Speech-perception aids for hearing-impaired people: Current status and needed research

Working Group on Communication Aids for the Hearing-Impaired\(^a\)

Committee on Hearing, Bioacoustics, and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council, Washington, DC 20418

(Received 14 July 1990; revised 17 January 1991; accepted 26 March 1991)

Both the overall aging of the population and its exposure to higher noise levels have increased the tendency to hearing loss and the importance of improved hearing aids for speech perception. This article reviews improvements in conventional electroacoustic hearing aids, as well as recently developed alternative classes of speech-perception aids, including surgically implanted cochlear stimulators, and vibrotactile, electrotactance and optical stimulating devices. It is concluded that the most effective aid for the vast majority of hearing-impaired persons is, and will remain for the immediate future, the electroacoustic hearing aid. In those cases for which no benefit is demonstrated for the electroacoustic aid, generally meaning persons with profound hearing loss, either the cochlear implant or a tactile aid may provide some improvement in the understanding of speech. In rare cases, some speech understanding in the absence of lip reading is achieved by patients with cochlear implants, for unexplained reasons. This and other pressing questions about speech processing need to be addressed by the research community if more effective aids are to be developed for the use of the 36.5 million hearing-impaired persons expected in the U.S. by the year 2050.

PACS numbers: 43.10.Ln, 43.66.Ts, 43.64.Me, 43.66.Wv

CONTENTS

Introduction .......................................................... 638
A. Magnitude of the problem ..................................... 638
B. The working group's approach ............................. 640
I. Conventional electroacoustic hearing aids ............... 641
   A. Introduction to conventional electroacoustic hearing aids ........... 641
      1. Major types currently available ..................... 641
      2. Frequency of use of major types .................. 642
      3. A recent development: Implantable electroacoustic aids ........ 643
   B. Dealing with reduced dynamic range and related audiological considerations ......... 643
      1. Frequency shaping .................................. 644
      2. Improving intelligibility ............................ 644
      3. Protecting against excessive amplification ...... 645
      4. Compression amplification .......................... 645
      5. Summary ............................................. 646
   C. New developments in conventional electroacoustic hearing aids .............. 646
      1. Digital hearing aids .................................. 647
      2. Speech-analyzing hearing aids ..................... 647
      3. Signal processing to reduce noise and reverberation .................. 649
   D. Who should use conventional aids? ..................... 652

II. Cochlear implants ............................................... 654
   A. Basic psychophysics of the electrical stimulation of hearing ................. 655
      1. Pitch .................................................. 655
      2. Loudness ............................................. 656
      3. Percepts with multiple electrodes ................ 657
   B. Risks involved in cochlear implants ...................... 657
   C. Patient selection criteria ................................ 657
   D. Acceptance by patients .................................. 658
   E. Patient results and performance: Postlingually deafened adults ............... 658
      1. Environmental sounds ............................... 658
      2. Discrimination of speech characteristics ....... 658
      3. Prosody .............................................. 658
      4. Consonant and vowel confusions ................. 660
      5. Word identification ................................ 660
      6. Sentence recognition .............................. 661
      7. Tracking ............................................. 661
   F. Patient results and performance: Prelingually deafened adults ................. 662
      1. Environmental sounds ............................... 663
      2. Stress/prosody ...................................... 663
      3. Speech discrimination .............................. 664
      4. Vowel/consonant recognition .................... 664
      5. Word recognition ................................... 664

\(^a\) Please see the end of this report for an important notice and the members of the Working Group on Communication Aids for the Hearing-Impaired and of the Committee on Hearing, Bioacoustics, and Biomechanics (CHABA).
INTRODUCTION

In the past 25 years, conventional acoustic hearing aids have been improved in a variety of ways. During the same period, two alternative classes of speech-perception aids have also been developed: (1) surgically implanted cochlear stimulators (cochlear prostheses) and (2) sensory-substitution systems, including vibrotactile and electrocutaneous stimulating devices. The purpose of this report is to describe the effectiveness of each of these three approaches to aiding speech reception by deaf or hearing-impaired people. In this introductory section, we begin with a few statistics that indicate the demographic magnitude of the problem. Then we discuss the specific questions the working group has attempted to answer and identify the target audience. Finally, we consider the problem of providing a report that accurately reflects the status of a topic about which scientific and clinical communities have not yet reached agreement.

Sections I–III are devoted to the three major classes of sensory aids: conventional hearing aids, cochlear implants, and sensory substitution. In Sec. IV, we present conclusions and recommendations of the working group when the evidence supports them and discuss the scientific arguments in areas in which the evidence is inconclusive.

The working group is aware that a large number of deaf people are committed to manual communication, and many of them have decided not to use oral–aural speech communication. This report does not address the very complex scientific and philosophical questions surrounding the choice, by a deaf person or by the parent of a deaf child, of manual, oral–aural, or mixed (“total communication”) modes of communication. It is explicitly directed toward those deaf and hearing-impaired people who choose to use speech communication.

Hearing loss is most broadly categorized into impairments that are either conductive or sensorineural in origin. Purely conductive losses are those resulting from dysfunction of the ear canal or middle-ear structures, so that less acoustic energy reaches the auditory receptors in the cochlea. Sensorineural loss involves reduced sensitivity or resolving power of the neural receptor mechanisms themselves. Today, most conductive losses are successfully treated through medical intervention and do not necessitate aids to speech understanding. Sensorineural hearing loss, however, is generally not amenable to medical intervention. Modern aids to speech perception, therefore, are designed primarily for individuals with sensorineural loss. Occasional mention is made, however, of aids that are particularly useful for those unusual individuals with conductive loss for whom surgery may be inappropriate.

A. Magnitude of the problem

Aids to speech perception are the primary form of treatment, therapy, or remediation (depending on the discipline of the clinician or practitioner) for enhancing the speech communication abilities of hearing-impaired or deaf people. While seldom a life-threatening condition, hearing impairment is likely to restrict social interaction and educational and career opportunities and thus significantly degrade quality of life. On the basis of the predicted numbers of affected people in the U.S. population, hearing impairment will be one of the major health problems to be faced in the remainder of this century. Table I illustrates the magnitude of the problem, as estimated by Fein (1983) on the basis of data collected by the National Center for Health Statistics and the U.S. Bureau of the Census. The predicted size of the hearing-impaired population—36 million hearing-impaired people by the mid-21st century—is the result of the increasing size of the total U.S. population as well as the aging of that population (it should be noted that these data are based on self reports).

Table II shows 1988 data on the numbers of hearing-impaired people in various age ranges (Adams and Hardy, 1988).
TABLE I. Estimated proportion of the U.S. population with hearing impairment (in millions), 1983.

<table>
<thead>
<tr>
<th>Year</th>
<th>U.S. population</th>
<th>Number with hearing impairments</th>
<th>Percentage hearing-impaired</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>180.6</td>
<td>13.0</td>
<td>7.2%</td>
</tr>
<tr>
<td>1970</td>
<td>205.0</td>
<td>15.2</td>
<td>7.4%</td>
</tr>
<tr>
<td>1980</td>
<td>227.7</td>
<td>18.1</td>
<td>7.9%</td>
</tr>
<tr>
<td>1990</td>
<td>249.7</td>
<td>21.1</td>
<td>8.4%</td>
</tr>
<tr>
<td>2000</td>
<td>268.0</td>
<td>24.0</td>
<td>9.0%</td>
</tr>
<tr>
<td>2025</td>
<td>301.0</td>
<td>32.7</td>
<td>10.9%</td>
</tr>
<tr>
<td>2050</td>
<td>308.9</td>
<td>36.5</td>
<td>11.8%</td>
</tr>
</tbody>
</table>

TABLE II. Age distribution and prevalence rates of hearing-impaired people in the United States, 1988. (Note: Total hearing-impaired population figures in Table I (1983) disagree with those shown in Table II for 1988, because of discrepancies used in projections for the former. The resulting errors do not exceed 2%–3% overall.)

<table>
<thead>
<tr>
<th>Age</th>
<th>Hearing-impaired people (millions)</th>
<th>Prevalence rate (%)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under 18</td>
<td>1.1</td>
<td>1.7</td>
</tr>
<tr>
<td>18–44</td>
<td>5.0</td>
<td>4.9</td>
</tr>
<tr>
<td>45–64</td>
<td>6.7</td>
<td>14.8</td>
</tr>
<tr>
<td>65–74</td>
<td>4.8</td>
<td>27.4</td>
</tr>
<tr>
<td>75 and over</td>
<td>4.2</td>
<td>38.1</td>
</tr>
<tr>
<td>All ages</td>
<td>21.9</td>
<td>9.1</td>
</tr>
</tbody>
</table>

¹ Ratio of number of cases divided by number of people in the total population in that age group. Source: National Center for Health Statistics data cited in Adams and Hardy (1989).

1989). Over half of the hearing-impaired people in the United States were working-age adults, and roughly one-third were over 65. These data, based on survey interviews, do not indicate the levels of severity of the hearing impairments. In 1985 about 8% of the U.S. population (about 19 million people) had a hearing impairment; roughly 3% (about 7 213 000 people) indicated that they could understand normal conversation only “with difficulty;” and 0.3% (about 481 000 people) said they could not hear normal conversation at all.

The primary treatment for most hearing-impaired people, namely those who have sensorineural hearing loss (“nerve deafness”) and cannot or do not choose to learn to live with it, has been some type of aid for speech perception. Most commonly these have been conventional electroacoustic hearing aids. However, modern technology has not neglected this problem, and a variety of new types of aids has been developed. Some of these have been commercially available for 10 to 15 years, while others have only recently been shown to be effective in laboratory studies and have just begun to be used in significant numbers. Some of these new devices have been the subject of a good deal of discussion in the general media, including, for example, reports of President Reagan’s in-the-canal acoustic hearing aids.

The primary treatment for most hearing-impaired people, namely those who have sensorineural hearing loss (“nerve deafness”) and cannot or do not choose to learn to live with it, has been some type of aid for speech perception. Most commonly these have been conventional electroacoustic hearing aids. However, modern technology has not neglected this problem, and a variety of new types of aids has been developed. Some of these have been commercially available for 10 to 15 years, while others have only recently been shown to be effective in laboratory studies and have just begun to be used in significant numbers. Some of these new devices have been the subject of a good deal of discussion in the general media, including, for example, reports of President Reagan’s in-the-canal acoustic hearing aids.

For the severely impaired and deaf population, there is no device that has caught the imagination of the public more than the cochlear implant or aural prosthesis. This device has been developed on the basis of animal work by Galambos and Rupert (1959), Simmons (1964), and others and applied to human patients by Simmons (1966), House (1974), Michelson (1971), and by many others in recent years. Newspaper headlines such as “New Ear Implant Liberates the Deaf from World of Silence” (Indianapolis Star, 30 January 1985) have announced this new device to the public. When computerlike memory banks are available in $30 wristwatches, the idea that a replacement for the organ of hearing might be devised and surgically implanted seems not at all farfetched, at least to some people. However, as is often the case, enthusiastic reports in the general media have failed to reflect a wider range of opinions in the clinical and scientific literature. For example:

Results have been very encouraging, with all but two of these patients (out of 15 implanted and tested with a UCSF/Storz multichannel cochlear implant) obtaining some degree of open-set auditory-only speech recognition. Most patients have demonstrated improvement over time without extensive rehabilitative intervention and all patients have attained an enhancement in lip-reading ability, as measured with a tracking procedure. According to a self-rated performance scale, all patients have experienced improvement in general communicative function since receiving the implant (Schindler and Kessler, 1987).

