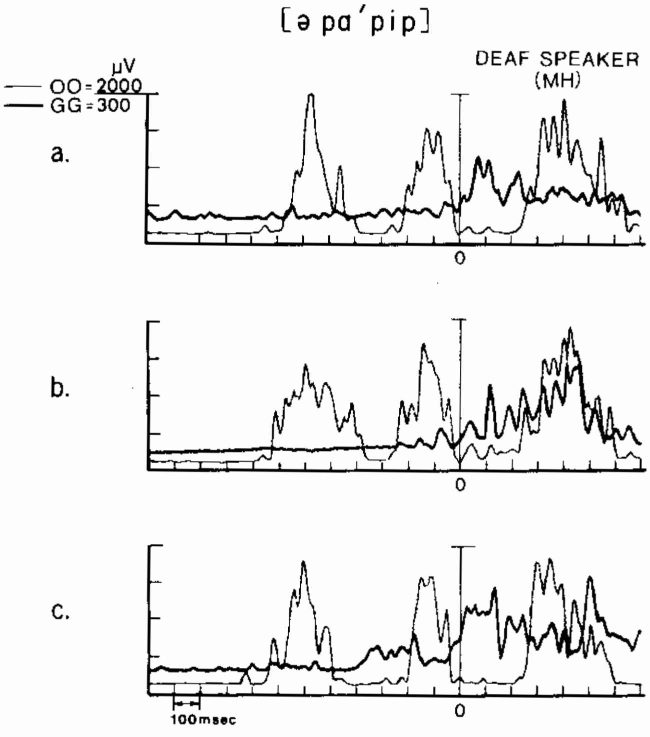


FIGURE 12. Three individual tokens of a simple nonsense utterance, showing 00 and GG outputs as a function of time, for a hearing-impaired talker. (Reproduced with permission from McGarr & Harris, 1983.)

tivity of the visible articulator. This could produce the kind of acoustic variability analyzed in Rubin's work. Similar interarticulator variability has also been described in our work with deaf speakers for larynx-upper articulators (McGarr & Löfqvist, 1982) and tongue-lip (McGarr & Gelfer, 1983) coordination.

One final result illustrates the extraordinary stability of interarticulator timing in normal adult speech production. Harris, Tuller, and Kelso, (in press); Tuller & Kelso, 1984; Tuller, Kelso, and Harris, (1982, 1983) have performed a series of experiments in which normal adult subjects produce simple nonsense syllables (again, of the form /papap/), with stress on either the first or second syllable and at two self-selected speaking rates. In a typical experiment, lip and jaw movements were monitored by fixing light-emitting diodes on these articulators. In an utterance such as /babab/, downward jaw movements can be associated with vowels, while upward lip movement can be associated with consonants. Tuller was thus able to examine the relationship of the temporal onset of the medial consonant to the duration of a vowel-to-vowel interval.

Figure 13 shows the data plots with the values of r and the slopes for a linear regression for four utterance types, /bapab/, babab/, /bawab/, and /bavab/ for a single speaker. The r values do not vary systematically with consonant. For the various measures analyzed, the Pearson product-moment correlation values range from +.84 to +.97 across the four subjects of the experiment. While the values of m show a trend towards flatter slopes, and thus, earlier consonant onsets for /v/ and /w/ as compared to /p/ and /b/, the ordering of slopes was not identical across subjects.

The substantial size of the linear correlations suggests that stability of the ratio over changes in vowel duration produced by stress and speaking rate changes is a characteristic of mature normal speech production. If we were to examine similar data for normal children, we would expect a systematic decrease in the scatter around the line of best fit with increasing articulatory maturity. For deaf speakers, we would expect even lower correlation values and we are presently analyzing data from a comparative study of deaf and normal speakers.

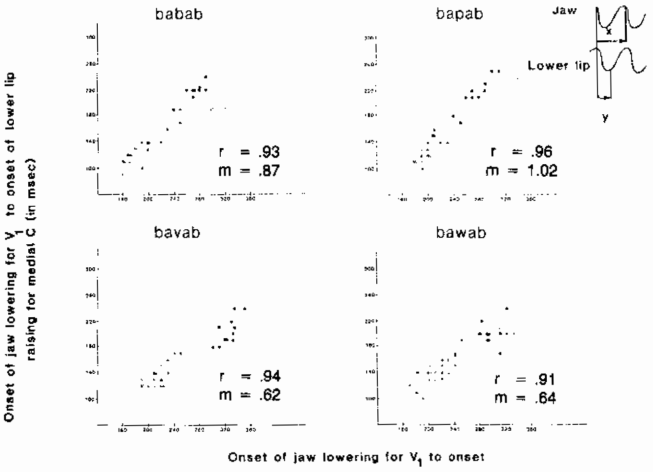


FIGURE 13. Period (jaw lowering) versus latency (lower lip raising) for nonsense disyllables differing in medial consonant for a single subject. Circles indicate utterances spoken at a conversational rate, triangles indicate a somewhat faster rate. Filled symbols have stress on the first syllable, open symbols have stress on the second syllable (Reproduced with permission from Tuller, Kelso, & Harris, in press).

Finally let us return to the beginning of this paper and point to the moral. Although "deaf speech" may have distinctive characteristics, the striking thing about the results reported here is the link between deaf speech and motorically immature speech. This relationship will in part be obscured by any description which ignores production variability as an essential characteristic of the speech production capabilities.

ACKNOWLEDGMENTS

A portion of this research was conducted in partial fulfillment of the requirements for the Ph.D. at The City University of New York by Judith Rubin-Spitz. This, and the other research cited, was supported by Grant Nos. NS-13617 and HD-01994 to Haskins Laboratories, and Grant No. NS-17764 to the Graduate School and University Center, City University of New York, as well as the Faculty Research Award Program, City University of New York.

REFERENCES

ANGELOCCI, A., KOPP, G., & HOLBROOK, A. (1964). The vowel formants of deaf and normal hearing eleven to fourteen year old boys. *Journal of Speech and Hearing Disorders, 29*, 156–170.

BUSH, M. (1981). *Vowel articulation and laryngeal control in the speech of the deaf.* Unpublished doctoral dissertation, Massachusetts Institute of Technology.

CALVERT, D. R. (1961). *Some acoustic characteristics of the speech of profoundly deaf individuals.* Unpublished doctoral dissertation, Stanford University.

EDWARDS, M. L., & SHRIBERG, L. D. (1983). *Phonology: Applications in communicative disorders.* San Diego: College-Hill Press.

EGUCHI, S., & HIRSH, I. J. (1969). Development of speech sounds in children. *Acta Otolaryngologica Supplement 257.*

EIMAS, P. D. (1982). Speech perception: A view of the initial state and perceptual mechanics. In J. Mehler, E. C. T. Walker, & M. Garrett (Eds.), *Perspectives on mental representation* (pp. 339–360). Hillsdale, NJ: Erlbaum.

ELBERT, M., DINNSEN, D. A., & WEISMER, G. (1984). Phonological theory and the misarticulating child. *ASHA Monograph 22*, 1–4.

FERGUSON, C. A., & FARWELL, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*, 419–439.

FISHER, J., KING, A., PARKER, A., & WRIGHT, R. (1983). Assessment of speech production and speech perception as a basis for therapy. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training and personnel preparation* (pp. 195–214). Baltimore, MD: University Park Press.

GULIAN, A. B., & HINDS, J. (1981). Response bias in the perception of deaf speech by experienced and naïve listeners. *British Journal of Audiology, 15*, 165–171.

HARRIS, K. S. (1983). A review of *Phonological interventions: Concepts and procedures. Applied Psycholinguistics, 4*, 294–297.

HARRIS, K. S., & MCGARR, N. S. (1980). Relationships between speech perception and speech production in normal hearing and hearing impaired subjects: Implications for training speech of the hearing impaired. In J. D. Subtelny (Ed.), *Speech assessment and speech improvement for the hearing impaired* (pp. 316–337). Washington, DC: Alexander Graham Bell Association.

HARRIS, K. S., TULLER, B., & KELSO, J. A. S. (in press). Temporal invariance in speech production. In J. Perkell (Ed.), *Invariance and variance in speech processes: Proceedings of Conference at MIT.* Hillsdale, NJ: Erlbaum.

INGRAM, D. (1976). *Phonological disability in children.* New York: Elsevier.

JAKOBSON, R. (1968). *Child language, aphasia, and phonological universals.* The Hague: Mouton.

KENT, R. D. (1976). Anatomical and neuromuscular and maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research, 19,* 421–447.

MANGAN, K. (1961). Speech improvement through articulation testing. *American Annals of the Deaf, 106,* 391–396.

MARKIDES, A. (1970). The speech of deaf and partially hearing children with special reference to factors affecting intelligibility. *British Journal of Disorders of Communication, 5,* 126–140.

MAXWELL, E. M. (1979). Competing analyses of deviant phonology. *Glossa, 13,* 181–214.

McGARR, N. S. (1978). *The differences between experienced and inexperienced listeners in understanding the speech of the deaf.* Unpublished doctoral dissertation, The City University of New York.

McGARR, N. S., & GELFER, C. (1983). Simultaneous measurements of vowels produced by a hearing-impaired speaker. *Language and Speech, 26,* 233–246.

McGARR, N. S., & HARRIS, K. S. (1983). Articulatory control in a deaf speaker. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation.* Baltimore: University Park Press.

McGARR, N. S., & LÖFQVIST, A. (1982). Obstruent production in hearing-impaired speakers: Interarticulator timing and acoustics. *Journal of the Acoustical Society of America, 72,* 34–42.

MENN, L. (1980). Phonological theory and child phonology. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson, (Eds.), *Child phonology production* Vol. 1 (pp. 23–41). New York: Academic Press.

MENYUK, P., & MENN, L. (1979). Early strategies for the perception and production of words and sounds. In P. Fletcher & M. Garman (Eds.), *Language acquisition* (pp. 49–70). Cambridge: Cambridge University Press.

MONSEN, R. B. (1976). Normal and reduced-phonological space: The production of English vowels by deaf adolescents. *Journal of Phonetics, 4,* 189–198.

MONSEN, R. B. (1983). General effects of deafness on phonation and articulation. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training and personnel preparation.* Baltimore, MD: University Park Press.

MONSEN, R., & ENGEBRETSON, M. (1983). The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear production. *Journal of Speech and Hearing Research, 26,* 89–97.

NICKERSON, R. (1973). *Computerized speech training aids for the deaf.* Report No. 2633. Bolt, Beranek & Newman.

OCHIAI, K., & FUJIMURA, O. (1971). Vowel identification and phonetic contexts. *Reports of the University of Electro-Communication 22,* 103–111.

PISONI, D., CARRELL, T. D., & SIMNICK, S. S. (1979). Does a listener need to recover the dynamic vocal tract gestures of a talker to recognize his vowels? *Journal of the Acoustical Society of America, 22,* 807–820.

RUBIN, J. (1983). *Static and dynamic information in vowels produced by the hearing impaired.* Bloomington, IN: IULC.

SMITH, B. L. (1978). Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics, 6,* 37–67.

SMITH, N. V. (1973). *The acquisition of phonology: A case study.* London: Cambridge University Press.

STEVENS, K. N., NICKERSON, R. S., & ROLLINS, A. M. (1983). Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training and personnel preparation.* Baltimore, MD: University Park Press.

STUDDERT-KENNEDY, M. (in press). Sources of variability in early speech development. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability of speech processes.* Hillsdale, NJ: Erlbaum.

THOMAS, W. (1963). Intelligibility of the speech of deaf children. *Proceedings of the International Congress on the Education of the Deaf.* Document No. 196.

TULLER, B., KELSO, J. A. S., & HARRIS, K. (1982). Interarticulator phasing as an index of temporal regularity in speech. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 460–472.

TULLER, B., KELSO, J. A. S., & HARRIS, K. S. (1983). Further evidence for the role of relative timing in speech. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 829–833.

TULLER, B., & KELSO, J. A. S. (1984). The timing of articulatory gestures: Evidence for relational invariance. *Journal of the Acoustical Society of America, 76,* 1030–1036.

VERBRUGGE, R. R., STRANGE, W., SHANKWEILER, D. P., & EDMAN, T. R. (1976). What information enables a listener to map a talker's vowel space? *Journal of the Acoustical Society of America, 60,* 198–212.

Chapter 10

RESPIRATORY FUNCTION IN SPEECH PRODUCTION BY NORMALLY-HEARING AND HEARING-IMPAIRED TALKERS: A REVIEW

JUDITH L. LAUTER

Central Institute for the Deaf, St. Louis, MO

Current models of speech production usually posit a fairly abstract semantic stage at one extreme, and a fairly specific stage of muscle-fiber control at the other, all referred to the sounds produced. In between, one or more intermediate stages are usually described, with various specifications. It might be suggested that one of these intermediate stages involves control of respiration, specifically, "breathing-for-speech." As we will see, control of respiration can be seen as involved intimately in a range of speech behaviors, from the articulation of phonemes to the rhythmic structuring of sentences. Also, problems that the severely hearing impaired have with speech production may be related to knowing how to breathe while talking.

DYNAMICS OF SPEECH BREATHING

It has been known for some time (e.g., Stetson, 1951) that speech-breathing is somewhat different from quiet or "tidal" breathing. Borden and Harris (1980) note that more air is inspired during breathing for speech and the proportions of the cycle devoted to expiration and inspiration are very different (Figure 1). Hixon and colleagues (1973, 1976, 1982) have demonstrated a number of distinctions between quiet and speech-breathing.

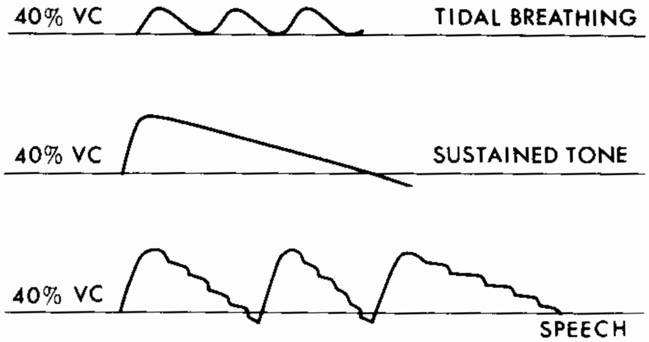


FIGURE 1. Use of lung capacity and rate of breathing compared for three different types of respiration: tidal (quiet) breathing, sustained vocalization, and normal speech. Lung capacity is indicated on the ordinate, and is shown in relation to 40% vital capacity, the lung volume at the end of a quiet expiration. Breathing rate is shown along the horizontal (time) axis, and is indicated by the relative slopes of inspiration versus expiration. (Reprinted with permission from Borden & Harris, 1980).