Although it is true that several implant recipients have reported achieving open-set speech discrimination, it is equally true that the auditory skills of some implant recipients do not exceed the skill level they achieved with a hearing aid. In fact, most persons who are involved in the clinical investigation of cochlear implants likely would be quick to point out that superior performance is probably the exception rather than the rule (Windmill et al., 1987).

While the attention of the medical community, as well as the general public, was understandably captured by the concept of a “bionic” cochlear implant, research has also continued on other approaches to the treatment of hearing impairment. In addition to a wide range of technological improvements in traditional acoustic hearing aids, significant advances have also been made in sensory-substitution aids. Rather than eliciting responses in the cochlea and VIIIth cranial nerves with intense sound (as in the hearing aid) or by direct electrical stimulation (as in the cochlear implant), this approach simply bypasses the auditory sections of the nervous system and presents speech information via an unimpaired sensory channel (vision or touch). That speech can be received through the tactile channel is demonstrated by the abilities of some deaf-blind individuals who communicate by tactile speechreading (the TADOMA method). These people are able to understand speech at slow conversational rates by placing a hand on the face of the talker and monitoring the mechanical actions of the face associated with speech production. An equivalent demon-
B. The working group's approach

This is a report on the status, in the late 1980s, of speech-perception aids for hearing-impaired or deaf people. It was prepared by a group of ten clinicians and scientists selected by the National Research Council’s Committee on Hearing, Bioacoustics, and Biomechanics (CHABA). The working group includes individuals who are widely experienced in research on, and in clinical applications of, each of the three classes of aid. The working group's charge was to “review and summarize the levels of speech perception that can be expected using various types of aid, by persons from infancy to old age, with various degrees of hearing loss, and who are either pre-lingually or post-lingually hearing impaired.”

As the body of this report demonstrates, some portions of the charge have been met more successfully than others. The charge can be thought of as a multidimensional experimental design, with the number of cells equal to the product of the number of categories in each dimension. There are three to four conceptually distinguishable levels of hearing impairment, of age, and of aid type, and at least two important categories of the age of onset of impairment, all deserving separate consideration. Thus, to respond fully to the charge, we would need to know the level of speech perception predicted for approximately 50 + combinations of listener characteristics and classes of aid. The difficulty in filling this weighty order is not, however, the sheer amount of information needed; it is the absence of the rigorously collected data required to address many of the cells.

Beyond simply reviewing the characteristics and performance of hearing aids, in this report we have also attempted to provide an overview of new approaches to solving the general problem of speech perception by hearing-impaired people, many of which are not yet available commercially. Certain unresolved research problems are also identified, which represent major obstacles to the development of more successful aids. Among these latter, two vital issues that are not reviewed in depth here are the related questions of (a) exactly how radical can transformations of the normal acoustic waveforms of speech be, before human beings cannot learn to recognize them as speech and (b) at what age must experience begin with a particular type of transformation of speech, delivered via a particular sensory channel, for that form of the speech code to be efficiently learned. These questions clearly involve issues of neural and developmental plasticity, “critical periods,” and many other topics beyond the scope of this report, despite their considerable relevance to the long-term solution of the problem of speech perception by hearing-impaired and deaf people.

Although very little hard evidence is available on hearing-impaired people, experiments on normal listeners suggest that the learning of new sensory codes may require lengths of time similar to those required to learn to recognize the sounds of continuous discourse in a foreign language (Watson, 1980). This prolonged time course of perceptual learning has two potential consequences for the problems of concern here. First, it suggests that hearing-impaired people may need to devote very long periods of training to their use of a new hearing aid before realizing its full potential. This required training time is probably longer as the aid’s output is perceptually less similar to the speech the person once knew (possibly particularly difficult in the case of sensory-substitution aids or cochlear implants). Second, and indirectly a more positive consequence, the long-time course of perceptual learning suggests that the human capacity to learn new sensory codes may be considerably greater than has been demonstrated in the performance of many experimental systems. Only rarely have experimental aids been used for the several months or possibly a year or two of training that may be required to reach the performance levels of which people are capable.

Despite the working group’s recognition of the importance of prolonged training in learning to use alternative codes, many of the studies reviewed here are presented with little mention of the effects of training or even of the times employed. The reason, as noted above, is that very few investigations have systematically studied the results of prolonged periods of training. In some cases, performance measures have been collected after several years of regular use of an aid, but often these same subjects had not been studied early in this period, or at periodic times through it, to document the actual course of training.

While the critical importance of the time course of perceptual learning, of the degree of developmental plasticity, and of the age and duration of possible critical periods for learning speech codes are acknowledged, both hard data and a generally accepted theory of these matters are lacking.

The heart of the report consists of three chapters devoted to the three major classes of speech-perception aids (acoustic, cochlear-implant, and sensory-substitution aids); each of those chapters was prepared by a subgroup experienced with that particular type of aid. This format guaranteed that the report would include the evidence considered most relevant to the functioning of each class of aid by those experts most familiar with that class. In general, it appeared easier to agree about the facts, as documented in the research literature, than about the most valid interpretation of those facts. One example may help the reader to appreciate the nature of the difficulty. Candidates for cochlear implants
are, in some programs, first fitted with a vibrotactile device. Unusually good performance with the sensory-substitution aid is seen by at least one implant team as an indication that a patient will also do well with the implant and therefore that the implant procedure should be undertaken. Others find such decisions difficult to understand, in the absence of a larger volume of evidence that the implant, on the average, can be predicted to provide better speech perception than a vibrotactile device.

This report is not intended as a substitute for the thorough reading of the pertinent literature, which is the responsibility of any clinician or scientist who wishes to achieve expertise on speech-perception aids. It does attempt, however, to articulate both the views of a representative sample of those who have studied these subjects actively for the past 20 to 30 years, as well as something of the variance in opinions about optimal treatment held by contemporary clinicians and scientists (despite their general agreement about the results of relevant research).

It is hoped that the report will provide some useful assistance to several groups who may lack either the time or the appropriate background to study carefully the literature on speech-perception aids for hearing-impaired and deaf people. One important group is deaf and hearing-impaired people with the technical background necessary to read it. Although we wish that this report could promise more help for those people than it does, a careful reading should assist them in making difficult choices. No one is likely to be a more dedicated advocate for a patient than the patient. A second target group is physicians in general practice and other clinicians who may not be actively involved in selecting or providing hearing-impaired people with speech-perception aids but who nevertheless have to advise hearing-impaired patients. Third, it is hoped that this report may be useful to the staff of, and advisers to, the federal agencies responsible for the support of research on the next generation of communication aids. Finally, hearing-aid manufacturers may find this report useful in their consideration of future product lines.

I. CONVENTIONAL ELECTROACOUSTIC HEARING AIDS

In this section, we first describe the characteristics of electroacoustic hearing aids that are currently available.1 We then discuss some of the major results of research efforts to improve electroacoustic aids, many of which have not yet been incorporated into commercially available aids. Finally, we discuss the characteristics of those hearing-impaired individuals for whom traditional electroacoustic aids are the most appropriate aid to speech perception.

Acoustic amplification is the method most commonly used to enhance the recognizability of speech and other information-bearing signals so as to improve communication for hearing-impaired people. The instrument most commonly used for this purpose is the conventional hearing aid. A conventional hearing aid has three key characteristics: (1) it is a personally fitted, wearable device; (2) it can be freely put on and taken off by the user, although intended for regular full-time use; and (3) the amplified signal is delivered acoustically to the external ear canal. Other acoustic amplification devices that do not meet these constraints are known more generally as "assistive listening devices;" this class of devices is not discussed here.

Conventional hearing aids typically consist of a microphone, electronic filter, controls for adjusting the amplification (or gain) and overall shape of the frequency response (e.g., a bass or treble boost), circuits for limiting the amplified signals to a comfortable or safe level, an earphone (commonly referred to as a "receiver" in the United States), a battery that serves as the power source, and various acoustic components, such as flexible tubing and an earmold, for coupling the output of the receiver to the external ear canal.

A. Introduction to conventional electroacoustic hearing aids

1. Major types currently available

In this section, we describe the major types of conventional electroacoustic hearing aids currently available. Discussions of the frequency of use of each of the types and their relative merits appear in later sections. Because of the wide variety of types of conventional aid and because merits or drawbacks often characterize more than one type, this organization seemed most efficient.

A common method of classifying conventional hearing aids is according to their relative size and how they are worn. The largest wearable hearing aid is the body aid, in which the electronic components are housed in a body-worn case, the amplified signals being delivered by wire to a receiver mounted in the ear. Most body aids are high-powered instruments.

A smaller and less conspicuous instrument is the eyeglass hearing aid. The electronic components of this device fit into one of the bows of a pair of eyeglasses. The microphone is mounted at the front of the bow next to the lens, the receiver is at the rear of the bow and its output is via a nozzle and attached tubing to an earmold.

A relatively small and very popular instrument is the behind-the-ear (BTE) hearing aid. The electronic components of a BTE aid are housed in a small elliptical case that, as the name implies, fits behind the ear. The acoustic signals generated by the receiver are delivered to the ear canal by means of a flexible acoustic tube terminating in an earmold.

An even smaller instrument is the in-the-ear (ITE) hearing aid. All the components of this aid are contained in a small plastic case that is molded to fit into the user's ear. It typically occupies the outer portion of the external ear canal and the concha (the innermost part of the external ear).

The smallest hearing aid of all is the in-the-canal (ITC) aid, which fits entirely in the ear canal. It is the least conspicuous of conventional hearing aids, the outer face of the unit being just visible at the opening to the ear canal. Unfortunately, the available power decreases with size, due primarily to the very small batteries required for these instruments and the limited electrical power they supply. As a result, ITC instruments are generally restricted in application to those patients with only mild amounts of hearing loss.

Hearing aids are also classified according to the number
and arrangement of microphones and receivers. The simplest and most common is the monaural hearing aid consisting of one microphone, one receiver, and associated electronic circuitry. Normally both the microphone and the receiver are mounted on the same side of the head. A variation of this arrangement is the CROS hearing aid, in which the microphone and receiver are on opposite sides of the head. Another is the BICROS, in which there are two microphones, one at each ear, and only one receiver. The BICROS systems are typically used on bilaterally hearing-impaired persons who have only one ear that can benefit from amplification. Another aid of this type is the POWERCROS which, like the conventional CROS aid, also uses one microphone and one receiver mounted on opposite sides of the head. For the POWERCROS instrument, greater separation of the microphone and receiver is needed to reduce acoustic feedback. (Acoustic feedback is a common problem in high-powered hearing aids: when the amplified sounds from a receiver are reintroduced directly into the microphone—i.e., are "fed back"—the level of the output of the receiver may rise quickly to the maximum of which the system is capable, resulting in an intense whistling or screaming sound that is very objectionable to the user.) Hearing aids with one microphone and two receivers are known as Y-cord aids and are sometimes used with persons having a bilateral hearing loss. The more common approach for bilateral amplification is the binaural hearing aid, which essentially consists of two monaural aids, one for each ear. A modification of the above is the CRISCROS hearing aid, in which the microphones and the receivers are connected contralaterally, also in order to reduce acoustic feedback.

All of the hearing aids described thus far rely on air conduction for transmitting sound from the receiver to the eardrum. In a bone-conduction hearing aid, sound is transmitted to the cochlea through the bones of the skull, usually from a bone vibrator placed on or tightly affixed to the mastoid bone behind one ear. Bone-conduction hearing aids are not widely used, although at one time they were extremely common, for reasons discussed in a later section.

2. Frequency of use of major types

The conventional hearing aid is the most widely used sensory aid for the hearing impaired (Levitt et al., 1980). Approximately 1.4 million hearing aids were sold in the United States last year (Cranmer, 1990). Of these 78% were either ITE or ITC hearing aids, 20% were BTE units, and roughly 3% were miscellaneous other types of instruments, such as eyeglass or body-worn hearing aids. This breakdown represents a significant change over previous years, when BTE hearing aids dominated the market. Eyeglass hearing aids, which at one time were quite popular in Europe, are no longer widely used and represent only a very small fraction of the American market. Similarly, bone-conduction hearing aids, which at one time were quite popular, are used only rarely.

The reasons for the growth or decline in usage of different kinds of hearing aids involve a complex mix of audiological, engineering, and marketing factors. Body aids, developed as the first electronic-acoustic aids, are used primarily by persons requiring very high-powered acoustic output and for whom cosmetic factors are of secondary importance. Many profoundly hearing-impaired children at schools for the deaf use body aids, as do some hearing-impaired adults requiring high-powered acoustic amplification. A common complaint with body aids is the inconvenience of the hard-wired connection between the body-worn unit and the earpiece. This electrical connection is a common source of trouble in the form of broken or intermittent connections. These factors, in addition to the cosmetic issue, help account for the low popularity of body aids.

The more recently developed BTE hearing aid represents a substantial advance over the body aid. It not only provides convenience and cosmetic appeal, but it also provides a more natural signal input, through its use of ear-level microphones. As a consequence, it soon replaced the body aid as the instrument of choice for the large majority of hearing-aid users. The BTE hearing aid has been replaced by ITE or ITC aids as the instrument of choice. The trend is not as marked in Europe, where ITE hearing aids are increasing in popularity but have not yet surpassed BTE hearing aids in new fittings. The great popularity of ITE and ITC hearing aids appears to be due in large measure to their cosmetic appeal. Although audiologically they have some advantages (and some technical disadvantages) over their larger BTE counterparts, it is clear from marketing successes that small size (and hence low visibility) is their major attraction (in part, probably, because of the marketing emphasis on cosmetics). Because of their small size, ITE and ITC aids (particularly the latter) are limited in terms of their total acoustic power output and are appropriate only for persons with mild or moderate hearing losses.

In addition to identifying which hearing aids are in common use, it is also relevant to consider which hearing aids are no longer widely used and why. The bone-conduction hearing aid was in fairly common use prior to the development of BTE units. At that time, bone vibrators were not significantly larger than other electroacoustic receivers and, since all hearing aids of that period were body worn, there was little difference cosmetically between the two types of hearing aids. The bone-conduction hearing aid was designed specifically for persons with middle-ear impairments (conductive loss), which 25 to 35 years ago accounted for the majority of hearing-aid users. With the development and refinement of BTE hearing aids in the 1960s and 1970s, there were also significant advances in middle-ear surgery and, as a consequence, there are relatively few persons today who are candidates for acoustic amplification because of a conductive impairment. Moreover, the power output of a modern conventional hearing aid can be made high enough to overcome the attenuation caused by a conductive impairment.

In summary, the conventional hearing aid is by far the most widely used sensory aid for hearing-impaired people. Nevertheless, even though the number of hearing-aid users is high (over 2.5 million in the United States alone) the census data in the Introduction suggest that fewer than 1 out of 5 persons who could benefit from acoustic amplification wears a hearing aid. Of the many different kinds of hearing aids, currently the most popular are the small, relatively inconsi-
cuous units that fit either behind or in the ear. Some body-worn aids are still used for special applications. Although outside the scope of this report, there are other commercially produced acoustic amplification systems for hearing-impaired people. These include special-purpose amplifiers for use on the telephone, radio, and television as well as personal FM or infrared transmission systems for use in classrooms, auditoriums, theaters, and other settings in which background noise or reverberation is a problem.

3. A recent development: Implantable electroacoustic aids

A recent development, used to date in relatively few patients, is implantable hearing devices (not to be confused with cochlear implants). These devices typically do not involve an air-conduction path but rather substitute direct mechanical stimulation for the amplified acoustical signal. There is at least one major advantage of direct stimulation of the middle ear—reduction in acoustic feedback—and various techniques for direct mechanical stimulation have been developed over the years. One approach is to attach a small magnet to the eardrum or at some other point in the ossicular chain and to drive the magnet electromagnetically by means of an induction coil (Watanabe, 1965; Glorig et al., 1972).

Another approach is to insert a metal pin into the temporal bone and to drive the pin electromagnetically (or by an external vibrator), the vibrations being transmitted to the cochlea by bone conduction (Hough et al., 1986). An implantable hearing device using this general approach is available in which a small magnet is implanted in the mastoid. The magnet is sealed in silicone and housed in a titanium disk attached to an orthopedic screw. The screw-magnet assembly, referred to as the internal unit, can be implanted under local anesthetic in an outpatient setting.

The internal unit is driven by an induction coil mounted externally. The magnetic core of the coil is used to hold the external unit in place directly over the implanted magnet. Electrical signals applied to the coil cause the magnet to vibrate. These vibrations are transmitted to the cochlea by bone conduction. The implantable bone-conduction hearing device has been designed primarily for persons with conduction hearing impairment and for whom surgical correction is inappropriate. The device has recently received marketing approval from the Food and Drug Administration and as of 1 May 1987, more than 120 persons had been implanted with this device.

B. Dealing with reduced dynamic range and related audiological considerations

As is evident from the preceding section, cosmetic and marketing factors have had a major influence on hearing-aid development. As a consequence, engineering efforts have focused primarily on issues of microminiaturization and power conservation (for reducing battery size). Audiological considerations, despite their importance, have remained of secondary concern to the industry. This section provides a review of the basic audiological constraints affecting hearing-aid performance and the various attempts to deal with these constraints.

The reduced dynamic range of the impaired auditory system is the most obvious audiological problem and has received the greatest attention. The nature of the problem is illustrated in Fig. 1. The solid curves show typical speech spectra as measured at a distance of 1 m from the speaker’s lips. The dashed curves show hearing thresholds for three typical sensorineural impairments. (Sensorineural impairment has been chosen for this illustrative example since it is by far the most common impairment for which hearing aids are prescribed; in this form of hearing impairment, greater losses are typically found in the high frequencies.) Curve A is for a person with a mild-to-moderate, high-frequency hearing loss, curve B is for a severe loss, and curve C is for a profound hearing loss. Curve D shows the normal threshold of hearing. All curves show the detectability of a 1/3 octave band of noise as a function of band center frequency. The dashed line at the top of the diagram is the loudness discomfort level curve. Loudness discomfort levels are fairly similar for both normal hearing and sensorineurally hearing-impaired persons and, for purposes of simplicity, a single curve is shown.

The effective dynamic range of the auditory system is defined by the distance between the threshold curve and the loudness discomfort level (LDL) curve. The area encompassed by these two curves, i.e., between the threshold of audible sound and the ceiling of too-loud sound, is known as the residual hearing area. Note that the residual hearing area becomes progressively smaller with increasing hearing loss. Under extreme conditions, particularly in the high frequencies, the dynamic range from threshold to discomfort level may be only a few decibels. Some hearing-impaired persons with a drastically reduced dynamic range will typically expe-
experience an abnormally rapid growth of loudness with increasing stimulus level, known as loudness recruitment.

All of the hearing-loss curves in this example show more severe losses at higher frequencies. While this is very common, not all hearing-aid users have a high-frequency loss. Some have a loss that is relatively flat or that changes slowly with frequency. A still smaller number have U-shaped losses in which the hearing loss in the region of 1 to 4 kHz is greater than that at either low or high frequencies (Levitt et al., 1987a, 1987b).

Also shown in Fig. 1 are several typical average speech spectra (i.e., curves representing the average intensity level of the speech waveform at various frequencies). These include curves for male and female speakers both at a normal and a raised-volume level for a typical speaker-to-listener distance of 3 ft. Note that in all cases the average speech spectrum falls at a rate of about 9 dB/octave for 1/3 octave bands above about 500 Hz. Individual speech sounds, however, differ markedly from the average spectrum. Nasal consonants have most of their energy in the low frequencies whereas fricatives such as /s/ and /ʃ/ have more energy in the high frequencies. It is also important to note that there are large variations in speech level between speakers and also as a function of voice level. Temporal variations in level in normal conversational speech are estimated to cover a range of at least 30 dB (Dunn and White, 1940). Between- and within-speaker differences in level are thus large compared with the available dynamic range for most hearing-impaired listeners.

For a sound to be heard, of course, it is necessary that its level exceed the threshold of audibility. Acoustic amplification is thus of value in that it can raise many of the weaker sounds of speech into the residual hearing area. Unfortunately, there is a limit to the amount of amplification that can be provided. Because of the reduced dynamic range of the impaired ear, the amplification needed to make the weak sounds of speech audible will also make the intense sounds uncomfortably loud. Prolonged stimulation above the loudness discomfort level can also damage whatever residual hearing exists. Since for sensorineural hearing loss the dynamic range decreases with increasing hearing impairment, these problems increase in severity with increasing hearing loss. An exception to this rule is a purely conductive loss in which the threshold is raised approximately equally at all frequencies but the dynamic range is essentially normal.

1. Frequency shaping

One approach to the dynamic-range problem is to provide different amounts of amplification at different frequencies ("frequency shaping") so as to fit as much of the speech signal as possible into the residual hearing area without causing discomfort. A very simple but practical approach is to provide gain that is equal to roughly half the hearing loss at any given frequency (Lybarger, 1978). Another technique is to determine the frequency-gain characteristic that will make speech at a normal voice level comfortably loud at all frequencies (Pascoe, 1978; Skinner et al., 1982; Cox, 1985). Although many different formulae for the ideal or most practical frequency-gain characteristic have been proposed over the years (Wallenfels, 1967; Byrne and Tonisson, 1976; Berger et al., 1978; and McCandless and Lyregaard, 1983), there is no general agreement as to which is the best (Humes, 1986; Sullivan et al., 1988). Another long-standing point of view, which is less popular clinically, is that a single frequency-gain characteristic (e.g., + 6 dB/octave above 80 Hz) is near optimum for the large majority of hearing aid users and that individualized shaping of the frequency-gain characteristic should be undertaken only with hard-to-fit cases (Davis et al., 1947; Davis et al., 1986).