These differences can be quite dramatic. Von Euler (1982) reports that while the muscles of the diaphragm continue to be active through about one half the expiration phase of quiet breathing, in speech-breathing, the muscles relax completely at the onset of expiration. Also, the metabolic reaction that occurs when subjects consciously hyperventilate without talking does not follow the hyperventilation that accompanies speech. Von Euler goes on to suggest that control of the two kinds of breathing may be partially separated in the CNS, quiet breathing depending on structures restricted to the brainstem and spinal cord, while voluntary breathing involves control centers in the cortex and basal ganglia as well (see also Abbs & Cole, 1982).

These and other data suggest that breathing-for-speech involves a set of motor skills that children must learn if they are to produce speech that sounds normal. Stetson (1951) presented his studies on speech production as an analysis of a set of "skilled movements." How might we think of the details of this skill as it relates to speech production? In a chapter published in 1973, Ron Netsell suggested that a useful description of the set of body structures used in speech production, the "speech apparatus," was as a system designed for generating and valving an airstream. The acts of control and coordination usually described with reference to the sounds thus produced, in this view are defined according to the effects on air flow and air pressure through the system.

Netsell went on to describe changes in airflow and pressure correlated with a range of linguistic events, from segmental to suprasegmental (prosody). He divided the "speech apparatus" into nine components (see Figure 2), and noted that the control of segmental aspects of speech in terms of this system required steady pressure maintained by the lower components, controlled modulation of the laryngeal "valve," modulation movements within the upper vocal tract, and extremely fine temporal-spatial coordination of all components—a coordination which must be able to comprehend within the same time frame the action of abdominal muscles as well as movements of the tip of the tongue.

Netsell noted how prosodic aspects of speech could be described in terms of the same components, with "valving" muscles and generated air pressure working together to achieve intonation—valve timing acting to achieve rhythm

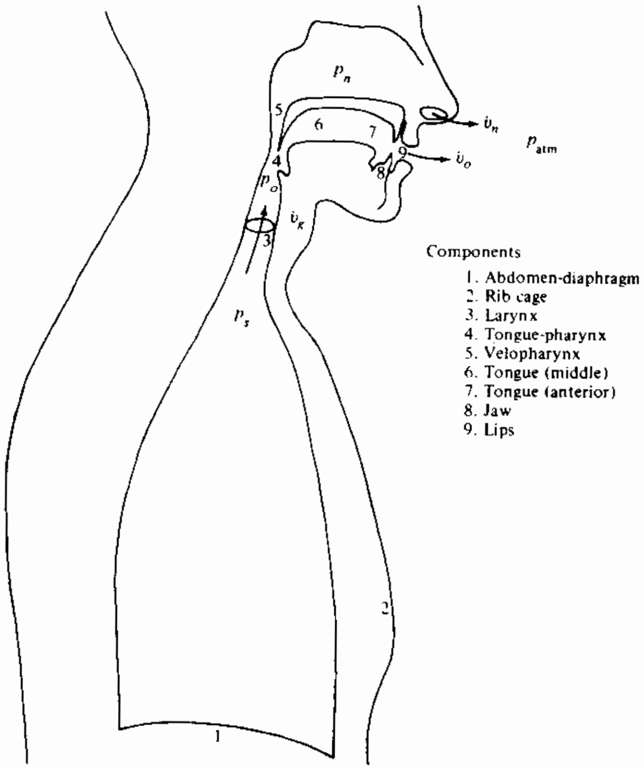


FIGURE 2. The nine components of the speech apparatus as described by Netsell (1973). The symbols /*v*/ and /*p*/ indicate various points where air *v*olume and air *p*ressure may be measured and compared. (Reprinted with permission from Netsell, 1973).

and pitch control at both the segmental and phrase level, and subglottal pressure, extent of movements within the vocal cavity, and contact force of the different valves changing with the amount of effort of an utterance.

Speech-Breathing and Deaf Speech

It is interesting to compare the linguistic effects of the actions of such a system with the characteristics of typical "deaf speech" such as described by Nickerson (1975) and Osberger and McGarr (1982). Nickerson divides these characteristics into classes that can be easily related to the segmental/suprasegmental distinction used by Netsell. Nickerson notes that deaf speakers often have poor articulation, including substandard velar control, a restricted range of F2 variation, problems with voiced/voiceless distinctions and with "continuous phonation." Many of these characteristic difficulties in deaf speech could be described as inadequate control (perhaps in terms of poorly learned control constraints) of Netsell's "valves"—the velum, the tongue, the vocal folds.

The same could be said for Nickerson's list of the characteristics of deaf "voice quality"—nasality (velar valve), breathiness (laryngeal valve), inappropriate loudness (perhaps compensation using changes in intensity controlled at the larynx instead of changes in fundamental frequency managed there), and durational distortions. Even more suggestive are Nickerson's deaf-speech characteristics having to do with suprasegmental aspects of speech—they read as though taken from Netsell's list of the speech aspects that depend intimately on temporal coordination of the nine speech-apparatus components. Deaf talkers' timing and rhythm are often abnormal in that these speakers may not provide clear duration distinctions between stressed and unstressed syllables, and segmental durations can be inaccurate. Also, pitch and intonation may be affected, in that base-line fundamental frequency is often too high, there is little variation in fundamental frequency, and modulations of intensity seem to be substituted for variations in pitch.

Respiration and Speech Planning

It is possible that Netsell's approach to speech production may bear both on questions regarding the planning and production of speech, as well as the problems of individuals with handicaps such as motor disorders or hearing impairment. First, emphasis on this "intermediate," perhaps underlying skill of speech-breathing—defined in Netsell's broad terms of *airstream generation* and *modulation*—may serve as a guide for studying the physiology of speech acts. Interrelations between neural control centers and patterns of movement control may be suggested by this approach that would not emerge from thinking only about the sounds produced. For example, poor use of muscles of the torso for controlling subglottal pressure may have direct effects (perhaps via open-loop feed-forward connections) on control of the larynx, which could result in abnormalities in voice pitch. Von Euler (1982) has pointed out that the cerebellum may use the input it receives from both the larynx and the lungs to coordinate laryngeal and lower respiratory motor activities in phonation.

Certainly current concepts of motor control such as "heterarchical organization" (cf. Turvey, 1982) are compatible with this view of speech production. For example, details of motor action occurring against a background of general system "tuning" might be exemplified by the pulsed actions of the intercostal muscles timed against the background of other muscles acting to maintain subglottal pressure at a generally constant level throughout an utterance. MacNeilage's (Chapter 4) concept of "frame/content" organization could be illustrated by individual gestures of different valves programmed to match details of an utterance stress contour, and Bernstein's (1967) idea of "interactive coordinative structures" could be used to describe the interaction of the larynx and sublaryngeal structures to maintain subglottal pressure in the face of laryngeal actions such as opening and closing for segmental differentiation.

With developments in technology, it has become possible to perform noninvasive studies of the speech respiratory activity of normal and hearing-impaired individuals. Woldring (1968) used pneumographs (re: thorax and abdomen) to compare breathing patterns in one normal and two deaf children, from 10 to 12 years of age. He reported that during phonation the deaf subjects showed an absence of controlled expiration, with either insufficient ventilation or hyperventilation. He suggested that their poor control was due to the lack of auditory feedback; Woldring noted that "deaf glassblowers, in whom the feedback process is visual and not disturbed," show good control of respiration skills needed in glassblowing. Forner and Hixon (1977), using the kinematic procedure developed by Hixon, et al. (1973), reported a study of 10 young male deaf students. Two pairs of magnetometer coils were used to measure movements of the chest and abdomen during a variety of respiratory maneuvers, including quiet breathing, and breathing during a series of speech tasks. The authors concluded that although the deaf speakers showed quiet breathing patterns that were within normal limits, their speech breathing was generally deviant. Departure from normal behaviors included: fewer syllables per breath than normals, less air inspired with each breath than normals, higher volume of air per syllable than normals, and inspirations taken at linguistically irrelevant points. Whitehead (1983) used similar measurement techniques to study 15 young deaf males, whose speech was rated as semiintelligible or unintelligible. He reported results similar to those seen by Forner and Hixon (1977), and suggested that speech intelligibility might be affected by a speaker's respiratory skill. Specifically, such practices as initiating speech at low lung volumes, and continuing speech beyond the lower limit of tidal breathing, could contribute directly to listeners' difficulty with comprehension.

Aerodynamic Feedback for Deaf Talkers

Second, as Woldring (1968) suggested, it is possible that individuals who cannot hear the effect of actions of the "speech apparatus" may benefit from feedback directly related to its aerodynamics. Forner and Hixon (1977) included in their report a final study where they showed one of the deaf talkers the display that formerly only the experimenters had seen; and taught the subject how movements of his torso could affect the tracing. After a few minutes of working with the display, the hearing-impaired speaker learned to: (a) produce a speech-breathing pattern more like that of a normal speaker, and (b) as a side effect, without direct attention by experimenters or subject, lower his abnormally high "deaf voice pitch" to a normal level.

Certainly the importance of "breathing exercises" is cited in the oldest treatises on oral education of the deaf. However, examination of these descriptions reveals a lack of understanding about the intimate relations between gestures within the respiratory tract and segmental *and* suprasegmental characteristics of speech. It is possible that the general failure in teaching the deaf to produce normal speech is based in part on the failure to teach them the motor skills involved in breathing for speech. Forner and Hixon (1977) reported that some of their hearing-impaired subjects said that they were only "taught to make speech sounds, never to breathe in a different way for speech than for quiet breathing." As we have suggested, the range of relevant motor skills involve a variety of details, from knowing that more air needs to be inspired for speech than for tidal breathing to being aware of the effect of leaving the velopharyngeal valve open.

CONCLUSION

In the past, conclusions about planning for speech production have been drawn from observations of speech dysfunction, as in aphasia, dysarthria, and spontaneous speech errors (for the latter cf. Fromkin, 1973; Shattuck-Hufnagel, 1985). It is possible that observations of the aerodynamics of deaf speech, both before and after relevant instruction, may provide new evidence for the stages of planning, centers of control, and details of coordination that are involved in creating the disturbances in the air that we hear as speech.

REFERENCES

Abbs, J. H., & Cole, K. J. (1982). Consideration of bulbar and suprabulbar afferent influences upon speech motor coordination and programming. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 75–93). Oxford: Pergamon Press.

Bernstein, N. (1967). *The coordination and regulation of movements.* Oxford: Pergamon Press.

Borden, G. J., & Harris, K. S. (1980). *Speech science primer.* Baltimore: Williams & Wilkins.

Forner, L. L., & Hixon, T. J. (1977). Respiratory kinematics in profoundly hearing-impaired speakers. *Journal of Speech and Hearing Research, 20,* 373–408.

Fromkin, V. A. (1973). *Speech errors as linguistic evidence.* The Hague: Mouton.

von Euler, C. (1982). Some aspects of speech breathing physiology. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 95–103). Oxford: Pergamon Press.

Hixon, T. J. (1982). Speech breathing kinematics and mechanism inferences therefrom. In S. Grillner, B. Lindblom, J. Lubker, & A. Persson (Eds.), *Speech motor control* (pp. 75–93). Oxford: Pergamon Press.

Hixon, T. J., Goldman, M., & Mead, J. (1973). Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *Journal of Speech and Hearing Research, 16,* 78–115.

Hixon, T. J., Mead, J., & Goldman, M. D. (1976). Dynamics of the chest wall during speech production: Function of the thorax, rib cage, diaphragm, and abdomen. *Journal of Speech and Hearing Research, 19,* 297–356.

Netsell, R. (1973). Speech physiology. In F. D. Minifie, T. J. Hixon, & F. Williams (Eds.), *Normal aspects of speech, hearing, and language* (pp. 211–234). Englewood Cliffs, NJ: Prentice-Hall.

Nickerson, R. S. (1975). Characteristics of the speech of deaf persons. *Volta Review, 77,* 342–362.

Osberger, M. J., & McGarr, N. S. (1982). Speech production characteristics of the hearing impaired. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice, Vol. 8* (pp. 221–283). New York: Academic Press.

Stetson, R. H. (1951). *Motor phonetics* (2nd ed.). Amsterdam: North-Holland Publishing.

Turvey, M. T. (1982). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing; toward an ecological psychology,* (pp. 211–265). Hillsdale, NJ: Lawrence Erlbaum Associates.

Whitehead, R. L. (1983). Some respiratory and aerodynamic patterns in the speech of the hearing impaired. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired.* (pp. 97–116). Baltimore: University Park Press.

Woldring, S. (1968). Breathing patterns during speech in deaf children. *Annals of the New York Academy of Sciences, 155,* 206–207.

Chapter 11

MEMORY ISSUES IN SPEECH PLANNING AND PRODUCTION

CRAIG A. BOWE

Central Institute for the Deaf, St. Louis, MO

What is the role of memory in the planning and production of speech by normally hearing and hearing-impaired people? Although this is certainly a question of interest to those concerned with normal and disordered speech planning and production, no definitive answer is yet available. In this paper, we will review some of the relevant research. The role of memory will be considered in two stages of speech behavior—planning and production. The planning stage of speech will be described as involving processes of interest to students of "verbal" learning and memory, while the production stage of speech will be considered as involving processes studied by those interested in "motor" learning and memory. Of course, the planning and production stages must overlap if speech is to occur, and during our separate consideration of these stages, we will offer some comments on how these stages may overlap.

MEMORY AND SPEECH PLANNING

Are there distinct stages of memory involved in the planning of an utterance? Do hearing-impaired speakers remember phrases, words, or phonemes differently than do normally hearing speakers? If speech is planned at different levels or stages of a memory system, is the planning done sequentially or in parallel? Is there a special kind of memory for speech? How does speech get so well-organized in memory? These are just a few of the possible questions about speech planning that can be raised within the context of verbal learning and memory research.

Models of Memory and Speech Planning

The study of verbal learning and memory was once dominated by a theoretical viewpoint known as stimulus-response associationism imported from the study of the learning and behavior of nonverbal animals. In essence, this viewpoint held that animals (including humans) learn and remember things in a more or less passive manner as a result of automatic connections that are formed between spatially- and temporally-contiguous stimuli and responses. About 25–30 years ago, however, a new approach to verbal learning and memory came into vogue that was based on an analogy between humans and computers. Simply put, this new approach maintained that humans are active, information processors that, like computers, accept only certain kinds of input (the stimulus), code the input in prespecified ways, perform various transformations on the coded input according to built-in structures and rules, change various internal conditions (or "programs") in response to the transformed input, and finally select and produce an appropriate output (the response).

Two general information-processing views of memory can be identified at present—the multiple-component view (Atkinson & Shiffrin, 1968, 1971), and the unitary view (Craik & Lockart 1972; Melton 1963). We will consider only the first. According to the Atkinson-Shiffrin model (see Figure 1), there are four separate but related components in the memory system: a sensory register, a short-term memory (STM), a long-term memory (LTM), and a response generator. The sensory register transforms the physical stimulus into an internal representation of the stimulus that lasts only a very brief time. STM can be roughly equated in common terms to one's "consciousness." Information passes from the sensory register into STM where various "control processes," like coding and rehearsal, operate on the information. Then, via a control process called retrieval (or search), STM determines which information will be transferred between STM and LTM and how it will be related to the other information already in the system. STM is "capacity-limited" (see Shiffrin, 1976, for a detailed discussion of capacity limits) in that only a small amount or "chunks" (Miller, 1956) of information can occupy STM at any given moment and the information in STM is constantly being replaced at a fairly rapid rate (e.g., a maximum-duration, unrehearsed existence in STM of about 18 s for the trigrams that were used in Peterson & Peterson's, 1959, classic experiment). Information that reaches LTM is al-

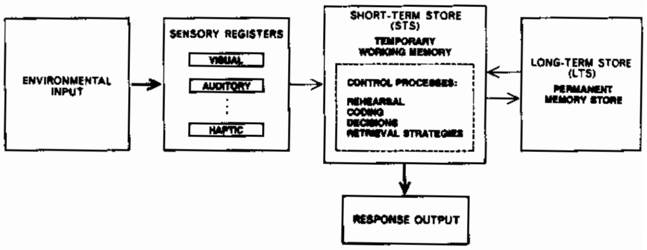


FIGURE 1. Schematic of the multiple-component information-processing model of memory. (Reprinted with permission from Atkinson and Shiffrin, 1971, The control of short-term memory. *Scientific American, 225*, 82–90.)

ways available for future reference. Therefore, whatever has been learned and is then forgotten by an individual is not the result of a decay process. Most forgetting occurs when the retrieval process that is initiated in STM either (a) "finds" incorrect information obscuring the correct information available in LTM (i.e., interference occurs) or (b) simply fails to find the correct information within a reasonable amount of time and so terminates its search of LTM. Decisions made in STM as a result of its interaction with LTM are finally passed to the response generator which then issues motor commands to the effectors for the appropriate action(s).

Consider, now, one way in which the above sketch of the Atkinson-Shiffrin model might deal with the processing of a single English phoneme, /a/, as heard and then repeated by a normally-hearing speaker of English. First, the acoustic stimulus is received in most, if not all, of its detail by the auditory processor in the sensory register. This "echoic" image of the physical stimulus then begins to rapidly decay within about 100-250 msec (see Efron, 1970, or Massaro, 1970). However, before the image disappears altogether, certain important characteristics of the /a/ have passed into STM. By the time the /a/ reaches STM, a great deal of the information in the original acoustic stimulus has been lost. In STM all other information except that necessary for correct identification of the phoneme can also disappear from the system. Presumably, identification of the phoneme occurs in STM as a result of coding, rehearsal, and retrieval control processes. The coding process (acoustic or articulatory) transforms the remaining echoic information from the sensory store into a form which can be maintained in STM via the rehearsal process. While maintenance rehearsal is ongoing, the retrieval process brings information from LTM into STM that can be used to identify the rehearsed information as an /a/ (note the parallel operation of control processes here). Once identification has occurred, a decision process that has a criterion set by prior instructions like, "Repeat the sounds you hear," determines the form of information that STM will pass to the response generator. Finally, based on the information that is passed to it from STM, the response generator issues the motor commands to the articulators that will produce the spoken /a/.

Obviously, this is a highly-idealized description of how a multiple-component memory system might process speech at the level of individual phonemes. Whether the system actually processes individual phonemes in the fashion proposed above, whether the phonemic level is only one of many levels processed serially or in parallel, or whether the phonemic level of speech simply reflects the processing of a more basic form of data within the system (e.g., F1 by F2 by F3 plots of entire words) can only be answered by future research.

LTM and Speech Planning

The structure and operation of LTM in its interactions with STM is still very much a mystery. It is generally agreed that LTM is, at least in part, structured semantically (e.g., Bartlett, 1932; Bransford, Barclay, & Franks, 1972; Clark & Clark, 1968; Katz & Fodor, 1963). Indeed, a number of memory models have been proposed for the semantic structure of LTM (e.g., Anderson & Bower, 1973; Quillian, 1968; Rumelhart, Lindsay, & Norman, 1972; Winograd, 1972). The "meaning" of information in the memory system is derived from the organization of LTM. A great amount of research has been devoted to exploring the effects of semantic organization imposed on to-be-remembered information by either the experimenter or the subject (e.g., Mandler, 1967; Tulving, 1962; Tulving & Donaldson, 1972). However, given the severe capacity limitations of STM, it must be true that important levels of speech other than the semantic, such as the syntactic level or the phonemic level, also find their representations within the structure of LTM.

One piece of laboratory evidence that speakers store information from levels of speech besides the semantic in their LTM comes from the "tip-of-the-tongue" phenomenon (Brown & McNeill, 1966). Brown and McNeill showed that, when subjects were instructed to respond with the word corresponding to a definition that was provided by the experimenter, the subjects often were able to correctly report nonsemantic characteristics of the word, like certain of its letters, the number of syllables, etc., even though they could not immediately reproduce the word. Subsequently, many of these words on the tip-of-the-tongue were recognized by the subjects when read by the experimenter. This phenomenon, by the way, not only supports the idea that LTM is not exclusively semantic, but also supports the notion that LTM is a permanent repository for all of one's learned information, which may then be "forgotten" as a result of the failure of retrieval processes initiated by STM.

Somewhat stronger evidence that acoustic, in addition to semantic, information is directly stored in LTM lies in our ability to remember the very complex acoustic patterns of music. Although one might argue that a semantic code is responsible for the maintenance of the information in LTM that yields the tip-of-the-tongue phenomenon, it seems very unlikely that the recognition and recall of musical pieces by nonmusicians, at least, is based on a semantic rather than an acoustic code. And, if music is stored acoustically in LTM, it seems economical for evolution also to have provided acoustic storage space in LTM for some aspects of one of our most important activities—speech. Thus, it is possible that basic elements of speech, such as phonemes or allophones, have an acoustic representation in LTM. We therefore concur with Baddeley's (1976, p. 118) statement that, "it is surely necessary to have a long-term acoustic memory in order to learn to speak and understand in the first place."

STM and Speech Planning

Although a distinguishing aspect of STM is its capacity limitations, none of the STM research dealing with capacity limitations has employed elementary units of speech in isolation, such as phonemes or subphonemic features, to determine, for example, whether only 7 plus or minus 2 of such units can simultaneously and separately occupy STM.[1] Nonetheless, assuming a separate coding for each unit and a way to present isolated units to a subject, there is no reason to believe that

---

[1] Of course, "chunking" or other coding efforts usually appear to override the separate existence of units like phonemes within STM.

the capacity limits for such units would be qualitatively different from those placed on the information typically employed in STM experiments (e.g., words or "nonsense syllables").

Hintzman (1967) and Wickelgren (1965) were among the first to suggest that the coding processes of STM were basically auditory or articulatory in nature. Even visually presented linguistic information was thought to be coded acoustically for subsequent use by rehearsal and by other control processes of STM, because in various STM procedures reliable confusion errors could be demonstrated that were based on the acoustic/articulatory properties of the visually-presented items. Other research (e.g., Baddeley, 1966), however, dispelled the notion that coding in STM is exclusively acoustic or articulatory. Nevertheless, one of the most important control processes of STM, "maintenance" or "rote" rehearsal, appears to depend heavily on such coding for its ability to keep vital information active in STM, despite the continuous influx of new information which displaces nonrehearsed information in STM.

One way in which STM might retrieve information from LTM is via a kind of sample-and-recognize model (e.g., Shiffrin, 1970). This model treats the processing of speech as if it were serial or sequential. However, serial processing seems much too slow to handle the high rate of information that we apparently can process during normal speech. In fact, it is probable that most of speech perception and production occurs automatically, without the need for any conscious or controlled efforts on the part of the speaker/listener. It is in this respect that the study of "controlled" versus "automatic" processing (Shiffrin, 1976; Shiffrin & Schneider, 1977) becomes important for an understanding of speech planning.

The crucial point about controlled and automatic processing for our present purposes is that given a great deal of fairly repetitive experience with a particular kind of information, the processing of that information becomes automatic in the sense that it is perceptually very salient and allows very accurate and rapid responding. For example, subjects in a visual-search task who are asked to detect various numbers of letter "targets" among "distractors" in rapidly-presented displays can do so more easily (rapidly) if the distractors are digits than if they are other letters. Finding letters among digits is automatic processing and finding letters among other letters is controlled processing because of the extensive preexperimental experience that the subjects (college sophomores) had in separately categorizing letters and digits.[2] Something like the same sort of perceptual learning that yields automatic processing probably occurs with respect to various levels of speech.

Often, the final information-processing task of STM is the delivery of a decision to the response generator for translation into action. It is at this point in speech that our conceptual boundary between planning and production becomes blurred. Unfortunately, research in verbal learning and memory, specifically that related to multiple-component approaches like the Atkinson-Shiffrin model, has been scarce (nonexistent?) concerning the nature of the transmission from STM to the response generator and thence to the effectors. Thus, we must wait for the appropriate theoretical and empirical efforts to be made before the overlap between the planning and the production of speech in the memory system can be fruitfully addressed in terms of a multiple-component view.

"Acoustic" Memory and Speech Planning

We have already mentioned a number of lines of evidence showing that components of a memory system, such as those of the Atkinson-Shiffrin model, are especially adapted for the processing of auditory information (e.g., the occurrence of echoic images and the tip-of-the-tongue phenomenon). STM experiments also provide evidence for an "acoustic memory" (see Darwin & Baddeley, 1974, for an extensive discussion of an acoustic memory resembling tape-recordings that decay with time).

Conrad and Hull (1968) reported a serial-position experiment using 6-digit sequences, where performance improved if the subject was able to read the sequences aloud instead of silently. Thus auditorily-presented linguistic information seems more suited to processing by STM than visually-presented information (even when such visually-presented information can be easily coded in an acoustic or articulatory fashion). More dramatically, perhaps, Crowder and Morton (1969) and others have demonstrated a "suffix effect," whereby immediate recall of the last digit of a 7-digit sequence in a serial-position experiment is severely disrupted when followed by speech-like stimuli but is not affected when followed by visual or nonspeech-like stimuli. It is worth noting that the semantic content of the suffix has no effect on the occurrence of this phenomenon. Darwin and Baddeley (1974) have concluded that the suffix effect is greatly influenced by the "acoustic similarity" among the to-be-recalled items. However, as Darwin and Baddeley noted, it may be difficult to provide an independent and objective scaling of the acoustic "similarity" of items. Therefore, it may not be possible to predict the occurrence of such an acoustic memory phenomenon.

Memory in the Planning of Hearing-Impaired Speech

Ling's (1978) comments on his experience with teaching hearing-impaired children to speak illustrate the importance of verbal STM for speech and vice versa:

> The value of speech to a hearing-impaired child is not limited to his production of it. A child who is able to speak is better able to understand the speech of others. Not only does his own speech provide a framework against which he can match incoming patterns, it provides him with a means to rehearse verbal material in short term memory. If I said a new word to you and asked you to remember it without writing it down, you would say it to yourself subvocally . . . ***If the hearing-impaired child is not taught speech skills, he is being deprived of an important tool for developing effective strategies of storing, rehearsing, and recalling verbal information.*** (pp. 115–116)

---

[2] Subjects can be taught to process automatically in tasks requiring detection of letters among other letters or digits among other digits if a "consistent mapping" (Shiffrin & Schneider, 1977) procedure is employed such that the items presented as targets and distractors are never mixed for a given subject. That is, for a given subject, a certain letter or a certain digit is always a target and never a distractor or always a distractor and never a target.

Laboratory research within the verbal learning and memory tradition has been conducted using hearing-impaired speakers (see Hunt, 1982, for a review of these studies). Selected results from this research are as follows:

1. Normally hearing subjects performed better than hearing-impaired subjects on a digit-span task with written item presentations and written responses by the subjects. But a memory-span task based on visual "nonsense" forms showed no differences between normally hearing and hearing-impaired subjects (Olsson & Furth, 1966).
2. A memory-span task based on a list of words spoken aloud to normally hearing subjects (via a tape-recorder) and signed to hearing-impaired subjects (via a video recording in ASL) revealed a span of 5.9 items for the hearing subjects and 4.9 items for the hearing-impaired subjects, a slight but statistically-significant difference between the two groups of subjects (Bellugi, Klima, & Siple, 1974).
3. An STM task in which consonant lists or word lists (some consonants and words were acoustically similar—B/V or past/passed—and others were visually similar—F/P or race/care) were presented to hearing-impaired subjects for immediate recall showed that some subjects relied on articulatory coding because they made more errors on acoustically similar consonant or word lists and some subjects relied on other (idiosyncratic) coding strategies because the systematic error patterns for each of these subjects were not correlated with the acoustic/visual division in the similarity of the presented items (Conrad, 1970; Conrad & Rush, 1965).
4. A serial-probe task in which the letter probe required subjects to report the position of the same letter within a previously presented, 6- or 7-letter list showed the performance of hearing-impaired subjects to be inferior to that of normally hearing subjects. However, upon subsequent exposure to instruction in the use of rehearsal strategies that involved labelling and clustering of the letters via finger spelling, the hearing-impaired subjects showed significant improvement in their serial-probe performances (Belmont, Karchmer, & Pilkonis, 1976).