A prompt resolution of the above controversies seems unlikely, particularly since the evaluation of different hearing-aid-fitting procedures has been confounded with several practical constraints. Hearing aids are not ideal amplifiers; they introduce both distortion and noise, which in some cases can have consequences more serious than a poorly chosen frequency-gain characteristic. Another complicating factor is that the frequency-gain characteristic as measured on a standard coupler (device used to calibrate earphones) may be quite different from the true frequency-gain characteristic of the hearing aid when mounted on the ear.

2. Improving intelligibility

The process of making a sound audible does not necessarily make it intelligible. Although audibility of the many different components of speech is an important requirement for intelligibility, it is not the only requirement. It is also necessary that the listener be able to discriminate between different supra-threshold sounds, to be able to identify them, and, finally, to understand them in context. Factors affecting these processes include the etiology of the hearing impairment, perceptual distortions associated with the impairment, previous experience in listening to amplified or processed speech (including auditory training), the availability and use of additional visual cues, as in speechreading, and the listener’s knowledge of the language and ability to integrate contextual cues with those obtained auditorially and visually, and/or tactually (Levitt et al., 1980).

In view of the above problems, some audiologists have adopted an empirical approach to hearing-aid fitting in which a subset of hearing aids is first chosen, each of which has roughly the correct frequency-gain characteristic. A series of comparative tests is then performed (including tests of speech discrimination in noise) in order to select the best of the available subset of hearing aids. This approach is essentially a variation of the hearing-aid selection procedure developed by Carhart (1946). Serious problems with this comparative approach to hearing aid selection, however, have been reported (Walden et al., 1983). Many audiologists, as a result, have turned to prescriptive formula-based methods to select the most appropriate frequency-gain characteristics.

It is reasonable to ask how much improvement in speech recognition can be expected for a hearing-impaired person fitted with conventional amplification. Although this is not a trivial task, recent research has shown that the articulation index (French and Steinberg, 1947; Fletcher and Galt, 1950; ANSI, 1969) may be used to estimate the speech-recognition
Of average hearing-impaired listeners (Pavlovic, 1984; Kamm et al., 1985; Humes et al., 1986). To illustrate expected speech-recognition performance, let us consider the three hearing-impaired listeners shown previously in Fig. 1. Articulation Index calculations were made for quiet listening conditions for each of these listeners. For these calculations, we have assumed that a representative prescriptive procedure is used (gain equal to one-half the hearing loss) and that the maximum real-ear gain realized is 50 dB at any frequency. We have further assumed a normal conversational speech level (70 dB SPL) and speech materials consisting of nonsense syllables (Resnick et al., 1975). The unaided and aided speech-recognition performance estimated for patients A, B, and C under these assumptions are as follows:

<table>
<thead>
<tr>
<th>Patient</th>
<th>Unaided</th>
<th>Aided</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (mild-to-moderate)</td>
<td>79%</td>
<td>93%</td>
</tr>
<tr>
<td>B (severe)</td>
<td>41%</td>
<td>66%</td>
</tr>
<tr>
<td>C (profound)</td>
<td>0%</td>
<td>12%</td>
</tr>
</tbody>
</table>

Clearly, as the severity of loss progresses, the aided speech-recognition performance decreases. This is true even though the amplification of the hypothetical instrument is considerably greater for the profoundly impaired listeners. Estimates of aided performance for meaningful sentences rather than nonsense syllables are 100%, 91%, and 15% for patients A, B, and C, respectively. All estimates of speech-recognition performance for these patients, moreover, are for auditory input only; visual cues are not considered. It is important to note again that, although these articulation index predictions have been shown to be accurate for average listeners with a given hearing loss, some differences remain for individuals with identical loss.

### 3. Protecting against excessive amplification

In addition to frequency shaping, it is also necessary to provide some form of protection against excessive amplification. Amplification is excessive when it either results in sound levels that are uncomfortably loud for the patient or causes further loss of hearing due to the high sound levels. It is inevitable, however, that hearing aids for severely impaired listeners will generate high sound levels in the listener's ear and could jeopardize their remaining hearing sensitivity. General clinical practice is to assume that, although the high gain required by severe-to-profound losses for optimal speech recognition may continue to do some damage to the remaining sensory and neural structures, this risk of treatment-induced loss is preferable to the social isolation resulting from the nonsense of amplification. It is also assumed, from the fairly modest changes in threshold that typically occur in patients with severe or profound hearing loss following use of high-gain aids, that the remaining healthy sensory and neural elements in such an ear are fairly resistant to intense stimulation (Humes and Bess, 1981). More research, however, is needed on this issue.

Two of the most common forms of protection against excessive amplification are peak clipping (eliminating all portions of the output of the aid that exceed some specified level) and compression limiting (reducing amplification for higher-level sounds). The latter form of protection produces less distortion. Compression amplification, in principle, can also be used to maximize the proportion of the time-varying speech spectrum that can be placed within the residual hearing area. Clinical protocols for the specification of hearing-aid output levels not exceeding loudness discomfort have been developed in recent years (Hawkins, 1980; Cox, 1981), as have guidelines for limiting the output to safe levels that minimize the risk of further loss of hearing (Humes and Bess, 1981).

### 4. Compression amplification

Methods of compression amplification can be divided into three broad categories: (1) compression limiting, (2) long-term automatic gain control, and (3) syllabic compression. Each of the above methods of compression can be implemented either within a single frequency band (wideband compression) or in several contiguous frequency bands (multiband compression). Unless otherwise stated, wideband compression is assumed in the discussion that follows.

The most common form of compression in modern hearing aids is that of compression limiting. This type of compression is designed primarily for protection and operates only at relatively intense sound levels. The hearing aid behaves as a conventional amplifier for signals below the threshold of compression. When the compression threshold is exceeded, which occurs only for signals approaching the discomfort level, the gain of the amplifier is reduced substantially so that the output does not exceed a hazardous or uncomfortable sound level.

Experimental evaluations have shown that, as a protective device, compression limiting is superior to simple peak clipping. There is less distortion of the amplified speech signal and, correspondingly, speech intelligibility is reduced less by compression limiting than by peak clipping (Davis et al., 1947). Clinical evaluations of compression limiting, however, have not been favorable (Blevigad, 1974; Edgardh, 1952). In a critical review of this topic, Braida et al. (1979) attribute the negative results obtained in clinical evaluations to poor choice of compression characteristics, lack of individualized fitting, and confounding with other, uncontrolled electroacoustic variables.

Long-term automatic gain control (reducing amplification in response to high output levels), also known as automatic volume control (AVC), involves the use of a relatively long time constant; i.e., a time constant much greater than the duration of individual syllables in speech. This form of compression is designed to adjust for long-term variations in speech level so that more of the speech signal lies within the available range of residual hearing. Although the potential value of long-term automatic gain control was recognized some time ago, few studies have been undertaken to evaluate this form of compression. In addition, for reasons that are unclear, only a small proportion of conventional hearing aids have long-term automatic gain control.

In syllabic compression, the parameters of the amplification system are chosen so as to alter the relative intensities of individual speech sounds. Many of these intensity changes...
speech signal in each band makes an independent contribution to overall intelligibility. Stevens and Blumstein (1978), however, have shown that important speech information, such as place of articulation in stop consonants, is conveyed by short-term changes in the slope of the speech spectrum. Information of this type is lost in a multiband compression system in which the compression amplifiers in each frequency band operate independently. It is possible to maintain spectral contrasts in multiband compression systems by not allowing the compressors in separate bands to operate independently. This has been tried experimentally, but with only modest results (Braida et al., 1979; Levitt et al., 1987a; Busamante and Braida, 1987). The question of whether multiband amplitude compression is superior to a properly prescribed conventional single-channel hearing aid remains unresolved. The diversity of results obtained thus far indicate that many additional factors need to be considered.

In summary, of the three types of compression amplification that have been described, long-term automatic gain control and compression limiting have been found to be advantageous. Compression limiting is often used instead of peak clipping in many new hearing aids; automatic gain control is used in only a small number of commercially produced instruments. Syllabic compression is the most interesting and, at least in theory, potentially useful form of amplitude compression. More research is needed to determine unambiguously whether syllabic compression can improve speech understanding and, if so, the number of channels that should be used and the conditions under which it is most effective.

5. Summary

In summary, although many factors affect the design, selection, and use of conventional electroacoustic hearing aids, we have concentrated in this section on the central problem of dealing with the reduced dynamic range of the impaired ear. Frequency-shaping techniques are commonly used to fit as much of the speech signal as possible into the residual hearing area without causing discomfort. Protection against excessive amplification can be provided by means of peak clipping or compression amplification; however, both of these have failed to yield significant improvements in speech intelligibility. Research is ongoing to improve each of these techniques.

C. New developments in conventional electroacoustic hearing aids

Considerably more space is devoted in this section to research that has not yet had an effect on commercially available electroacoustic aids than is devoted, in Secs. II and III, to corresponding research on cochlear implants or sensory-substitution aids. This disproportionate attention to studies that may seem unduly futuristic is explained by three important differences between research on traditional aids and that on the two alternatives. First, because research on electroacoustic aids has been under way for 75–80 years, many more general lines of investigation have been developed and have yielded a significant literature, i.e., much
more scientific work has been done with them. Second, as noted in the final conclusions and elsewhere in this report, the conventional electroacoustic aid will almost certainly be the treatment of choice for most hearing-impaired people for the foreseeable future, and thus research on those devices is aimed at a far larger population of prospective users. Third—and probably most important—many of the special signal-processing strategies currently being investigated for use in electroacoustic aids will, once perfected, very likely be applicable to both of the other types of aid (cochlear implants or sensory-substitution aids).

1. Digital hearing aids

A very promising new development is that of the digital hearing aid. Digital techniques are now commonly used in instruments for hearing-aid measurement and calibration. As a result, powerful new measuring tools have been developed. These include instruments for the rapid and convenient measurement of sound transmission in the ear canal, measurement of in situ gain of hearing aids, and more accurate measurement of acoustic impedance. The extension of digital technology to the hearing aid itself appears to be primarily a matter of time and several experimental digital hearing aids have already been developed (Levitt, 1982, 1987; Nunley et al., 1983; Engebretson et al., 1986a, 1987; Cummins and Hecox, 1987).

Three types of digital hearing aids have been developed: (1) a quasidigital hearing aid in which digital circuitry is used to control analog amplifiers and filters; (2) a sampled-data system in which the audio signal is sampled at discrete intervals in time, the samples remaining in analog form during processing; and (3) the all-digital hearing aid in which the audio signal is sampled, converted to binary form, and then reconverted back to a continuous analog waveform after processing.

Current research is focused on developing application-specific chips that are both small enough and have sufficiently low power consumption to provide a viable alternative to the conventional BTE hearing aid. Several leading industrial research laboratories are actively engaged in the development of digital hearing aids. A practical compromise that appears to be feasible for the immediate future is that of a quasidigital hearing aid in which the audio signals remain in analog form but are controlled by digital circuitry. A chip for adaptive noise reduction in hearing aids has been developed using this approach (Graupe et al., 1987).