With respect to STM and its control processes, these results suggest that if the information to be processed is primarily auditory/verbal, then hearing-impaired subjects will usually show greater capacity limitations than will normally hearing subjects. However, if the information to be processed is in some mode other than auditory/verbal or if subjects receive explicit instruction in strategies of information-processing, then hearing-impaired subjects can greatly improve their performances on STM tasks, sometimes to levels near those of normally-hearing subjects.

MEMORY AND SPEECH PRODUCTION

Is speech a "continuous" motor skill that is not subject to forgetting or is it a "discrete" motor skill, aspects of which can be forgotten? What is the role of feedback, auditory, tactile, proprioceptive, and kinesthetic, in the movements that accompany speech? How is the "invariance problem" to be solved? Do hearing-impaired speakers exhibit deficits in their motor, as well as their sensory, abilities that could be remedied in the absence of audition? What is the nature of the motor commands to the articulators? The study of motor learning and memory can provide hints for the answers to these and other questions about the production of speech.

Theories of Motor Control

Lashley's Motor-Program Theory. There is an obvious similarity between the origins of modern theories of motor learning and memory and those of verbal learning and memory. Both began as reactions against the same prevailing theoretical account of their respective phenomena—stimulus-response (S-R) associationism. The reaction of the motor theorists, though, occurred much earlier than did that of the verbal theorists. Lashley (1917, 1951) first formulated what is now termed the *motor program* approach to the understanding of motor control. Drawing mainly on his observations of deafferentated human patients and animal subjects, Lashley believed that many actions were controlled by central processes (i.e., "open-loop" control) as opposed to control by processes that employ peripheral feedback in directing action (i.e., "closed-loop" control). His belief has been echoed and amplified by current proponents of the motor program idea (e.g., Keele, 1968; Keele & Summers, 1976; Pew, 1974).

Adams' Closed-Loop Theory. In opposition to past and present motor-program notions, Adams (1971, 1976) provided a kind of modern defense of peripheralistic accounts of motor control by means of his "closed-loop theory of motor learning." Adams (1976) argued that motor-program theories (representing open-loop systems) are:

1. Not as strongly supported by deafferentation experiments (e.g., those conducted by Lashley and by Taub and his colleagues; Cohn, Jakniunas, & Taub, 1972; Knapp, Taub, & Berman, 1963; Taub, Bacon, & Berman, 1965; Taub & Berman, 1963) as had been suggested by motor-program proponents;
2. not supported in the case of learned human activities by the evidence for feedback-free behavior in insects (e.g., Wilson, 1964, 1965), and
3. not supported by the supposed sluggishness of proprioceptive feedback with respect to rapid-movement control (e.g., Lashley's, 1951, example of the rapid finger movements of a pianist that are presumably too swift for neural feedback control). Because of the rapid movement of the articulators during speech production, this latter belief of motor-program theorists that proprioception is not involved in controlling rapid movement holds particular importance for the study of speech. Adams cited Sussman (1972) for proof that the tongue, at least, can be controlled during the movements of speech by proprioceptive feedback having loop times of about 10 msec from the tongue's muscle spindles to a more central location (via hypoglossal and lingual nerves), and then back to the muscles of the tongue. Such speed, contends Adams, is sufficient for the closed-loop control of speech.

Adams' closed-loop theory of motor control is modelled after servo theory in engineering but is based on data from the laboratory study of motor skills. He outlines the essentials for a closed-loop system as one having feedback, error detection, and error correction as the key elements. There is a reference that specifies the desired value for the system, and the output of the system is fed back and compared to the reference for error detection. Once detected, errors are corrected. A common example of a closed-loop system is the automatic home furnace. The thermostat setting is the desired value, and the heat output of the furnace is fed back and compared against this reference. If there is a discrepancy the furnace turns on or off until the error is zero.

A closed-loop system of motor control achieves self-regulation by compensating for deviations from the reference (Adams, 1971). Response feedback in Adams' theory is of two basic sorts, "knowledge of results" and proprioception. Knowledge of results is the more or less external information (visual, auditory, tactile, etc.) one receives about the consequences of responding.[3] Error detection is accomplished via a comparison between the "reference mechanism" in the system and the response feedback. Mismatches denote error. When such mismatches are detected, adjustments in responding are affected to correct the error, presumably by issuing new neural commands to the appropriate effectors.

There are two aspects of Adams' theory that require memory. First, the reference mechanism in his theory is called the "perceptual trace." It is the stored representation of previous response feedback. A second aspect requiring memory is the agent for the initiation and selection of movements, which he called the "memory trace."

Schmidt's Motor-Schema Theory. As an alternative to motor-program theories and Adams' closed-loop theory, Schmidt (1975, 1976) proposed a "schema" theory of motor control which incorporates elements from both of its predecessors. In addition to his use of previous motor-program concepts, Schmidt borrowed his "schema" concept from perception research (e.g., Posner & Keele, 1968) and memory research (e.g., Bartlett, 1932).

Schmidt (1976) advanced two basic arguments that supposedly create difficulties for both motor-program and closed-loop conceptions of motor control: (a) the "storage problem" and (b) the "novelty problem." The storage problem has often been used to discredit motor-program theories. The argument is that the number of possible movements which an organism can perform is simply too large to allow storage of the appropriate program for each within the CNS. Thus, for example, MacNeilage and MacNeilage (1973) have estimated that English speakers can produce 100,000 different speech sounds, each of which requires a distinct movement that is presumably caused by its own motor program. Unfortunately, the storage-problem argument rests on the implication that the nervous system is insufficiently complex and/or detailed to store vast numbers of motor programs, and this implication is neither supported nor refuted by physiological evidence.[4] Nevertheless, Schmidt claimed that Adams' closed-loop theory has the same fatal storage problem as do motor-program theories, because memory must store, "as many references of correctness with which response-produced feedback is compared as there are movements" (p. 43). The novelty problem is that in the execution of most motor skills (e.g., driving a car or engaging in athletic events), the same response is almost never repeated exactly, even under nearly identical stimulus conditions. The novelty problem afflicts motor-program and closed-loop theories equally, because neither approach to motor control has a mechanism capable of producing novel movements, regardless of whether or not external circumstances change.

Schmidt's solution to the storage problem and the novelty problem lies in his schema theory of motor control. The motor schema is a highly generalized and abstract memorial representation of the actual movements required in a given situation. According to Schmidt's theory, the precise details of any movement depend on an interaction between the motor schema and the situational specifications of the movement. These specifications include such constructs as the "response specifications, "the "initial conditions" for the response, the "desired outcomes" of the response, the "actual outcomes" of past and present responses, the "sensory consequences" of past and present responses, and the "expected sensory consequences" of the response. The situational specifications not only combine with the motor schema to determine the performance of a response at a given point in time, but also are the variables that, in certain functional relations, constitute the two kinds of motor schema in Schmidt's theory: recall and recognition. Because the motor schema (recognition or recall) is stored in memory without the precise values for its variables, very little space (i.e., capacity of memory or of the CNS) is occupied by the schema, despite the large number of possible movements that the schema can generate in different situations given different values for its variables. Thus, motor-schema theory solves the storage problem. And, this flexibility in motor output as determined by situational demands upon a generalized motor schema also supposedly solves the novelty problem.

With respect to speech production, probably the most interesting aspect of Schmidt's theory is the claim that schemas are not best established via repetitive practice on a particular motor task. Rather, a kind of controlled variability of practice is better:

> One of the major predictions of the recall schema idea is that increased variability in practicing a number of variations of a movement class should result in increased transfer to a new, and as yet unpracticed, member of that same class. (Schmidt, 1976, p. 52)

Motor Memory

Components of Motor Memory. A fairly recent proponent of a multiple-component approach to motor memory is Martenuik (1976), who wrote:

> The transformation of movement information into a permanent internalized form is seen as a two-stage process. First, incoming sensory information is coded by an appropriately developed long-term integrated store that results in a temporary

---

[3] Adams echoed other motor theorists in taking great pains to distinguish knowledge of results from traditional learning theory's response consequences, rewards or reinforcers. We need not be concerned with his efforts in this regard.

[4] Schmidt also admitted this point, and yet seems to adhere to the storage problem as a real one for motor theories other than his.

> code being stored in motor short-term memory. Since there are opportunities for the development of several integrated stores, motor short-term memory is seen as containing multiple codes of movement information. The second stage in the transformation process concerns the development of the motor schema. Here, codes from motor short-term memory are seen as being integrated and abstracted so that the final form of movement information is analogous to an intellectual process capable of producing a wide range of movements. (p. 185)

It is commonly believed that one never forgets how to ride a bicycle. Such a belief, when generalized, yields the unlikely proposition that long-term memory for motor skills is infallible (i.e., in terms of the earlier discussion of LTM, retrieval processes from STM always succeed in finding the correct motor information in LTM). Although the laboratory evidence from continuous motor tasks, like tracking, seems to support the infallibility of long-term memory (e.g., Fleishman & Parker, 1962), that from discrete motor taksks, like typing or linear positioning, does not support this notion (see Adams', 1967, discussion of memory for discrete versus continuous motor tasks).

Baddeley (1976) summarized his brief review of motor STM as follows:

> Although there is clear evidence of short-term forgetting, attempts to explore the short-term motor memory system in greater detail have proved disappointing. It is, however, generally true that forgetting occurs without interpolated activity and is frequently unaffected by interpolated mental activity. Both these results suggest that covert rehearsal is not a major factor, at least in the simple task of remembering the extent of a movement. (p. 260)

Adams and Dijkstra's (1966) replication and extension of Scripture's (1905) experiment was cited by Baddeley as strong evidence that forgetting of a simple, linear, motor response (i.e., running a finger along the edge of a meter rule) occurs over unfilled delays of 20-120. A number of experiments have shown the inability of "filler" tasks such as digit-classification or even writing, during the retention intervals to disrupt motor STM (e.g., Posner, 1967; Posner & Rossman, 1965; Williams, Beaver, Spence, & Rundell, 1969). Nevertheless, other studies (e.g., Kantowitz, 1972; Posner & Keele, 1969; Stelmach & Wilson, 1970; Williams et al., 1969) have shown disruption of motor STM via filler tasks. Thus, one must conclude that motor STM is not impervious to interference but is less affected by various kinds of interference than is verbal STM.

Organization of Motor Memory

As we have seen, students of verbal learning and memory have found that organization is a very salient characteristic of verbal long-term memory. Thus, those who study motor learning and memory have recently begun to investigate the effects of organization (e.g., Diewert & Stelmach, 1978; Gentile & Nacson, 1976; Stelmach, 1977). Diewert and Stelmach reported that subjects' attempts to reproduce five distances to be moved in a linear-positioning task were more accurate if the subjects had initially been exposed to the task "sequentially" (e.g., 10, 20, 30, 40, 50 cm) than if they had initially been exposed to the task "randomly" (e.g., 50, 10, 40, 20, 30 cm). Furthermore, even the subjects who had received "random" exposures tended to sequence or "cluster" (see Bousfield & Bousfield's, 1966, discussion of various metrics of clustering as indices of organization in verbal memory) responses during recall performances, thus demonstrating a kind of spontaneous or "subjective" organization (cf., Tulving's, 1962, use of this idea in verbal-memory theory) of motor memory.

THEORIES OF SPEECH PRODUCTION

Theories of speech production include to greater or lesser extents hypotheses related to issues of verbal and motor memory. Liberman, Cooper, Shankweiler, and Studdert-Kennedy (1967) presented the schematic representation of speech production shown in Figure 2. According to our present division of speech into planning and production stages, only those stages at and below the level in Figure 2 labelled "Neuromotor Rules" are considered to be speech-production stages. We believe the "Syntactic Rules" level in Figure 2 to be a part of speech planning (probably involving an interplay between STM control processes and the highly organized syntactic structures of LTM).

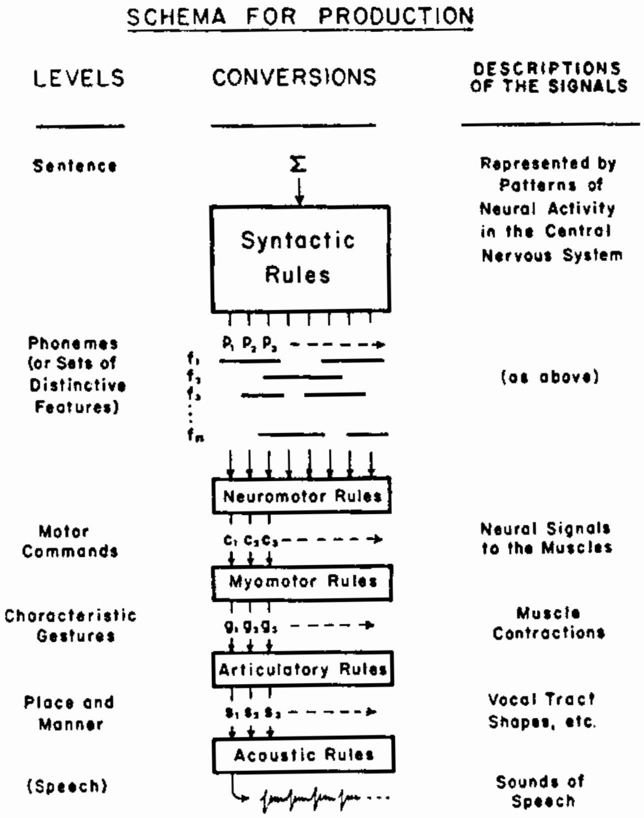


FIGURE 2. Model of speech production from the sentence level to the physical stimulus of the speech signal, (Reprinted with permission from Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967, Perception of the speech code. *Psychological Review*, *74*, 431–461. Copyright 1967, the American Psychological Association.)

Liberman et al. (1967) argued that, once the Syntactic

Rules had been utilized by a speaker to plan his/her sequence of phonemes in an utterance,[5] the transformations of the signals from the Syntactic Rules through the Neuromotor Rules to the Myomotor Rules are invariant. In other words, the contractions of the muscles controlling the articulators (as measured by EMG signals) will be invariant for a given phoneme (or subphonemic feature), regardless of the phonemic context. According to Liberman et al. (1967) context-sensitive variability occurs at the level of the "Articulatory Rules." It is here that spatial and temporal mechanical constraints upon the articulators are imposed by the preceding and following phonemes (or features) in an utterance to force the observed changes in the acoustic product, in spite of the prior invariance of the signals from the Myomotor Rules (i.e., muscle contractions) and the subsequent invariance of the signals from the "Acoustic Rules" (i.e., sounds). This proposal that speech perception and production are determined by invariant motor commands, presumably represented and stored somewhere in memory for each phoneme (or feature), is a characteristic of an open-loop motor-control system.[6]

Wickelgren (1969) invoked a "context-sensitive associative memory" to explain the serial ordering of speech. Specifically, he proposed that the individual sounds of the word are not simply a set of phonemes or subphonemic features, as had been suggested by Lashley (1951) or by Liberman et al. (1967). Rather, each sound is coded in an associative memory with respect to its preceding and following sound, thus yielding an ordered set of allophones, (e.g., /#kr, kru, rus, ust#/ represents the word "crust"). Wickelgren's theory resembles that of Liberman et al. (1967) in that both solve the invariance problem by positing an open-loop, motor-control system for speech. Liberman et al. (1967) assumed that memory stores a large set of motor commands, corresponding to all possible phonemes or subphonemic features of speech, that are issued to the articulators and determine speech movements without recourse to peripheral feedback. Wickelgren assumed that memory stores a large set of context-sensitive allophones that also determine speech without recourse to peripheral feedback.[7]

The basic difference between these theories is that Wickelgren's is more explicitly concerned with memory issues. In particular, the associative chaining of context-sensitive allophones into an ordered sequence of the sounds in a word, described in detail by Wickelgren, has no counterpart within the model proposed by Liberman et al. (1967). At best, the Liberman et al. (1967) theory would handle the serial ordering of speech by reference to the Syntactic Rules stage of speech production in which the phonemes (or features) must be arranged into the correct order before signals are passed to the Neuromotor Rules stage.

MacNeilage (1970) reviewed the EMG research conducted by his colleagues at the Haskins Laboratories and suggested a reformulation[8] of the invariance problem:

> One of the main conclusions of this paper is that the essence of the speech production process is not an inefficient response to invariant central signals, but an elegantly controlled variability of response to the demand for a relatively constant end. (p. 184)

MacNeilage's alternative to the theories of speech production proposed by Liberman et al. (1967) and by Wickelgren (1969) was based on the idea that speakers have "internalized" the "spatial targets" or positions of their articulators prior to an utterance. Supposedly, "phonological information" from what we earlier termed a *speech planning stage* is received by a centrally-located "space coordinate system" for the articulators. This system creates a series of "spatial target specifications" from the received information. The specifications are then transferred as "motor command patterns" to the articulators. Because these motor command patterns are based on the series of spatial target specifications, the sequence of sounds emitted by the articulators reflects the internal spatial targets. Additionally, because these targets are selected well in advance of the actual movements, the necessary modifications of movements required by varying phonemic contexts can be built in, as it were, to the targets and so to the motor command patterns. Thus, both the serial ordering of speech and the invariance problem are explained.

Some motor-control theorists (e.g., Adams, 1976) believe that MacNeilage's ideas constitute an open-loop, motor-control system. Although according to MacNeilage (1970, pp. 189–190), it is true that aspects of his theory (i.e., generation of motor command patterns by the motor system control mechanism) are open-loop in that peripheral feedback is unimportant, MacNeilage, Studdert-Kennedy, and Lindblom (Chapter 4) also believes that some closed-loop control of speech production is possible.

Memory in the Production of Hearing-Impaired Speech

Other than numerous studies about problems of speech production for the hearing-impaired, we are aware of no investigations concerning motor-skill deficiencies not directly related to the articulators that might cause or be correlated with the problems of these speakers. That is, there seems to be no bridge between research on motor learning and memory and that on the motoric abilities of hearing-impaired talkers. Thus, until future efforts are made to discern potential difficulties in the nonspeech motor performance of hearing-impaired people, one can only speculate as to the effects of impoverished auditory experience on the motor-memory system.

CONCLUSION

Most of the questions concerning the role of memory in

---

[5] Technically, they assumed that the phonemes themselves were composed of "subphonemic features" like those discussed by Fant or Stevens (e.g., Fant, 1973; Stevens, 1973).

[6] In their Footnote 30, they explicitly rejected any role for peripheral feedback in the control of speech perception and production.

[7] Though Wickelgren (1969) did not make his denial of feedback control in speech explicit, we think he would have maintained such a position due to his apparent belief that the context-sensitive associative memory suffices to fully determine word pronunciation.

[8] This reformulation, as MacNeilage noted, is actually a special case of "motor equivalence" (Hebb, 1949), which has itself spawned a great deal of theoretical and empirical study in psychology.

speech planning and production remain to be answered. This review has summarized some of the initial investigations in two areas of memory research (verbal and motor) that address some relevant issues. Hopefully future research on speech planning and production will represent a coordination of our knowledge of verbal and motor memory, brought to bear on the problems of speech behavior.

REFERENCES

ADAMS, J. A. (1967). *Human memory.* New York: McGraw-Hill.

ADAMS, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3,* 111–149.

ADAMS, J. A. (1976). *Learning and memory: An introduction.* Homewood, IL: Dorsey Press.

ADAMS, J. A., & DIJKSTRA, S. (1966). Short-term memory for motor responses. *Journal of Experimental Psychology, 71,* 314–318.

ANDERSON, J. R., & BOWER, G. H. (1973). *Human associative memory.* New York: Wiley, Halstead Press.

ATKINSON, R. C., & SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). New York: Academic Press.

ATKINSON, R. C., & SHIFFRIN, R. M. (1971). The control of short-term memory. *Scientific American, 225,* 82–90.

BADDELEY, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18,* 302–309.

BADDELEY, A. D. (1976). *The psychology of memory.* New York: Basic Books.

BARTLETT, F. C. (1932). *Remembering.* Cambridge: Cambridge University Press.

BELLUGI, U., KLIMA, E. S., & SIPLE, P. (1974). Remembering in signs. *Cognition, 3(2),* 93–125.

BELMONT, J. M., KARCHMER, M. A., & PILKONIS, P. A. (1976). Instructed rehearsal strategies' influence on deaf memory processing. *Journal of Speech and Hearing Research, 19(1),* 36–47.

BOUSFIELD, A. K., & BOUSFIELD, W. A. (1966). Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports, 19,* 935–942.

BRANSFORD, J. D., BARCLAY, J. R., & FRANKS, J. J. (1972). Sentence memory. A constructive versus interpretive approach. *Cognitive Psychology, 3,* 193–209.

BROWN, R., & MCNEILL, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

CLARK, H. H., & CLARK, E. V. (1968). Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology, 20,* 129–138.

COHN, R., JAKNIUNAS, A., & TAUB, E. (1972). Summated cortical evoked response testing in the deafferentated primate. *Science, 178,* 1113–1115.

CONRAD, R. (1970). Short-term memory processes in the deaf. *British Journal of Psychology, 61,* 179–195.

CONRAD, R., & HULL, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science, 10,* 135–136.

CONRAD, R., & RUSH, M. L. (1965). On the nature of short-term memory encoding by the deaf. *Journal of Speech and Hearing Disorders, 30(4),* 336–343.

CRAIK, F. I. M., & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception and Psychophysics, 5,* 365–373.

DARWIN, C. J., & BADDELEY, A. D. (1974). Acoustic memory and the perception of speech. *Cognitive Psychology, 6,* 41–60.

DIEWERT, G. L., & STELMACH, G. E. (1978). Perceptual organization in motor learning. In G. E. Stelmach (Ed.), *Information processing in motor control and learning.* (pp. 241–265). New York: Academic Press.

EFRON, R. (1970). The minimum duration of a perception. *Neuropsychologia, 8,* 57–63.

FANT, G. (1973) *Speech sounds and features.* Cambridge: MIT Press.

FLEISHMAN, E. A., & PARKER, R. F., JR. (1962). Factors in the retention and relearning of perceptual motor skill. *Journal of Experimental Psychology, 64,* 215–226.

GENTILE, A. M., & NACSON, J. (1976). Organizational processes in motor control. In J. Keogh & R. S. Hutton (Eds.), *Exercise and sport sciences reviews* (Vol. 4). Santa Barbara, CA: Journal Publishing Affiliates.

HEBB, D. O. (1949). *The organization of behavior.* New York: Wiley.

HINTZMAN, D. L. (1967). Articulatory coding in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 6,* 312–316.

HUNT, N. (1982, June). *Research on memory in deaf children: Implications for teachers.* Paper presented at the Bi-Annual Meeting of the Alexander Graham Bell Association for the Deaf, Toronto, Canada.

KANTOWITZ, B. H. (1972). Interference in short-term motor memory: Interpolated task difficulty, similarity or activity? *Journal of Experimental Psychology, 95,* 264–274.

KATZ, J. J., & FODOR, J. A. (1963). The structure of a semantic theory. *Language, 39,* 170–210.

KEELE, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin, 70,* 387–403.

KEELE, S. W., & SUMMERS, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

KNAPP, H. D., TAUB, E., & BERMAN, A. J. (1963). Movements in monkeys with deafferentated forelimbs. *Experimental Neurobiology, 7,* 305–315.

LASHLEY, K. S. (1917). The accuracy of movement in the absence of excitation from the moving organ. *American Journal of Physiology, 43,* 169–194.

LASHLEY, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–146). New York: Wiley.

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. G. (1967). Perception of the speech code. *Psychological Review, 74,* 431–461.

LING, D. (1978). Forward trends—Toward order in teaching verbal skills. *The Teacher of the Deaf, 2,* 110–116.

MACNEILAGE, P. F. (1970). Motor control of serial ordering of speech. *Psychological Review, 77,* 182–196.

MACNEILAGE, P. F., & MACNEILAGE, L. A. (1973). Central processes controlling speech production in sleep and waking. In F. J. McGuigan (Ed.), *The psychophysiology of thinking* (pp. 417–439). New York: Academic Press.

MANDLER, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation.* (Vol. 1) (pp. 328–372). New York: Academic Press.

MARTENIUK, R. G. (1976). Cognitive information processes in motor short-term memory and movement production. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 175–186). New York: Academic Press.

MASSARO, D. W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology, 85,* 411–417.

MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior, 2,* 1–21.

MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

OLSSON, J. E., & FURTH, H. G. (1966). Visual memory-span in the deaf. *American Journal of Psychology, 79,* 480–484.

PETERSON, L. R., & PETERSON, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 58,* 193–198.

PEW, R. W. (1974). Human perceptual-motor performance. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* New York: Erlbaum.

POSNER, M. I. (1967). Characteristics of visual and kinesthetic memory codes. *Journal of Experimental Psychology, 75,* 103–107.

POSNER, M. I., & KEELE, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

POSNER, M. I., & KEELE, S. W. (1969). Attention demands of movements. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: Swets & Zeitlinger.

POSNER, M. I., & ROSSMAN, E. (1965). Effect of size and location of informational transforms on short-term memory. *Journal of Experimental Psychology, 70*, 496–505.

QUILLIAN, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 216–270). Cambridge: MIT Press.

RUMELHART, D. E., LINDSAY, P. H., & NORMAN, D. A. (1972). A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press.

SCHMIDT, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225–260.

SCHMIDT, R. A. (1976). The schema as a solution to some persistent problems in motor learning theory. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 41–65). New York: Academic Press.

SCRIPTURE, E. W. (1905). *The new psychology*. London: Scott.

SHIFFRIN, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science, 168*, 1601–1603.

SHIFFRIN, R. M. (1976). Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 177–236). Hillsdale, N.J.: Lawrence Erlbaum Associates.

SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

STELMACH, G. E. (1977). Prior organization in motor control. *Journal of Human Movement Studies, 3*, 157–168.

STELMACH, G. E., & WILSON, M. (1970). Kinesthetic retention, movement extent, and information processing. *Journal of Experimental Psychology, 85*, 425–430.

SUSSMAN, H. M. (1972). What the tongue tells the brain. *Psychological Bulletin, 77*, 262–272.

TAUB, E., BACON, R. C., & BERMAN, A. J. (1965). Acquisition of a trace-conditioning avoidance response after deafferentation of the responding limb. *Journal of Comparative and Physiological Psychology, 59*, 275–279.

TAUB, E., & BERMAN, A. J. (1963). Avoidance conditioning in the absence of relevant proprioceptive and exteroceptive feedback. *Journal of Comparative and Physiological Psychology, 56*, 1012–1016.

TULVING, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review, 69*, 344–354.

TULVING, E., & DONALDSON, W. (1972). *Organization and memory*. New York: Academic Press.

WICKELGREN, W. A. (1965). Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology, 70*, 102–108.

WICKELGREN, W. A. (1969). Context-sensitive coding, associative memory, and serial ordering in (speech) behavior. *Psychological Review, 76*, 1–15.

WILLIAMS, H. L., BEAVER, W. S., SPENCE, M. T., & RUNDELL, O. H. (1969). Digital and kinesthetic memory with interpolated information processing. *Journal of Experimental Psychology, 80*, 530–536.

WILSON, D. M. (1964). The origin of the flight-motor command in grasshoppers. In R. F. Reiss (Ed.), *Neural theory and modeling* (pp. 331–345). Stanford, CA: Stanford University Press.

WILSON, D. M. (1965). The nervous coordination of insect locomotion. In J. E. Treherne & J. W. L. Beament (Eds.), *Physiology of the insect central nervous system* (pp. 125–140). New York: Academic Press.

WINOGRAD, T. (1972). *Understanding natural language*. Edinburgh: *Edinburgh University Press*.