Digital hearing aids promise many advantages over conventional analog instruments (Levitt, 1987). In addition to providing greater accuracy and flexibility in the choice of electroacoustic parameters, they can be programmed by an external computer, thereby allowing for the introduction of new and more effective approaches to prescriptive fitting and evaluation of hearing aids (Popelka and Engebretson, 1983; Engebretson et al., 1986b). Moreover, powerful new signal-processing techniques can be used for reducing acoustic feedback (Preves et al., 1986) as well as enhancing speech intelligibility and reducing the effects of background noise and reverberation (Lim, 1983), as discussed shortly. It is unlikely that many of these features will be available in the first generation of commercially produced digital hearing aids, but once a practical, wearable digital amplification system has been developed, all of the above-mentioned features represent viable short-term goals.

2. Speech-analyzing hearing aids

A hearing aid incorporating some degree of speech-specific signal processing is, by definition, a speech-analyzing hearing aid. Conventional hearing aids that shape the speech spectrum so as to be comfortably loud at all frequencies are thus a very simple form of speech-analyzing hearing aids. Of particular interest in this section are experimental speech-analyzing hearing aids or, more generally, methods of speech analysis that could be usefully incorporated into such aids.

a. Frequency-lowering techniques. A broad class of (experimental) speech-analyzing hearing aids involves a process known as frequency lowering. One of the most common characteristics of a sensorineural hearing impairment is that hearing loss is greater at higher frequencies. In particular, hearing loss due to aging (presbycusis) is characteristic of this type. Translating the acoustic spectrum of speech downward (to lower frequencies) is thus an appealing prospect, since it would transfer speech energy in the high frequencies (where it is either not audible or poorly resolved) into the low-frequency region, where it is both audible and resolvable.

A number of frequency-lowering systems have been developed over the years using various strategies to accomplish the changes, using both selective and total-waveform lowering (see Braida et al., 1979, for a review of many different systems). Despite the number of frequency-lowering devices that have been invented or reinvented, a crucial question remains: How effective are these devices in improving the intelligibility of speech for hearing-impaired persons?

Experimental evaluations of frequency lowering for hearing-impaired people have yielded mixed results. Gunman and van Bergeijk (1959) reported that some improvement in speech reception could be obtained using a channel VOCODER for frequency lowering, but that learning to achieve a substantial level of improvement would be slow and time consuming. In an early evaluation of his transposer system, Johannson (1966) obtained improvements in the discrimination of fricatives and other phonemes by profoundly hearing-impaired children. In contrast, Ling (1969) reported a series of experiments in which no significant advantages over conventional amplification were obtained for the Johannson-type frequency transposer for either speech reception or speech training. Foust and Gengel (1973), however, showed significant improvements in speech discrimination ability (relative to conventional amplification) for individual subjects, but only after a fair amount of training. Reed et al. (1983, 1985b) have studied pitch-invariant frequency lowering, using nonuniform frequency compression of the short-term spectral envelope, but without any large-scale improvement in speech discrimination performance.

A special form of speech-feature transposition is that in which the fundamental frequency of the voice is recoded so
as to make it perceptually more salient. (Fundamental frequency, abbreviated $F_o$, is the lowest harmonic frequency in the speech waveform that typically determines vocal pitch, the vibration rate of the vocal folds.) In one version of such a system, the amplified speech is interrupted or amplitude modulated at a rate proportional to the voice fundamental frequency (Villechur and Kilion, 1976). These interruptions are perceptible even to profoundly hearing-impaired people.

Several experimental studies have shown significant improvements in lipreading ability with the addition of supplemental information about $F_o$ presented either auditorily (Grant et al., 1985) or tactually (Hanin et al., 1988). A signal-processing hearing aid providing pitch information has been developed by Rosen et al. (1987), in which the voice fundamental frequency is coded as a frequency-modulated, low-frequency sine wave. No other auditory cues are provided and even a severely-to-profoundly hearing-impaired listener can perceive these frequency modulations. Note that in order to facilitate perception, the measured voice fundamental frequency is lowered by a fixed amount prior to being recoded as a frequency-modulated sinewave. In this way, the frequencies of the coded signals are reduced to well within the range of the subject's low-frequency residual hearing.

In summary, frequency lowering has not, for the most part, yielded major benefits. The improvements that have been obtained have been modest and limited to specific applications, such as speech training, or to specific types of impairment, most commonly through provision of fundamental frequency cues to profoundly hearing-impaired people not able to benefit from conventional amplification. The amount of training required is an important factor that deserves further attention in experimental evaluation of frequency-lowering strategies. One clear indication that this very logical approach to exploiting residual low-frequency hearing is not easily implemented is that there are no frequency-lowering aids available commercially.

b. Other speech-analyzing techniques. Until recently, methods of processing signals for improving intelligibility were severely limited by technological constraints. As a consequence, frequency lowering, which is relatively easy to implement, has received a fair amount of attention. In contrast, other more complex methods of signal processing, such as techniques for enhancing specific phonetic features, have not received much attention. The increased availability of sophisticated digital signal-processing techniques has changed this situation, and there are now several ongoing research studies investigating the effects of manipulating speech parameters in order to improve intelligibility.

A relatively simple technique that holds some promise involves adjusting the consonant–vowel intensity ratio. This type of processing may be viewed as a phonetically based form of amplitude compression. Most of the consonantal sounds are weaker than the vowels and are rendered inaudible because of insufficient amplification (amplification may be limited in order to prevent the vowels from being too loud) or because of forward or backward masking effects in which weak consonants are masked by intense neighboring vowels. Experimental evaluations of adjustments to the consonant–vowel ratio have shown improved intelligibility for speech in noise with normal-hearing listeners (Montgomery et al., 1987). This technique holds some promise for improving speech intelligibility for hearing-impaired people, although it is not entirely clear that the same effects might not be achieved by simply amplifying the high frequencies, the frequency region containing most of the acoustic energy for consonants.

Signal processing to exaggerate specific phonetic features also appears to be a promising technique. Revoile et al. (1986, 1987) have shown that the discrimination of consonantal voicing by hearing-impaired listeners can be improved by exaggerating the durational cues associated with voiced and voiceless consonants. They have also shown improved place-of-articulation discrimination in fricative consonants by exaggerating the spectral shape associated with each of these consonants. Similar results have been reported by Guelke (1987).

In another interesting approach to this problem, Picshen et al. (1985, 1986) have assessed the improved intelligibility achieved when speakers make an effort to produce clear speech. There is a considerable increase in intelligibility scores for this sort of speech, compared with normal conversation. Unfortunately, however, the speech waveform of clear speech differs from conversational speech in ways that could not easily be accomplished by even a sophisticated digital aid (e.g., selective lengthening of some phonemes and of interword intervals).

There have also been attempts at resynthesizing speech sounds so as to emphasize the formant structure, but these experiments did not show improved intelligibility (Haggard, 1977; Summerfield et al., 1985). A recent attempt at emphasizing formant structure by means of multiband amplitude expansion, as opposed to amplitude compression, also resulted in poorer intelligibility (Bustamante and Braida, 1987). For the case of speech (vowels) in noise, however, improving the salience of the formants by decreasing the acoustic energy in the valleys between spectral peaks has resulted in improved intelligibility (Leek et al., 1987).

In principle, it should be possible to improve intelligibility by exaggerating those phonetic cues not easily perceived by the hearing-aid user. Thus far, the results have been mixed: Improvements have been reported for some forms of signal processing and not for others.

An important practical consideration in systems of this type is that the processing be automatic and errorless. If the speech feature to be processed is extracted incorrectly, the effect on speech intelligibility can be far more damaging than that of the hearing impairment itself. Erroneous cues can be more damaging than missing cues. Automatic, reliable extraction of speech features is not a trivial problem but an important issue to be addressed if practical, feature-based speech-analyzing aids are to be developed successfully.

A related issue of great theoretical importance is whether the recoding of speech features introduces distortions that may be more difficult for the hearing-impaired person to deal with than the unprocessed speech signal. If the speech signal is altered radically such that a new code has to be learned, then even if the recoded signal lies entirely
within the available range of residual hearing, there may be inherent cognitive limitations in learning the new code. It has been argued, for example, that the negative results obtained with most speech transposition systems are due to the transposed speech code’s being too complex to be learned, even though all of the speech information may be available in the coded signal. An alternative view is that the time course of auditory perceptual learning is so long that few laboratory experiments have revealed the extent of the average human listener’s ability to learn new codes (Watson, 1980).

As mentioned in the Introduction, the question whether hearing-impaired persons can learn to recognize radically distorted speech signals raises several very basic issues regarding learning, plasticity, and the information capacity of the impaired auditory system. One point of view is that, if a radically different speech code is to be learned, then this learning should occur during the early years of life, when a child is first learning the sounds of speech. The assumptions underlying this approach are very difficult to test, and there are substantial ethical implications in doing so. It may well be that a deaf child could acquire speech and language more readily if all of the important speech cues were transposed to lie within the child’s region of residual hearing. If, however, the experiment is a failure, the child’s acquisition of speech and language may be retarded even further than would otherwise be the case.

A counterargument to the above approach is that the auditory system may have inherent speech-feature detectors. The ability (of a child or an adult) to learn to recognize radical transformations that did not code speech with these features might therefore be severely restricted. The thrust of this argument is that acoustically normal speech is in some sense special and that the normal auditory system is uniquely structured for the processing of that form of information. Accordingly, any major deviation from the normal speech code might be extremely difficult to learn and, from this view, it would be questionable whether rates of communication comparable to that of normal speech could ever be achieved using such codes. Essentially the same argument can be made regarding the difficulties encountered in learning to communicate using visual and tactile representations of speech. While this argument may be of some theoretical interest, the success of highly skilled lipreaders and of deaf-blind users of the TADOMA method of speech perception appears to contradict a strict (acoustic) speech-is-special argument, as does the performance of some implanted children.

In commenting on an earlier draft of this report, Liberman pointed out yet another version of the speech-is-special argument. He has stressed that the perceptual unit may be the speech gesture, by which is meant neuromotor-control sequences normally associated with the production of phonemes or other speech units. From this view, any transformation of speech—acoustic, visual, tactile or electrical—might be successful as long as it preserved these gestures in perceptually salient forms. The abilities of excellent lipreaders, deaf-blind users of TADOMA, and some outstanding cochlear-implant patients Liberman considers consistent with this view, because each of those modes, unlike telegraphy or printed text, does maintain the integrity of the “code” (speech gestures). Liberman has developed this basic position, originally termed the “motor theory of speech,” in numerous articles over the past three decades (e.g., Liberman et al., 1967; Liberman and Mattingly, 1985). If groups of congenitally deaf children were eventually taught to communicate using a variety of different transforms of speech, some of which preserve the speech gestures and some that do not, the validity of Liberman’s and other “speech-is-special” arguments might be tested in a way that was never before possible. Meanwhile, the practice of choosing transformations that are as perceptually similar to speech as possible will probably continue, and it seems likely that none of the major speech theorists will find this objectionable.