GROUP DISCUSSIONS

GROUP #1: SPEECH PLANNING AND PRODUCTION: STEP-BY-STEP DEFINITION OF THE TASK.

Members: Peter F. MacNeilage (Chairman), Jean Moog, Mary Ellen Nevins, Katherine Harris, Colin Painter, Stefanie Shattuck-Hufnagel, Judith L. Lauter (Editor)

P. MacNeilage (PFM) began the discussion by suggesting that there were three levels of activity involved in speech production: (a) cognitive processes involving ideas and meaning, a level that is not language specific; (b) an intermediate stage, which is language-specific and includes a lexicon with many subdivisions such as content versus function words, etc.; and (c) muscle fibers which constitute the "bottom end of the system." Reactions to this outline included S. Shattuck-Hufnagel's (SSH) observation that the intermediate level must also involve phonological representation, and C. Painter's (CP) suggestion that language-switching might be mediated at the highest level. All agreed it was next to impossible to discuss the details of "intention." We should focus our attention for now on the other two levels.

CP said there must be several levels between PFM's intermediate and "bottom end," and that one might use *langue* to refer to upper and intermediate levels, and *parole* to refer to the physiological realization of *langue* into eventual movements. PFM speculated that the operations related to intention and manifestation ("competence and performance") are kept separate early in the course of planning and producing speech, and that only at later stages are the two sets of operations blended.

A brief mention of how speech errors bear on this distinction between levels of processing led to a discussion of errors made by deaf talkers. J. Moog (JM) noted that deaf children will place verb endings on nouns, a type of error that does not occur in slips of the tongue like those SSH studies. Also, deaf children employ a kind of "telegraphic speech," leaving out function words, but maintaining a sentence contour with continued phonation or pauses where omitted speech sounds would be. JM called these "space holders." J. Lauter (JLL) suggested that this may be evidence that the larger pattern of a sentence (its prosodic structure) is planned separately from the sequence of segmental elements of phonemes and syllables.

Other types of errors—made by hearing children—such as "goed" for "went" are seldom made by deaf children. JM said deaf educators used to be proud of this; now they know lack of such errors indicates a failure in the deaf child's acquisition of grammatical rules. CP suggested these types of errors represented a kind of overloading, which could also account for the errors seen in second language learning. SSH observed that some form of "overloading" might be involved in the difference between the number of errors made in formal and informal speech situations.

Discussion of the details of deaf and hearing errors led PFM to suggest an elaboration of the three levels of planning and production: (a) intention, (b) a morpheme-ordering mechanism, (c) a segmental/phoneme ordering mechanism to fill in the "frames" of (b), and (d) organization of motor control. SSH offered some details regarding the segmental organization within a sentence, and the errors that interchange segments. A question by D. Calvert (DC) as to whether such errors occur in deaf children's speech (our deaf educators C. Nevins (CN) and JM said no) led to a discussion of how speech errors reflect language development. Segmental errors, common with adults, do not seem to occur in children until about 5 years of age. K. Harris (KH) wondered whether children learn "word shapes" or phonemes, and PFM and CP seemed to think one could describe stages of "phoneme detachability" through language acquisition. For example, Leopold (1947) reported that his daughter could produce "pretty" perfectly at a very early age, but "lost" this good production as she learned the rules for pronouncing the individual phonemes, and how they had to be produced in different sequences.

It is possible of course that there are individual differences in how people produce segmental errors (e.g., in treatment of clusters, but the evidence needed is lacking). KH mentioned that a suprasegmental window for planning might involve breathing, since linguistic structure can be observed to be matched to breath patterns. Thus one can describe the biological mechanism of a "default intonation contour," as breath is slowly released through a tensed glottis. JLL noted that there is evidence that deaf talkers often do not show such language/respiratory matching. It is not known whether deaf speakers can be taught to "breathe-for-speech" and thus help their speech intelligibility, though JLL notes that an observation by Tom Hixon indicates that this may be a promising way to approach the problems of speech planning and production faced by individuals lacking auditory feedback. CP noted that laryngectomees read with normal respiratory "speech breathing," even though it is not needed by them since their voicing is supplied not by pressure changes across the laryngeal folds, but by an electrolarynx.

We embarked on a short discussion about whether the "spreading" phenomenon of speech errors is based on features or phonemes. JLL suggested that planning must take biomechanical constraints into account, and that analysis of features that do "spread" should reflect such limitations on possible movement combinations. SSH noted that errors can affect speech movements not only at the level of features, but also above this, involving phonemes as a whole, clusters, etc.

At our second meeting, the discussion began considering "what is controlled" in speech production (i.e., what is the object of planning). Possible candidates are: velocity, displacement, cavity size, size/shape, rate of change, points in relative space, and sound targets. Given some or all of such goals, how does the motor system achieve them? One must also be able to take into account the fact that movement planning uses both stereotypy and flexibility. JLL and SSH noted that a speaker is probably always compensating to some extent, and "never does the same thing twice." JLL noted that there is no direct (i.e., monosynaptic) path between motor cortex and the periphery, that all fibers are involved with collateral interactions within the brain stem. This point is related to issues of motor control "loop" mechanisms. PFM pointed out that there are three alternatives for describing how the speech production periphery, as a motor system, works:

1. Open loop, with no feedback (in motor systems, a good design for fast 'ballistic' movements where no need for adjustment is anticipated);
2. closed loop with constant on-line adjustment, and
3. interactive: predominantly open loop, but with closed-loop connections available for adjustments.

PFM believes that number 3 has the least explanatory power, but JLL thought that this was not only the best description of not only the motor system, but related to sensory system function, as well: analogous issues involve the contributions of top-down versus bottom-up mechanisms.

PFM suggested we close by getting back to the top end of the system. He said that consideration of even the periphery must take into account that everything done in speech production is based at least in part on the "knowledge of the listener." JLL observed that this was a natural part of a system that involved interaction with outside influences, and so should depend like many other biological systems on "redundancy-reducing operations." PFM wondered where this awareness of the listener could go in the model we are evolving.

SSH suggested that the speech production mechanism "sets a dial" ahead of every step taken, as part of "intention," thus providing two ways to run the system, either with little feedback (sloppy running but fast) or with a lot of feedback (to be sure of getting a clear output). PFM said there might be a preprogrammed version of each event, even to the extent of storing separate entities for the variations of "quick speech," such as doncha, wanna, etc. KH noted that tradeoffs between amplitude of movement excursion and speed of movement were related to this, and JLL offered the example of the word father spoken quickly with a flap in the middle: this is not an error, just a tradeoff between amplitude and speed, with an undershot target for more speed. SSH asked if it was possible that production might involve both over- and undershooting.

PFM brought up one kind of redundancy problem related to deaf speech, in which function words are often omitted. JM said that unstressed syllables are omitted in general, and that function words will stay in, "once the child knows they're there," but the teacher must emphasize them in the beginning. She observed that -ing is easier for deaf children to learn to include than are plurals or past tenses; JLL suggested this was because it was an additional syllable (even though unstressed), and CN suggested this was because a participle is marked with a modal: it has a syntactic companion, and the child learns the two-part pattern.

In a final return to the highest levels of the system, CP suggested that some of the decisions made here involve such issues as, what language do I want to speak? What dialect? What style? Shall I speak as myself or mimic someone else? Shall I sound like myself sober or drunk? etc. PFM agreed that such decisions must be among those made early (i.e., high in the system), but that they can affect one, few, or many of the other levels, depending on how much must be changed: syntactic patterns, lexical constraints, phonological representations, etc.

PFM's summary of the group's deliberations stressed the notion of levels of planning/production, and what is included at each level. The levels are: nonverbal intention, mental lexicon (consisting of the stems of content words), and the execution mechanisms of the motor system. Speech errors can be used to help us think about intermediate levels, and lead us to a qualification of the definition of the lexicon level: intention leads to a selection of both (a) lexical items, and (b) a syntactic frame of function words. For the lexical stems, representation is understood as consisting of a segmental list; errors such as spoonerisms affect this level of planning. From this point on, the motor system "realizes phonological structure." As to how the motor system accomplishes this, the options are: open-loop only (everything is "preprogrammed"), an extensive system of exclusively closed-loop connections (feedback and feed-forward), and some combination of the two. Finally, it should be emphasized that the speaker plans and produces with a listener in mind, and reduces redundancy (in terms of everything from semantics to phonology) to the extent that he expects it will not be needed.

REFERENCES

LEOPOLD, W. F. (1947). *Speech development of a bilingual child.* Chicago: Northwestern University Press.

GROUP #2: THE ROLE OF MEMORY IN SPEECH PLANNING AND PRODUCTION

Members: George E. Stelmach (Chairman), Myra Aubuchon, Ann Geers, Ruth Geier, Ginger Kuehn, James D. Miller, Mary Joe Osberger, Joan Slein, Craig A. Bowe (Editor)

The focus of our discussion about the role of memory in speech planning and production was largely practical. That is, we were mostly interested in discovering how facts, theories, and ideas about memory (verbal and motor) can help us design more effective methods of teaching hearing-impaired people, in particular, hearing-impaired children, to speak easily and intelligibly. We therefore have nothing to report about the more "basic science" issue of what role memory plays in the planning and production of speech by normally-hearing people (for relevant findings, see Bowe, Chapter 11).

The discussion resulted in four suggestions about where future efforts might be directed to determine the nature of the interactions between memory processes of hearing-impaired children and their speech. Each of these suggestions is summarized below.

Suggestion 1: Rather than extensive efforts being directed at modeling or prompting various aspects of speech in the hearing-impaired child, more effort should be directed at establishing self-initiated speech. This suggestion is motivated by a number of research findings that self-initiated behavior is in general more easily organized and thus more easily remembered.

Suggestion 2: We need more data concerning errors made in the speech of hearing-impaired speakers. Specifically, we need to know about those errors where there is independent evidence that the speaker "knows" how to produce the correct sound. As one example of this kind of error, many hearing-impaired children can be taught to produce a highly intelligible "s" (e.g., as in "yes"), and yet they may often fail to produce this sound in the appropriate circumstance. Systematic data collection about such errors could prove very useful in diagnosing particular deficits in memory that might then be remediable.

Suggestion 3: Research concerning motor learning and memory, in particular the research related to Schmidt's (1975) "schema" theory (see Bowe review, Chapter 11), has indicated that memory for certain actions can be enhanced by a kind of controlled variability of practice. The idea is that during the learning of a given target action the learner should practice by performing actions that vary by known (via appropriate feedback to the learner) amounts around the target action. This kind of learning experience, when contrasted with the more typical rote repetition of the target action, has been shown in certain circumstances to result in better memory (and so, better performance) of the target action. The implication of all this for speech training of hearing-impaired children is that appropriate feedback systems should be developed and employed to allow the children to practice "controlled variability" around target speech sounds. Assessments could then be made to determine whether such training proves more effective than the usual attempts to practice only the specific target sound.

Suggestion 4: The sounds of speech carry a great deal of information. To plan any given utterance, therefore, requires that the information represented in the sound be appropriately "chunked." Otherwise, without the ability to chunk information, the stream of sound that is speech would exceed the limited capacity of our memory systems to process information, and we would understand little of it.

Because the ability to chunk information is so vital for our memory systems, its study is important for the understanding of how memory affects speech planning and production. Unfortunately, rather little research concerning chunking has been conducted using hearing-impaired subjects (but see Bowe, Chapter 11, for some relevant references). We will need much more of such research before we can design teaching strategies that will help hearing-impaired children to remember how to speak effectively.

REFERENCES

SCHMIDT, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82, 225–260.

GROUP #3: THE ROLE OF SENSORY FACTORS IN SPEECH ACQUISITION, MONITORING, AND PRODUCTION

Members: Arthur Boothroyd (Chairman), Donald Eldredge, Robert Gilkey, Robert Lutfi, Gerald Popelka, Ellen Rajtar, Karen Stein, Marcia Tash, Kenneth Grant (Editor)

This discussion group dealt with sensory feedback and its role in learning to talk. The group discussed a number of issues relevant to this topic, and generated a catalog of research areas that need to be explored before we can fully understand the many variables that influence a talker's awareness of his productions.

There are several channels of sensory feedback that can help a talker shape his/her productions. The most important seems to be audition. Also, there are channels connected directly with the motor system itself, such as somatosensory and kinesthetic feedback that can help a speaker learn the "feel" of a production. The use of these mechanisms for speech monitoring may involve not only sensations associated with the upper vocal tract (jaw movements, velar position, etc.), but may also depend on sensations from lower parts of the speech production system (i.e., those structures and physiological controls dealing with respiration). Finally, visual cues can be used by talkers to monitor speech production. Hearing-impaired individuals have long been taught to lipread others' productions for comprehension, and to try to mimic visible (and tactual) vocal-tract shapes and movements themselves.

There are many details associated with each of these modes of feedback, and most are poorly understood. Popelka presented one example of a common misunderstanding about a loss in auditory feedback. He pointed out that it is generally believed that in children, problems with producing speech are highly correlated with conductive loss. However, a simple comparison of a typical conductive audiogram and the energy at different frequency regions within the speech spectrum indicates that for such children, much of speech is still audible. He suggested that more sensitive tests of hearing capability, such as measures of most comfortable and uncomfortable listening level, would be more predictive of speech problems associated with conductive loss. Thus simple descriptions of a subject's ability to detect sounds are not sufficient to help us understand how auditory feedback contributes to learning to talk; rather, we need to ask more sophisticated questions such as "Are loudness cues important?", and "Do the psychophysical characteristics of the auditory system change with signal level?"

Also, we need to know whether the auditory capacity of a cochlea damaged in sensorineural hearing loss is specific to features of speech. Little work has been done relating different types of sensorineural hearing loss to feature analysis in speech perception. This could be of obvious importance for children who suffer from sensorineural problems, with or without conductive losses.

Attempts to measure hearing in terms of psychoacoustic abilities must be pursued in the context of the relevance of these abilities to speech perception. The integration of psychoacoustic aspects of sounds with speech perception has yet to be accomplished. Only when this is done will we be able to understand how to extrapolate from "hearing tests" to predictions of speech hearing.