3. Signal processing to reduce noise and reverberation

One of the most common complaints made by hearing-aid users is that speech in noise, or speech in a reverberant room, is particularly difficult to understand. Poor speech reception in noise and/or reverberation is to be expected since a reduced set of speech cues is available to hearing-impaired persons. They are therefore less able to make use of the normal redundancy in speech to compensate for the information degraded by the noise or reverberation. This problem is not unique to hearing aids and applies equally well to cochlear implants and to tactile and visual sensory aids.

For the common situation in which speaker-to-microphone distance cannot be controlled effectively, a major reduction in the level of the background noise or of the reverberation is very difficult to achieve. There are, however, several available techniques by which small improvements can be obtained. One method is the use of a directional microphone. Many modern hearing aids use directional microphones that provide several decibels of improvement in speech-to-noise ratio over omnidirectional microphones. User reaction to directional microphones in hearing aids has been mixed. Some find directional microphones very helpful, others do not. It should be noted that directional microphones are not effective in a highly reverberant room, and also that, for those conditions under which a directional microphone works well, it is necessary to point the microphone steadily toward the speaker, which is not always a convenient maneuver.

A second improvement, which can be implemented fairly simply, is to use two microphones. The typical approach is to place one microphone on each ear. In a true binaural hearing aid, the output of each microphone is amplified and delivered to the corresponding (ipsilateral) ear. In a quasibinaural aid, the outputs of the microphones are simply added, amplified, and routed to one, or possibly both, ears in parallel (Harris, 1965).

One of the major advantages of the two-microphone systems is that, for any given frequency, at any given instant in time, the speech-to-noise ratio at one ear will be greater than that at the opposite ear. Which ear has the larger speech-to-noise ratio will depend on the relative spatial locations of the speech and noise, the acoustic shadow produced by head diffraction, and the relative spectra of the speech and noise. Simply adding the two microphone outputs will improve the
speech-to-noise ratio by a small amount, on the average. Routing the amplified output of each microphone separately to each ear (as in a true binaural hearing aid) provides additional interaural cues that the auditory system can extract when processing the binaural output. For a normal-hearing person, the effect of such binaural processing can produce a significant increase in intelligibility (Licklider, 1948; Hirsh, 1950; Levitt and Rabiner, 1967).

In order to take advantage of this binaural processing capability, it is essential that the speech and noise originate from different spatial locations. In this way the interaural time and intensity differences associated with each sound source will be different, thereby providing the necessary binaural cues for the auditory system to extract. The magnitude of the improvement in speech intelligibility resulting from the binaural processing of interaural time or phase differences is relatively small (comparable to roughly a 3-dB reduction in background noise) but is not negligible. In contrast, the improvement in speech-to-noise ratio that can be produced by head-shadow effects (interaural intensity differences) is on the order of 6 to 7 dB (Markides, 1977; Hawkins and Yacullo, 1984). For theoretical treatments of the equivalent gain in speech-to-noise resulting from the use of binaural amplification, see Levitt and Rabiner (1967) and Zurek (1982).

a. Benefits of localization. In addition to reducing the effects of background noise, binaural hearing aids also provide important localization cues (Hirsh, 1950; Markides, 1977). These cues are not only important in their own right but also, by allowing the hearing-aid user to rapidly identify the direction of the signal source, the head can be turned appropriately to further improve the signal-to-noise ratio. Speech is not necessarily the signal of greatest interest in such situations. The combination of improved localization ability and improved detectability of warning sounds is particularly important, for example, when crossing a busy street.

b. Digital techniques for noise reduction or cancellation. Digital signal-processing techniques for improving speech-to-noise ratio or reducing reverberation are currently of great interest, since recent advances in the development of digital hearing aids make the implementation of sophisticated signal-processing methods technically feasible. The most effective of these techniques requires the use of two or more microphones. Noise-reduction techniques have also been developed for a single microphone input but, as will be discussed shortly, these techniques introduce other distortions that limit the overall effectiveness of such systems.

c. Multiple-microphone noise reduction. For those situations in which one of two microphones can be placed near the noise source so that it effectively picks up noise only, then an adaptive noise-cancellation technique can be used (Widrow et al., 1975). The signal picked up by the noise microphone (the reference) is adaptively filtered and subtracted from the corrupted speech picked up by the second microphone. The adaptive filter is adjusted until the difference is minimized in a least-squares sense. The correlation between the noise reference and the noise corrupting the speech signal need not be known initially in order for this method of noise cancellation to work. In essence, the adaptive filter is adjusted until this correlation is determined and the noise component cancelled by subtraction. This technique has been found to produce significant improvements in intelligibility for hearing-impaired listeners (Chabries et al., 1982, 1987; Brey et al., 1987), but of course it cannot be realized in a wearable hearing aid.

d. Dual-mounted head microphones. A more realistic situation for the hearing-aid user is that in which both microphones are mounted on the head. An experimental two-microphone hearing aid of this type using the Widrow method of adaptive noise cancellation has shown significant improvements in speech intelligibility (Schwander and Levitt, 1987). Several research groups are currently working on variations of this approach to noise reduction (Peterson et al., 1987; Brey et al., 1987; Chazan et al., 1987; Weiss, 1987). Although promising, these techniques also have practical disadvantages. Adaptive noise cancellation techniques are not instantaneous. If both microphones are mounted on the head (or body) and the head (or body) is constantly moving, then the system may not have time to adapt. A more serious problem is that these techniques, as in the case of directional microphones, provide little or no advantage in a highly reverberant room.

e. Single-microphone noise reduction. Reducing noise or reverberation with single-microphone systems is the most difficult case. A relatively simple technique is that of time-invariant linear filtering. This approach can be quite effective when the speech and noise spectra do not overlap significantly or if the interference is a narrowband noise or sinewave. Unfortunately, in most cases of practical interest, there is considerable overlap between the speech and noise spectra—and there are no simple solutions in this case. Because of its relative simplicity, time-invariant linear filtering represents a reference condition against which other methods of noise reduction should be compared. The optimum filter for maximizing signal-to-noise ratio, with stationary signals, is the Wiener filter (Lee, 1960). Speech, however, is not a stationary process and the ideal, noncausal Wiener filter cannot be derived for this case. It is possible, however, to approximate the ideal filter by a time-varying Wiener filter that adapts to changes in the statistical properties of the speech signal (Lim and Oppenheim, 1979).

f. Time-varying filters. Simple time-varying filters have been used in hearing-aid applications. In one such implementation, a high-pass filter with variable cutoff frequency is used. When the low-frequency content of the incoming signal is relatively high, it is assumed that environmental noise is present and the filter automatically adjusts its cutoff frequency to exclude these low-frequency components. If the relative power in the low frequencies is comparable to that of speech, the filter broadens its bandwidth to include these signals (Iwasaki, 1981). Many types of environmental noise are predominantly low frequency, so this type of hearing aid appears promising. Initial experimental evaluations have shown a significant listener preference for the self-adapting filter, although intelligibility scores did not show an improvement (Ono et al., 1980).

g. Graupe-Causey filters. A self-adaptive filter specifi-
cally designed for hearing-impaired persons has been developed by Graupe and Causey (1975). Preliminary evaluations of this self-adaptive noise filter show significant improvements in speech intelligibility (Stein and Dempsey-Hart, 1984). These studies, however, did not use time-invariant filtering as a reference condition, and therefore it is not known how much of the improvement is due to the adaptive characteristics of the filter and how much to simple filtering. Tyler (1988) describes very little improvement with this type of noise filter, or with a passive reduction in low-frequency gain. Similar findings have been reported recently by Van Tasell et al. (1988) and Klein (1989).

The technique developed by Graupe and Causey has been implemented on a single chip (Graupe et al., 1987), and several manufacturers have recently introduced noise-reducing hearing aids into their product line using this chip. Consumer reaction to this form of noise reduction has yet to be assessed.

h. Spectrum subtraction. Another approach to noise reduction is that of spectrum subtraction (Lim, 1978; Boll, 1979). According to this procedure, a running estimate of the noise spectrum is obtained, typically during pauses in the speech. The estimated noise spectrum is then subtracted from the short-term speech-plus-noise spectrum yielding a spectrum with a much reduced noise component. All of the above spectra are necessarily short term, i.e., time windows of finite duration are used. Since the procedure is designed to track time-varying changes in the short-term speech spectrum, these time windows should be shorter than syllabic durations in speech. Lim and Oppenheim (1979) have shown that, under certain conditions involving linear transformations of the signals, the short-term spectrum subtraction method reduces to that of adaptive Wiener filtering.

A key element in the spectrum subtraction method, as in the Wiener-filtering approach, is the accuracy with which the speech-plus-noise and noise-only spectra can be estimated. The mechanism used for deciding that noise only is present is thus of critical importance; the more accurate this decision, the more reliable is the subsequent noise reduction. Empirical evaluations of the spectrum subtraction method have generally shown improvements in speech-to-noise ratio, but no significant improvements in speech intelligibility, although improvements in overall quality have been reported (Lim and Oppenheim, 1979).

The subtraction process need not be restricted to the frequency domain. Furthermore, in addition to spectral transformations such as the Fourier transform, other nonlinear transformations can be used. In the approach used by Weiss et al. (1974), the square root of the amplitude spectrum is transformed to the time domain, by analogy with cepstrum analysis (Noll, 1967), and after nonlinear weighting (determined empirically) the subtraction process takes place. This method of noise reduction has yielded good results in terms of overall speech quality and in reducing fatigue when listening to speech in noise for long periods of time. Data obtained on hearing-impaired subjects show similar results (Levitt et al., 1986).

i. Comb filters: Speech-specifying filtering. In principle, the more that is known about the signal, the more effectively it can be extracted from the noise. A number of noise reduction techniques have focused on known features of the speech signal. One class of such methods utilizes the quasi-periodic structure of the speech waveform. If the voice pitch is known, a comb filter can be used to extract all of the harmonically related components of the speech signal and to reject the noise between these harmonics. The effectiveness of this approach, however, is critically dependent on the accuracy with which the fundamental frequency of the voice has been estimated. This creates an inherent problem in that in order to reject the noise effectively a comb filter with high-frequency resolution is needed, but this requires a filtering time that is fairly long, which, in turn, may not be able to effectively track temporal variations in the voice pitch.

In one implementation of the comb-filter technique (Lim et al., 1978), it was found that increasing the filter length improved the speech-to-noise ratio but reduced intelligibility. The longest filter length used was equal to 13 pitch periods of the speech. This produced an improvement of 9.8 dB in the speech-to-noise ratio, but intelligibility was reduced to less than half of that for the unprocessed condition. The shortest filter length used was equal to 3 pitch periods. This produced only a 3.4-dB reduction in speech-to-noise ratio, and no significant change in speech intelligibility. On the positive side, subjects reported that the quality of the processed signals was superior to that of the unprocessed speech in noise. [Note: Acoustical engineers typically attempt to process waveforms to maximize either speech quality, or speech-to-noise ratio (S/N), possibly because intelligibility is often at or near 100% in the systems under study. In the case of hearing-impaired listeners, however, the most important measure of performance is clearly the accuracy with which a listener can identify a message that was spoken.]

A case of special interest is that in which the interference is another speech signal. Parsons (1976) developed a method of harmonic selection for this particular problem. Although intelligibility data are not provided, Parsons reports that "suppression of the unwanted talker is virtually complete, except in a few cases where shared peaks produce some residual crosstalk."

j. Speech-model-based filtering. A second class of speech-specific procedures is concerned with modeling the speech production process and then estimating the parameters of this model. This approach can be used for both bandwidth reduction and noise reduction. Evaluations of this approach have been confined largely to the accuracy with which the model parameters can be estimated under noisy conditions. Although data show reliable estimation of certain speech parameters under noisy conditions (Lim and Oppenheim, 1979; Kobatake et al., 1978; Wise et al., 1976), no data have yet been reported showing an improvement in intelligibility of the reconstructed speech signal.

k. Phoneme-specific filtering. A third approach focuses on the acoustic structure of different classes of speech sounds. Drucker (1968) proposed a noise-reduction technique in which different filters are used for each of five classes of speech sounds (stop, fricative, glide, vowel, nasal). Each filter is designed to emphasize the salient characteristics.
tics of its respective sound class. This technique is attractive in that it takes advantage of a priori information about the acoustic–phonetic structure of speech. A practical problem with this approach, however, is that it is first necessary to automatically segment the incoming speech signal into these different sound classes. This is not a trivial problem. The development of reliable automatic segmentation has been a major stumbling block in the development of effective speech recognition by machine. If, however, the automatic-recognition aspect of the problem can be resolved, then substantial improvements in intelligibility can be obtained. Drucker (1968), for example, showed an increase in intelligibility from 50% to 75% when the segmentation was done by hand. Intense research efforts are under way on the segmentation problem, because of its importance to many industrial and military applications, but progress has been much slower than was originally expected.

Summary. In summary, noise-reduction amplification systems for hearing-impaired persons can be divided into two distinct groups: (1) relatively simple systems that work well under restricted conditions (e.g., personal infrared or FM-transmission systems that work well because they employ a microphone close to the speaker and a relatively noise-free transmission line) and (2) experimental systems that require a high degree of signal processing and for which substantial improvements in performance have yet to be demonstrated. Several manufacturers have recently introduced noise-reducing hearing aids using fairly sophisticated methods of signal processing (e.g., the Graupe-Causey filter), but data on relative performance of these hearing aids and consumer preferences have yet to be obtained.

D. Who should use conventional aids?

Conventional hearing aids have two important advantages in the habilitation and rehabilitation of hearing loss: (1) when conventional aids help even to a modest degree, that help commonly exceeds that which can be achieved by any other prosthetic device and (2) treatment can be terminated at any time with a loss only of the time spent (not a trivial matter in the case of small children) and the relatively modest cost of the device. For these reasons, conventional aids are the treatment of choice unless and until we can be certain that an alternative will be better.

Table III displays the general classes of aids most commonly recommended by present-day clinicians, as a function of age of onset and degree of hearing loss. This table summarizes recommendations typically made and considers only the most common four options. These are "conventional aid" (CA), "consider other prosthetic devices" (CO), "consider only other noninvasive fully reversible alternatives" (FR), and "none" (N). Age of onset has been divided into five categories based on the approximate ages at which significant changes occur with respect to hearing habilitation and rehabilitation. For example, if a child loses his or her hearing during the prelingual period (<3 years), the prognosis regarding language and speech development and the decisions regarding what ought to be done are profoundly affected (Levitt et al., 1987b).

<table>
<thead>
<tr>
<th>Age at onset</th>
<th>Mild &lt; 40 dB</th>
<th>Moderate 40–70 dB</th>
<th>Severe 70–90 dB</th>
<th>Profound &gt; 90 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very young</td>
<td>N</td>
<td>CA</td>
<td>CA</td>
<td>FR</td>
</tr>
<tr>
<td>&lt; 3 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>FR, CO</td>
</tr>
<tr>
<td>3–18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>CO</td>
</tr>
<tr>
<td>18–55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old</td>
<td>N</td>
<td>CA</td>
<td>CA</td>
<td>N</td>
</tr>
<tr>
<td>&gt; 75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*: Depends on age and developmental factors.

Although age of onset is a critical factor, the current age and status of an individual are also important. For example, if a person is currently over 75 years of age and has a variety of age-related disabilities, one needs to take these complicating factors into consideration before recommending a particular rehabilitative strategy.

The second dimension in Table III is degree of hearing loss. It is divided into four categories. Decibel values of hearing level (average loss at 500, 1000, and 2000 Hz) are attached to each descriptor of loss. These are intended as a general guide, not as definitive values.

Entries of a particular symbol in a cell indicate that this is the most commonly recommended aid. In some instances, individual characteristics suggest the need for a different course of action from that indicated in the table. However, the strength of the evidence needed to justify an alternative recommendation varies widely across the cells. For example, some individuals over 75 years are appropriately treated like younger persons and vice versa. Such variations are easily justified and are generally encouraged, where supported by reliable clinical assessments. However, a surgical procedure on a child under 3 years of age that would destroy residual cochlear functioning would require much stronger justification. Note that an entry of "CO" indicates only that alternatives to conventional amplification should be considered. Whether an alternative is actually implemented or not should depend on a thorough evaluation of the person's performance at the present time in relation to realistic expectations for the alternative under consideration.

Two additional qualifications regarding Table III are needed. First, we have not described a relatively new class of sensory aid for the hearing impaired known as assistive listening devices (ALDs). ALDs range from mild-gain electroacoustic instruments designed for very specific listening situations, such as aided television viewing, to visually oriented alerting/warning devices designed for the profoundly impaired. Individuals falling in any cell of Table III could potentially benefit from the use of ALDs. Second, we note that many school-age children (5–6 years of age and older)
with average hearing loss less than 25 dB may require the use of amplification. The negative impact of mild hearing loss, in the range 15–25 dB, and of unilateral hearing loss on the educational development of children has received much attention in recent years (Bess, 1985, 1986; Oyler et al., 1987).

As noted in Sec. IV, these recommendations, while typical of most practitioners, are not universally accepted, particularly by those advocates of cochlear implants who favor early intracochlear implantation of profoundly deaf children.

1. The effect of age

Age is an important variable when considering the application of prosthetic devices. In the case of young children, educators generally believe that (conventional) hearing aids should be applied as soon as possible. This includes children as young as a few months of age (Rodel, 1985; Krantz, 1985). It is not so clear whether aids should be applied in the case of a mild loss in a child under 3, although it is certainly not categorically ruled out. Few such losses are identified at this early age, in any case, and the stresses associated with keeping an aid on such a young child are considerable for both the child and the parent. Training the parent in more effective communication skills is probably a better choice in most situations.

For the somewhat older child, hearing aids are more likely to be applied successfully. This together with evidence that even mild hearing losses can significantly affect language and learning skills of school-age children (Bess, 1985; Davis et al., 1986) supports the need for prosthetic devices for virtually all hearing-impaired persons in this age category.

The candidates for prosthetic devices other than conventional hearing aids are found mainly among the profoundly hearing impaired. (This point is discussed in more detail in the next two chapters.) However, in the case of very young children (birth to 2–3 years), there are two considerations that appear to preclude some options: (1) it is not possible to determine the hearing status of a child in this age range with as much certainty as in the case of older persons and (2) because these children are young, they will live a long time and during that period revolutionary new developments will probably occur. Therefore, procedures should not be undertaken that cause permanent detrimental changes in the auditory system either directly or through a lack of utilization.

In the case of many older individuals, special considerations include a general inability to cope with amplification devices. For this reason very elderly persons with mild losses may not profit enough from amplification to offset the difficulties of using it. The same may be true of the very old with very severe losses. In general, however, each person is very much an individual case for whom the best course of action depends on a variety of variables. At this time, there is no method of aid selection that is clearly more effective than a trial period with an aid that is well selected in relation to the hearing loss and to the personal needs and capabilities of the individual.

2. The effect of hearing loss on speech intelligibility

As discussed in earlier sections, conventional hearing aids make an acoustic signal more intense overall, make it relatively more intense in some frequency regions than others, limit its maximum intensity level, and reduce its intensity range. Furthermore, signals arriving from some directions may be reduced relative to those arriving from another. However, for all that, the basic nature of the signals is unchanged from the original. That is, it is a time-varying acoustic waveform that encodes, among other things, those features most listeners recognize as speech. For persons with mild to moderate hearing losses, and even for some with severe losses, this simple alteration of the stimulus processing is very much preferred to more drastic recoding because it retains the essential temporal and spectral features of the normal speech signal. Preservation of a relatively familiar signal also obviates the need for lengthy, difficult, and often only partially successful training of the user to understand a new information encoding scheme. For these reasons and because those with mild-to-moderate losses often hear quite well with conventional aids, they are the aids of choice for these groups.

With greater degrees of hearing loss, several factors combine to reduce the effectiveness of the conventional hearing aid. A greater hearing loss requires greater amplification to achieve audibility. This means that both the hearing aid and the ear itself may produce more distortion. Upward spread of masking may also be a problem at high sound levels. Furthermore, because the threshold of discomfort does not increase as much as the threshold of detection, the dynamic range of the ear can be very much reduced. For these reasons, a hearing aid must reduce the signal's dynamic range, sometimes sharply, if the hearing loss is great. As a result, the features that contribute to intelligibility are reduced further. Finally, in cases of severe-to-profound hearing loss, damaged or missing peripheral auditory-system components may be incapable of performing preliminary analyses of the signal that are necessary if the properties of the speech waveform are to be successfully communicated, via the auditory nerve, to the higher components of the auditory nervous system.

The combined effects described above may render many persons with profound hearing loss unable to derive substantial benefit from the acoustic signals provided by conventional hearing aids. Even persons with severe hearing losses may receive only very limited help from conventional hearing aids, but the amount of help provided still may be greater than that available by any other means at present. This may be true even for some persons with profound hearing losses. When a hearing loss is only moderate-to-severe in extent, it is almost certainly true that a conventional aid will provide better results than any current alternatives.

It follows from this discussion that the candidates for alternative prosthetic devices to the conventional aid are found among those with the greatest hearing losses, for the reason that those are the persons who process speech poorly, or not at all, with conventional devices. The basic question is which technology will provide the best results for each individual. As alternative methods of aiding the hearing im-
paired (e.g., vibrotactile aids, cochlear implants) are improved, more and more people who previously were candidates only for conventional aids might become candidates for the alternative treatments. Improvements in the performance of conventional hearing aids will, of course, have the opposite effect. A vigorous and continuing competition to improve both conventional and alternative aids will clearly benefit the population of hearing-impaired people.

E. The future of conventional hearing aids

The conventional electroacoustic hearing aid is clearly the current treatment of choice for the overwhelming majority of hearing-impaired people. The same conclusion was reached by the recent National Institutes of Health Consensus Panel on Cochlear Implants (National Institutes of Health, 1988). For this reason, ongoing research on both analog and digital signal processing, especially speech-specific processing, on noise reduction, and on selecting and evaluating these devices should be generously supported. That work holds the promise of improved speech perception for many millions of hearing-impaired people living today (see Table I). Whether the alternatives to the conventional aid eventually will begin to challenge its position as the primary personal aid to speech perception for any but the profoundly deaf remains to be determined. It is clear, however, that work in those areas also deserves support. Finally, any discussion of conventional electroacoustic hearing aids would be incomplete if it did not end with a statement of the critical importance of proper hearing-aid selection, fitting, and follow-up training and counseling. Clinical investigations have repeatedly demonstrated that the percentage of successful wearers of conventional aids can be improved considerably by proper fitting and follow-up (Brooks, 1979; Brooks and Johnson, 1981). Comprehensive training and counseling of patients may be as important as any other single variables in achieving patient success with hearing aids.