Grant emphasized the importance of understanding how to recode speech from its auditory form into other modalities. He pointed out that the quality and salience of recoding schemes will depend on the psychophysics of the target modality, and on its pattern-recognition characteristics. If a modality is to be used for recoding speech sounds, we need to know its basic "map" and salient proportions (intensity, spatial dimensions, etc.), and how the modality interacts with other modalities. Recoding problems are present even within one modality, such as when researchers attempt to extract auditory cues from speech to present as aids for lipreading. Grant's work with extraction of fundamental frequency cues, in terms of both spectral and amplitude information, indicates that such aids may have the capacity to be tailored to fit different hearing losses, and have the potential to provide significant boosts in speech comprehension.

Boothroyd pointed out that we know little about the role of attention in integrating hearing with other aspects of perception factors, especially regarding the stability of auditory input. The problem of stability is an important one for children who have fluctuating conductive hearing loss, and for both children and adults with hearing aids, the performance of which may fluctuate over time due to mechanical limitations.

The interaction between sensory modalities as they work in concert to provide cues to the talker has not been studied in detail. We need to know how well residual hearing can be integrated with other sensory channels. For example, lipreading is commonly used to supplement deaf children's hearing capacity, and yet we know little about auditory/visual interactive psychophysics. We also know little about how acquisition of phonology is related to self-monitoring of signals in motor and sensory feedback channels. It is also possible that extrasensory factors such as attention may be important in modality interactions.

Finally, the experience of teachers of the hearing impaired indicates that there are everyday practical problems yet to be solved. For instance, when asked "What do you need?", Rajtar reported that she would like a device to use in the classroom that would provide specific information for each day's task: for one session, an aid to help the child perceive /s/; on another day, an aid to help him/her discriminate /a/ versus /i/, and then go on to learn /i/ versus /ɪ/. Boothroyd noted that we don't yet have an answer to the all-important educational issue of "carryover." Children who learn a speech sound in a classroom situation may fail to produce that same sound in "communication situations" outside structured learning sessions. Factors contributing to this often-observed phenomenon may include:

1. Children might be less motivated outside a structured situation.
2. Teachers often allow a child to produce ("practice") many more instances of an erroneous production than of a correct one—the teacher asks for repetitions until a correct target is achieved, says "good" and goes on to the next item.
3. Children learning to speak may depend on the teacher as a "monitoring device," as a substitute for their inadequate sensory feedback systems.

We need to find ways to solve these problems, including means by which children can self-monitor their own productions, so that learning to talk can proceed whenever the child is awake and communicating.

GROUP #4: INSTRUCTIONAL STRATEGIES FOR TEACHING DEAF CHILDREN TO SPEAK

Members: Ray D. Kent (Chairman), Chris Clark, William Clark, Saundra Daugherty, Mary Lou Koelkebeck, Nancy McGarr, David Mason, Michael Mudrovic, Cindy Thomas, and Roanne Karzon (Editor)

Speech and language are naturally and spontaneously acquired in the normal-hearing child. In contrast, the deaf child has, by virtue of his or her sensory deficit, missed many months of auditory stimulation. Although rudimentary communication strategies within families usually develop to substitute for the lack of an oral-aural channel, the deaf child cannot learn speech and language spontaneously. At some point the loss is diagnosed, an aid is fitted and the formal training necessary for speech and language development begins. With respect to teaching speech and language, the initial question is—to what extent should normal development serve as a model for teaching speech to the deaf child? According to C. Clark, the present strategy is to follow the pattern of development for normal-hearing children, but to consider fully the consequences of the auditory deficit. For the most part, this approach suggests that the auditory deficit has delayed speech and language acquisition and, therefore, the proper approach is to systematically teach speech and language according to the known norms of development. For example, when choosing which speech sounds to work on first, the teacher is guided by two principles. First, choose from among the first speech sounds acquired by normal-hearing children. Second, because of the auditory deficit choose those sounds which are highly visible.

Contrary to this view, it was mentioned that there is some evidence that deaf children develop a pattern of vowel production based on three target positions. Thus, their pattern or vowel development may be different from that of normal-hearing children. If a child's speech and language is different rather than delayed in some areas of phonology and grammar, then perhaps a different overall approach to training is needed.

Marked variability in performance was observed during the training films shown during the seminar. At present, we do not know whether this variability is greater among deaf children with deficient articulation then with normal-hearing children with deficient articulation. Actually, the variability may facilitate instruction by allowing flexibility in attempts to approximate the target production.

PERCEPTION

At present audiologic evaluation provides insufficient information with respect to the speech processing ability of severely and profoundly hearing-impaired children. Mason pointed out that recent work by Charles Berlin of the Kresge Research Laboratories indicates that some hearing-impaired individuals have usable hearing in the high frequencies, beyond the standard audiometric frequencies (i.e., >8000 Hz). Teachers of the deaf need to know which patterns and features of speech can be perceived auditorily by the hearing-impaired student in order to develop effective strategies to teach speech production. There are few tests of speech discrimination and recognition that are appropriate for use with the child with a severe to profound hearing-impairment. It is also important to understand the relative contributions made by speechreading and tactile aids to receptive language processing.

Research has demonstrated that children with severe and profound sensorineural hearing loss have poor frequency resolution. It was speculated by Karzon that selective enhancement of intonation contours in the speech addressed to deaf children may assist in grouping syntactic units receptively and in achieving more varied intonation patterns in speech production. Grant at CID is currently investigating the effects that augmented pitch contours have on speech perception in adult listeners with severe-to-profound hearing impairment. Koekelbeck expressed concern with regard to the effects that speech models with exaggerated intonation may have on the training of speech production training of deaf children.

With respect to time and intensity, Erber has written that even profoundly deaf children can use these cues to process speech. Therefore, it is possible that in some cases these cues may be emphasized to teach prosodic patterning. Rather than relying on fundamental frequency as the key to stress in English (as normal listeners do), the work of Rubin and McGarr suggests that the hearing-impaired child may need to focus on the intensity and timing of the speech sequence to process the prosody of the speech. Although, there are no formal studies comparing the prosodic patterning of teachers of the deaf to that of teachers of normal-hearing students, W. Clark noted that the structured approach to teaching speech may inadvertently result in exposure to a restricted set of prosodic patterns. Thus, deaf students may benefit from both enhancement of selected prosodic features and modeling of a variety of prosodic patterns. Although there is some ongoing research on the ability of deaf students to perceive various features of speech, much more empirical work is needed before the findings can be applied to teaching speech to the deaf.

PRODUCTION

Teachers in the working group indicated that the primary goal for training of speech production is *intelligibility*, with voice quality as a secondary concern. However, S. R. Silverman pointed out that there is a strong relationship between articulation and voice. For example, excessive air-flow on stop consonants leads to insufficient air for the vowels that follow. Therefore, articulation and voice training need to be combined to produce the best possible speech.

McGarr reported that the physiology of deaf speech has been studied only in a very basic way. Recent work by

McGarr and Höfqvist (1982) showed that the interarticulator timing between vocal fold abduction/adduction and oral articulation is often inappropriate in the speech of the deaf.

TRAINING

Reinforcement

Observation of speech/language training films of CID teachers and students prompted W. Clark to comment on a common pattern of reinforcement. Often the teacher would elicit several productions until an acceptable token was produced and then she would positively reinforce the child and progress to the next task. There are two potential problems with this approach. First, the reward for the child may be "stopping the drill" rather than producing an acceptable response. Second, one acceptable production amidst several less optimal productions does not allow the child the opportunity to learn what contributed to the acceptable production. It was suggested that practice should continue until several acceptable productions are obtained, indicating that the child has learned how to produce an acceptable production at will.

The use of parents to reinforce speech productions at home was discussed. The clinicians and teachers of the group stated that without guidance many parents tend to be too critical of speech patterns and do not supply adequate positive reinforcement. Therefore, it was recommended that parents be instructed with respect to methods of working with their child. Furthermore, the tasks assigned need to be carefully selected; assignments should consist of activities the child has mastered but needs to practice. Thus, homework for speech production is primarily a device for carryover activities. However, parents are also excellent resources for vocabulary development.

Sensory Devices

Hearing aids have been greatly improved during the past decade. Combined with improved methods for fitting of aids, all but the profoundly deaf can rely primarily on the auditory channel for speech and language training. Future research may investigate other signal processing techniques to enhance the audibility and discriminability of a variety of speech features. For example, if expansion in the frequency domain facilitates speech perception, perhaps aids could be developed to automatically expand this dimension. For the profoundly deaf who have little or no residual hearing, tactile aids promise to be a useful supplement to visual processing to teach many of the features of speech.

McGarr emphasized the need for interaction between teachers and technologists in order to develop sensory devices that provide adequate feedback. However, the device should not replace the teacher who must provide speech training and must know how and when to use speech aids.

Computers

Kent emphasized the many potential applications that computers have in the speech training of deaf children. Computers are now within the range of clinical and school budgets and have already been successfully applied in teaching reading and other areas of special education in the public schools.

The working group agreed that computers should be a supplement for the teacher, not a replacement. Children would use the computer to solidify their mastery of a particular speech task and to generalize it to a variety of different contexts.

The computer can be programmed to collect and analyze response data as well as to reinforce correct production of speech sounds. McGarr cited a study by Boothroyd which showed that in traditional speech training, teachers talk twice as much as the students. With a computer, a greater percentage of the training session would consist of student vocalizations. Working for even brief periods with a computer, a child can experience many trials. Because of the computer's finely tuned data analysis and increased number of trials, teachers may be able to see small changes in performance on a daily or weekly basis that would otherwise go undetected and unrewarded. However, one problem with implementing computers for speech production training is the need to develop objective criteria for judging the student's performance.

REFERENCES

ERBER, N. P. (1982). *Auditory training*. Washington, D.C: A. G. Bell Association.

MCGARR, N. S., & LÖFQVIST, A. (1982). Obstruent production in hearing-impaired speakers: Interarticulator timing and acoustics. *Journal of the Acoustical Society of America*, 72, 34–42.

GROUP #5: STRATEGIES FOR SPEECH TRAINING AND SPEECH MONITORING AIDS FOR THE DEAF

Members: Kenneth N. Stevens (Chairman), Margaret W. Skinner, A. Maynard Engebretson, Christine Gustus, Arnold Heidbreder, Victoria Kozak, Arthur R. Niemoeller, Deborah Servi, Janet M. Weisenberger (Editor)

The task assigned to Working Group 5 was to evaluate the present status of, and make recommendations of future directions for, speech training and speech monitoring aids for the deaf. The comments of the working group members are summarized below, by topic.

ADVANCES IN HEARING AID TECHNOLOGY

The first issue around which discussion centered concerned the question of what speech information hearing-impaired persons need beyond what can be provided by a hearing aid. Engebretson stated that the technology available for the design and construction of hearing aids, particularly in the areas of signal processing and packaging, was more than adequate, but that more information was needed regarding what improvements would provide most benefit to the hearing-impaired population. Some suggested improvements included greater ease of use, particularly for the very old and very young hearing-impaired populations who find it difficult to make manual adjustments on small hearing aids, and the use of computer software programming of hearing aid characteristics to allow for adjustments required by changing environmental characteristics such as noise and reverberation. Still other issues were improvements in cosmetic appeal by size decreases, possibly to the point of having an "implantable" aid in the ear canal; remediation of overdrive distortion; increases in the number of independently addressable frequency channels; and elimination of feedback to the wearer arising from his own productions.

The discussion then turned to group hearing aids, used mostly in classroom settings. Niemoeller stated that to his knowledge there had not been much research into ways of improving group hearing aids since the development of infrared systems several years ago. Because of this relative lack of research, the gap between the performance of the group hearing aid and the individual wearable aid is decreasing, due largely to advances in wearable aids. However, the consensus was that there were still some advantages to using group aids.

The group concluded that for persons with any residual hearing, the conventional hearing aid was still superior to aids utilizing any other modality, and the fact that more and more hearing-impaired adults seek and use hearing aids than ever before suggests that users are deriving benefit from them.

Tactile Aids for Speech Reception and Speech Production

Despite the enormous advances in hearing aid technology in recent decades, there are still a substantial number of persons whose hearing loss is so profound that they derive essentially no benefit from conventional auditory amplification. For these persons, the idea of "sensory substitution" of an intact modality to transduce the information normally provided by the impaired modality may be feasible. Weisenberger stated that recent developments in the design and evaluation of aids using the tactile system as the substitute modality showed considerable promise for the application of such devices for deaf persons. However, the relatively small number of researchers in this field and the resulting lack of large amounts of data mean that many parameters remain at present untested, including which speech parameters to encode, what part of the body to stimulate, what type of stimulation to provide (vibrotactile vs. electrotactile), how many channels to use (single vs. multiple), what type of signal (sinusoid, noise, amplitude modulation, frequency modulation, frequency as frequency or as place, time variation, etc.), and what type of training procedure to employ. In addition, there is still a relative lack of information about the basic capabilities of the tactile system to provide insight for the design of tactile aids.

Heidbreder described some major technological limitations of currently available tactile aids, including the lack of an adequate transducer that is quiet, has good frequency response, is small and lightweight and has low power consumption and thus a long battery life. These limits have thus far kept most tactile aids in the laboratory and not allowed their use by large numbers of the deaf population.

The discussion then turned to a comparison of the relative merits and disadvantages of single-channel and multichannel tactile aids that have been tested to date.

Single-channel Aids

It was noted that single-channel tactile aids could be useful for profoundly deaf persons. Weisenberger suggested that a single-channel aid was most helpful in providing information about the occurrence of sounds in the environment, and about simple aspects of the rhythm and voicing of sounds. Thus, a single-channel aid would have the greatest advantage for a child at the very beginning of language acquisition, and might decline in usefulness as the child proceeded to more complex tasks.

Gustus described a case study currently underway at CID in which a 2-year-old profoundly deaf child was given a commercially available single channel aid, the Tactaid I. The child, at the beginning of training with the aid, had few vocalizations and little or no response to sound through conventional amplification. After approximately 6 months using the tactile aid, her receptive vocabulary had increased remarkably, and her vocalizations improved to the point where she could be taught words. Weisenberger noted that similar results had previously been obtained by Moise Goldstein at Johns Hopkins with another child of comparable age and hearing loss.

Servi suggested that one might set up a protocol for progressively more sophisticated tactile aids to be given to a child as he moved into progressively more complex stages of language learning. Kozak added that very simple single channel aids can become largely superfluous once a task has been learned, since many children seem to generalize simple tasks to the nonaided case with no difficulty. However, Weisenberger noted that the results from studies of lipreading with and without single-channel tactile aids show a clear advantage for the case in which the tactile aid is used, and thus there is a use for the single-channel aid for persons who have a good command of language as well as for the beginning learner.

Multichannel Aids

A number of concerns were discussed with respect to the design and evaluation of multichannel aids. First, the problems previously described for the single channel case (i.e., transducer problems, wearability) are now essentially multiplied by the number of channels contained in the aid. A more serious concern was the possibility that extensive training might well be necessary for optimal use of a multichannel tactile aid, since most of the designs that had thus far shown promise had done some sort of mapping of auditory frequency onto tactile place on the skin, and this constituted a very extensive recoding of the speech signal.

Engebretson raised a question about the use of electrical stimulation rather than vibratory stimulation, particularly with children. The group discussed the limited dynamic range and potential difficulties of maintaining consistent and safe levels of electrotactile stimulation. It was noted, however, that the electrotactile belt array developed by Saunders (1980) utilized a biphasic electrical pulse that was relatively safe, well-grounded, and not painful, and that this method might hold some promise.

Weisenberger noted that deaf-blind users of the Tadoma method of speech reception seem to have both better speech perception and better speech production than many deaf children, so that it appeared that the tactile system could be very useful in providing speech information to an individual, even if presently available mechanical tactile aids did not produce overwhelming results. Thus, research into their capabilities should be continued.

The major problem, according to Kozak, was that current tactile aids simply were not available to teachers of the deaf, who were eager to use them in the classroom if even a small improvement could be demonstrated for the aided case. Kozak complained that most tactile aids never left the laboratory, and that their long-term possibilities had not been tested. Weisenberger responded that there was not a large enough demand for tactile aids for the deaf to interest businesses in the commercial feasibility of investing the necessary expertise, time and money for their further improvement, miniaturization, and wearability, and that this factor might account for the relative lack of progress in this field as compared to hearing aid development.

Stevens suggested that it might be reasonable to approach the federal government for funds to develop and market a good tactile aid, in a fashion similar to the "orphan drug" strategy, in which the development and production of drugs to alleviate rare diseases was federally subsidized even though the commercial benefits were negligible. The difficulty arose in deciding exactly which tactile aid was the most promising for intensive development, and then of convincing the government of the need for such a device.

The group consensus was that tactile aids showed at least as much improvement in speech perception as was found with the cochlear implants studied to date, and therefore it was important to continue research into the design and evaluation of new tactile aids, particularly in view of their noninvasive character.

Visual Aids for Speech Reception and Production

Stevens noted that another alternative modality for presenting speech information was the visual system. Kozak and Gustus argued that the visual system was extensively used in lipreading by deaf children to great advantage, and that attempting to provide additional information might reduce the information gained by lipreading. For this reason, a visual aid for the deaf might be most appropriate as a training aid rather than a conversational aid.

Just as with multichannel tactile aids, observed Stevens, a substantial amount of training was necessary in order to use most multichannel visual aids, because of the extensive recoding of the signal involved. In fact, it was theoretically possible to extract any desired information from the acoustic signal that one wished, and to display it visually.

Some suggested parameters for a visual display were: (a) fundamental frequency (in a pitch vs. time contour); (b) nasalization (measured by an accelerometer on the nose); (c) breathy voice quality; (d) amplitude; (e) instantaneous spectrum. Stevens pointed out that several of these had been or were currently being investigated, and that while they showed some promise, none had thus far met with overwhelming success.

Stevens asked whether visual speech aids are "better" than teachers of the deaf in training a specific skill. Kozak answered that she did not envision the visual aid as a replacement for the teacher, but rather as a supplement to the teacher, particularly to help the child practice skills already learned and improve skills through simple repetition. Thus, the teacher's time could be spent in transmitting new skills to the child, rather than on drills.

Evaluation and Training Strategies for Sensory Aids

The final topic discussed by the working group was the question of how to evaluate speech aids for the deaf and how to determine the optimal training procedure for a given aid.

The general feeling of the group was that the evaluation of the capabilities of any aid would be considerably facilitated if there were a general test battery designed to assess all sensory aids, auditory, tactile, or visual. It was decided that such a test battery should include both basic psychophysical measures and various tests of speech perception.

It was noted that some of the simplest sensory aids would not yield any results with complex speech tests, so that one should be able to select appropriate subtests from this general test battery without reducing its validity.

The use of a general test battery for aid evaluation would make possible direct comparisons of different aids, something which cannot at present be done, and would quickly show the limitations and strengths of any aid, and thus provide for more rapid advancement of research into sensory aids. However, care should be taken that the results of any tests not be influenced by the language level of the user, and that proper training with a device be considered before any final conclusions from testing were drawn.

CONCLUSIONS

The overall conclusions of Working Group 5 are summarized below:

1. The design of hearing aids could derive further benefit from already existing technological advances in miniaturization and computerization (this includes the use of computerized fitting procedures). More work is needed in deciding the optimal characteristics of hearing aids with respect to compression, number of bands, bandwidth, etc. Group hearing aids would likewise benefit from increased

research and the available technology, if more persons were working on the problem.

2. There appear to be benefits from the use of both single-channel and multichannel speech aids for profoundly deaf individuals. Again, more work is necessary to determine optimal training procedures and stimulus parameters for any specific aid. However, a major problem that must first be overcome is the lack of a good transducer that has a frequency response tailored to the skin's capabilities, no acoustic emission, low power consumption, and long battery life. The high cost but low consumer demand for such a device suggest that the quickest and most efficient way to promote development of a wearable tactile aid might be to enlist the support of the federal government for the project, in a fashion similar to that of the "orphan drugs" policy presently implemented.
3. Visual aids to speech reception and production are potentially quite useful in the school setting, particularly to serve as a means for students to practice already acquired skills. A whole range of such aids, from very simple to very complex, could be utilized for a variety of speech tasks and language levels. Such a device would permit the teacher of the deaf to spend more time teaching new skills and less time on rote repetition and practice. In addition, a visual aid could be designed to provide extra motivation for a child if it could be used as a game.
4. Progress in the area of development of sensory aids for the deaf would be considerably facilitated by the development of a comprehensive and general-purpose test of auditory abilities, which would include both psychophysical measures for simple and complex detection and discrimination and more involved speech testing. Such a test battery would allow the evaluation and comparison of any sensory aid and permit direct assessment of the relative potential of any two aids, regardless of their degree of design similarity. This evaluation is presently not possible because of major differences across laboratories in training and testing techniques. Every effort should be made to construct tests that are language-independent and to allow selection of subtests without impairment of validity.

Sensory aids represent an enormous potential benefit to the hearing-impaired or deaf individual. Research should continue into improvements in auditory, tactile, and visual aids, and an attempt should be made to provide a variety of devices to the teacher of the deaf for use in the classroom as rapidly as possible. Even the amount of improvement provided by currently available devices should benefit some number of deaf individuals, while the potential for future benefits is truly impressive.

REFERENCES

Saunders, F. A., Hill, W. A., & Simpson, C. A. (1980). Speech perception via the tactile mode: Progress report. In H. Leavitt, J. M. Pickett, & R. A. Honde (Eds.), *Sensory aids for the hearing-impaired* (pp. 278–281). New York: IEEE Press.

CONFERENCE SUMMARY

IRA J. HIRSH

Central Institute for the Deaf, St. Louis, MO

After listening to these two days of lectures and discussions, and after 30 years of being frustrated by teachers' questions concerning why the Research Department did not do some things relevant to the teaching, I am led to adopt, for this summary, the position of a teacher. We teachers have told you through words and demonstrations what we do, and we have also told you about our frustrations about many aspects of teaching speech to the deaf. How can we now apply to that enterprise what you have told us in these 2 days?

The first question concerns what it is that we are changing, or working on, or teaching. MacNeilage outlined three possible compartments: (a) the realm of ideas or intentions, (b) a system of lexicon with syntactic frames, and (c) some motor executors. Mostly, we behave as if we were working on the motor-executor stage, manipulating directly or trying to instruct movement with (often inadequate) words and gestures. In a sense we must work at that level because we do not believe that the auto-organization constituting the output of the linguistic operators (lexicon and frames) learned early normal hearing can function in the case of auditory deprivation. Some of that organization ought to be available from those sensory feedback channels that are still functioning, but the principal one is not available and so we, the teachers, become the child's ears and provide some degraded signals from this second-hand auditory monitoring. In the development of the normally hearing, that auditory monitoring begins very early, even before the stages presented in Calvert's diagram, where the consequences of vocal and lingual play are available to several sensory channels so that, to borrow a concept from Held and Hein 1963, auditory and somasthetic fields are laid upon each other.

Your discussions have put great emphasis on the basic-unit size, and we have similar worries in our speech teaching. There appears to be a hierarchy from phonemic feature, to phoneme, to syllable, to word, and to phrase and sentence. From the demonstrations you note that the teacher's ear serves to pick up a feature that appears to be inappropriate, she works on a particular phone, employs that phone in modeling a word, and then repeats the accomplished motor execution in the context of a sentence. It is a bit discouraging to find that the features, so important to Stevens, are best described by him in either acoustical terms or perceptual ones. It would be more helpful to us if the features were definable in dimensions of production. Also his encouraging remarks about the redundancy in the system and the fact that phonemic identification did not require the identification of all features all of the time, was also a bit disappointing from the point of view of practice. We might have preferred if he could tell which features could most often be dropped, and perhaps an ordered list from the most important to the least important.

We will want to focus on the sentence, but in doing so we want to fix some smaller unit along the way, and we hope that we are working on the right units. Lest you think that this is a new issue, I would refer you back to a book, mentioned by Boothroyd, titled *The Teaching of Speech*, by G. Sibley Haycock, published in 1933, the year that Silverman entered this profession. Thus it is a book that shares with him a 50th anniversary. On page 34 we find the statement: "It is impossible to listen to ordinary speech without realizing that, in its phonetic aspect, the true speech-unit is the simple sentence."

The problem of choice of unit is joined also by the need for extending particular training at one unit level to speech productions at higher levels—eventually the sentence. As Haycock put it 50 years ago "There are two methods of teaching speech to deaf children, namely the analytic method and the synthetic method. Briefly described, the analytic method presents individual speech sounds to the deaf child for imitation and when these individual sounds can be rendered satisfactorily they are combined into words. But no encouragement is given to the children to reproduce words before the constituent elements can be correctly articulated. The synthetic method, on the other hand, encourages from the beginning, the imitation by the pupil of whole words before the constituent elements have been taught individually . . ." This dichotomy is still with us and still part of our theoretical discussions. Boothroyd has emphasized the notion of drill, particularly as applied to the smaller units in order to insure a maximum of habitual or automatic control over those units. Even beyond the "synthetic method" mentioned by Haycock, there has been in recent decades extension to sentences as a whole, not only as an example of the appropriate large linguistic unit, but also as an application of letting the child know and experience the consequences of successful social communication, as in the auditory-global approach of Simmons-Martin. (It was interesting to observe Moog employing this approach with our visiting lecturers who were learning her new, artificial language.)

Perhaps more by way of commentary than summary, I would like to emphasize the vast territory that lies between the two extremes of automaticity and cognitive activities. Those notions appear to be the poles of extremes of a continuum and much of our discussion pertains to which aspects of the speech or language system can be taught and controlled at the automatic level, and which require decisions. To set up those two alternatives as a dichotomy may prevent us from discovering, within the large territory between them, sets of perceptual rules and similar patterns of motor organization, both of which can be characterized as hierarchies and at the same time represent learned or acquired properties of the perceiving and the producing systems.

We have been fascinated to learn about different kinds of words, particularly from Stephanie Shattuck-Hufnagel whose observations on speech errors lead us to assign different roles to different kinds of words with respect to stages of the speech-production hierarchy. It was some 20 years ago when Martin Braine described the language systems of very young children as consisting of two word classes—the open class and the pivot class. These were to be the precursors of the later dichotomy between function words and content words, which dichotomy is now importantly returned as the representations respectively of frames and of content. The speech-error support for these classes is fascinating and begs to be extended. Before I would assign the speech-error information to a status within linguistic theory, I would want to know whether a more parsimonious interpretation might be based on the fact that, first, the function words are probably the smallest morphs, and second, having been members of the pivot class (thus small in number and often used) they have been the most practiced. Briefly put, have these function words become automatic, because of the role they play in the language or because they are the most practiced? After all this is settled, then you must also tell us whether because the normal feedback channels were not functioning for the deaf child, our methods of speech instruction should be different for those two general word classes.

Finally one is led to ask what new ideas have come forth in this conference to shed light on the speech-teaching process. Surely the pulling together of a theory of word classes or different parts of the production system, with systematized observations on speech output is an important new area of knowledge. I had thought to say also that things aren't very different from the Haycock-Silverman days of 50 years ago, but they are. On the engineering side, display instruments and hearing aids themselves are better. The potential recoding of some aspects of the message for a substitute sensory modality also is closer to being a reality.

On the developmental side, we still need to know what aspects of the planning system are different because the child does not hear, and what aspects are shared with children who do hear. The enriched experience of a young child in providing the nutrients for his ideas or intentions must not be neglected in favor of the vehicle itself, lest our students develop good speech-production systems without having much to talk about.

Almost all of our lecturers have emphasized the fact that the picture is not completely clear, that more knowledge is necessary. I would only add to that, that I hope that over the next couple of decades empirical contributions will become as numerous as theoretical ones.

REFERENCES

BRAINE, M. (1963) The ontogency of English phrase structure. *Language*. 39, 1–13.

HAYCOCK, G. S., (1933). *The teaching of speech*. Washington: Vofta Bureau.

HELD, R., & HEIN, A. V. (1963). Movement-produced stimulation in the development of visually-guided behavior. *Journal of Comparative and Physiological Psychology*, *56*, 872–876.

SIMMONS-MARTIN, A. (1977). Natural language and auditory input. In F. H. Bess (Ed.), *Childhood deafness* (pp. 305–312). New York: Grune & Stratton.