II. COCHLEAR IMPLANTS

In this section, we describe the characteristics of implantable cochlear stimulators that are currently available for clinical use. After a brief general description of this class of cochlear prosthesis, we review the auditory sensitivity and discrimination capabilities, or resolving power, that can be achieved with it. We then discuss clinical considerations, including potential risks, patient selection criteria, patient acceptance, and clinical results. The relative merits of single versus multichannel implants are described; the use of these devices in children is discussed; and the typical cost of a cochlear implant, compared with conventional hearing aids, is estimated.

A very brief, nontechnical description of the cochlear implant is as follows. In the simplest of these devices, a stimulating electrode is surgically inserted about 6–15 mm into one of the tubelike ducts of the snail-shaped cochlea. This placement locates the tip of the electrode close to the auditory nerve cells normally responsible for converting mid- and high-frequency sounds into neural impulses. In this simple, single-channel configuration, as developed by House and Urban (House et al., 1976), the electrodes are connected to an induction coil that is implanted under the skin of the mastoid. This internal coil, in turn, can be activated by current passing through an external coil; only a thin fold of skin separates the internal and external coils. Current is delivered to the external coil from a special signal processor, which translates the acoustic energies of sounds, especially speech, into an electrical waveform intended to be effective as a stimulus for auditory neural structures.

The major alternative to this simple, single-channel stimulation system is the multichannel device, which consists of a bundle or array of electrodes, up to 20 or more, each of which, when implanted, terminates at a different distance into the cochlear duct. These multichannel devices are used in conjunction with electronic systems that analyze incoming sounds into acoustic categories. Each type of sound causes a different electrode, or group of electrodes, to be activated. The simplest, and thus far most common, of the multielectrode systems employs a frequency-coding scheme (low to high) in an effort to return some pitch-discrimination abilities to the deaf patient. Electrical connection from the signal-processing unit to the multichannel implant is made either via a percutaneous connector (a direct electrical connection through the skin), or using the induction-coil strategy described for single-channel devices. Another form of direct cochlear stimulation is the extracochlear stimulator, in which an electrode is placed near the entrance to the cochlear duct, but not actually inserted into it. Because the purpose of this report is primarily to describe levels of speech identification that can be achieved with various devices, and also because the electronics used in some implants are proprietary, only very general descriptions of the devices and their operation are provided.

In 30 years, cochlear implants have gone from a speculative possibility to a therapy that is considered routine by many of its practitioners. Although this rapid development and the diverse interests, therapeutic and experimental, that have been involved have resulted in a large body of published literature on implants, many fundamental questions remain unanswered. The practical consequence of this situation is that data from different devices and laboratories cannot be compared unequivocally. Despite this situation, several facts are clear: Electrical stimulation of the cochlea can produce sensations that postlingually deafened implantees agree are auditory; these sensations can be interpreted in terms of sounds, to the benefit of the patient; and, finally, most subjects are able to use electric stimulation as an effective aid to speechreading (lipreading), while a limited number are able to understand some speech without the aid of visual or contextual cues.

With this background, the following compilation of results attempts to define the type, degree, and range of benefits obtained from cochlear implants. These results represent a snapshot of a rapidly evolving field, in which empiricism sometimes far outdistances theoretical understanding. Research on cochlear implants is likely to benefit our understanding of hearing and communication in all modalities, as well as contribute to refinements in future generations of implants, to the greater benefit of those who receive them.
A. Basic psychophysics of the electrical stimulation of hearing

This review provides an overview of basic psychophysical measures collected from patients with several different single- and multichannel cochlear-implant devices. Detailed descriptions of the implants are provided in the articles cited. In general they were intracochlear implants, with the stimulating electrodes placed about 15 to 25 mm into the basal turn of the scala tympani. Single-channel stimulation is generally accomplished with the shorter electrodes, multichannel with longer. Whenever possible we have chosen results from subjects with direct percutaneous connections to their electrodes in order to avoid possible transmitter/receiver artifacts. There is, however, no reason to believe that there are any substantial differences in the responses of subjects for similar signals actually reaching the electrode, whether or not the signal is internally demodulated by a totally implanted stimulating system. Some reported results, as noted in the text, have favored the best observations rather than the average or worst. The intent of much of the published literature on cochlear implants has been to characterize the sensations that can be elicited by electrical stimulation of the peripheral auditory system, rather than to document the successes or failures with all implanted patients.

1. Pitch

a. Rate pitch. The perceived pitch of stimulation increases with stimulation rate (the rate of pulses delivered to the implanted electrode or electrodes) from about 60 Hz through 350–400 Hz for most subjects. Below 60 Hz, subjects may report pulsating sounds (e.g., telephone ringing, ratchets and such). Above 400 Hz, there is fusion of the input pulses, with no further increase in pitch (Simmons, 1966; Bilger, 1977, 1983; Tong et al., 1982; Townshend et al., 1987). These sounds above fusion frequency often generate perceptions of considerably higher pitch than the sensation a normal hearer might describe for a 400-Hz sine wave, sometimes by several octaves. Some subjects are able to discriminate slightly higher rates, but claims of rate discrimination at higher frequencies (such as 1000 Hz) may be attributable to changes in stimulus intensity.

The best frequency difference limens (DLs) are about 5% for stimuli at 100 Hz, 5%–10% at 200 Hz, and 8%–15% at 300 Hz, when collected with a two-alternative, forced-choice paradigm (Simmons, 1966; Hochmair-Desoyer et al., 1983). These compare to approximately 0.2% for unimpaired people listening to acoustic waveforms (Wier et al., 1977). Many patients do considerably worse. Waveform—sine versus rectangular pulses—does not alter DLs substantially. Most of the data have been collected from subjects at a "comfortable loudness." The effects of stimulus duration have not been thoroughly studied. The DLs do worsen slowly when burst durations are decreased below 100 ms and are probably optimal at 200–300 ms. DLs not only differ among subjects, but can also differ among electrodes on the same subject. The reliability of many DL measurements may be marred by contamination with loudness changes.

b. Place pitch. Subjects implanted with more than one electrode report (with varying degrees of conviction) that stimulation of individual electrodes by pulses having no rate–pitch information (single pulses or pulse trains above the fusion frequency) produces different pitch sensations (Townshend et al., 1987). The ordering of these characteristic pitches roughly corresponds with the electrode location within the scala tympani. That is, when subjects are asked to rank pitch—as a floor-to-ceiling height, by anecdotes, or by a sharpness-dullness continuum—the lowest pitch is likely to be obtained at the most apical electrode and the highest pitch at the most basal electrode (Tong et al., 1982). The sensations accompanying the stimulation itself, whether or not rate is also varied, are not described as "pure," in the sense of listening to a sinusoid, but seem rich in harmonics. One subject was able to compare electrically stimulated pitch with acoustic matching in his opposite, normal-hearing ear, but no exact matches were possible (Eddington et al., 1978).

There are very few data available on the consistency of pitch ranking among subjects. It is clear that some perceive these electrode-specific sensations more easily than others. Some can "correctly" rank all electrodes and are seldom confused during paired-comparison trials. Others have electrodes in which the low-to-high ranking, apex to base, has discontinuities or local reversals. Still others report very little pitch difference among electrodes, and this is confirmed by poor or inconsistent ranking scores.

c. Range of percepts. The variety of anecdotal descriptions suggests ranges of several kilohertz for some subjects to only assorted "buzzes" for others. Both rate and place of stimulation affect anecdotal ranking decisions within the rate–pitch range (Eddington et al., 1978; Atlas et al., 1983). For example, a 200-Hz stimulus on one electrode may be ranked lower in pitch than 200 Hz on another electrode, and the corresponding anecdotal description might describe a large truck horn versus a car horn. There have been anecdotal reports of experimenters feeding back these presumed pitches as simple tunes, with limited success.

d. Intensity effects. In most, but not all subjects, stimulus intensity has important effects on pitch. The pitch can either increase or decrease with intensity, depending on the subject (Townshend et al., 1987). Occasionally, pitch increments created by as little as a 2-dB intensity increment can produce anecdotal descriptions suggesting rises of an octave and more. These intensity–pitch effects, aside from not being totally predictable from one electrode to the next, have probably led to errors in some reported psychophysical results. It may be that pitch judgments are so level dependent that intensity must be randomized above and below the level of equally loud stimuli across frequency, to avoid confounding effects.
2. Loudness

a. Waveform/charge considerations. Investigators have chosen a variety of waveforms (typically sine or some form of rectangular pulse) and measures of intensity (volts, rms current, peak current, coulombs, etc.) in reporting their results. Across the range of frequencies, growth of loudness is typically a power function, and values of the exponent as high as 3.5 have been obtained, although the exact value varies widely from measurement to measurement (Pfingst, 1984). The exponent varies with frequency; for example, values of 3.36 at 1000 Hz and 1.6 at 100 Hz have been reported by Shannon (1983a,b) for sinusoidal stimulation.

The rapid growth of the intensity of the sensation of loudness implied by exponents of 3.0 or larger is similar to that described for the feeling of electric shock applied to the skin. This may be contrasted to the considerably lower rates of growth of the sensations of loudness or brightness, in response to increasing sound or light intensity, for which the exponents are approximately 0.66 (loudness) and 0.33 (brightness) (Stevens, 1975). Most simply, the range of physical intensities between stimuli that are barely detectable, and those that are so intense they are painful, is very large for lights and sounds but quite small for direct electrical stimulation.

b. Thresholds. The literature contains a surprisingly large range of electrical thresholds (Shannon, 1983a), which presumably reflects not only waveform differences, but nerve survival, location of electrodes, and modes of stimulation [mono- or bipolar with narrow (2 mm) and wider (10 + mm) separations of the electrode pair]. On the average, from lowest to highest thresholds by place of implantation are: within the nerve (modiolus); within scala tympani; on the round window membrane; in the round window niche; and (highest) on the promontory. There is considerable overlap in absolute current requirements among the first three locations. When calculated in coulombs, to correct for waveform differences, that range is still about 20 dB for comparable stimulation. The lowest thresholds, occasionally less than 1 μA, are obtained with 50–100 Hz sine waves and monopolar stimulation (Pfingst and Sutton, 1983). The highest have been in the vicinity of 600 μA for bipolar stimulation at 1.5-mm separation, with 0.36-ms biphasic pulses (Tong et al., 1982).

c. Dynamic range. As defined in Sec. I, the dynamic range refers to the difference, in decibels (dB), between threshold and maximum tolerable loudness of electrical stimulation. This range may vary from less than 1.0 to 30 dB and depends on the patient, the stimulation frequency, the mode of stimulation (bipolar or monopolar), and the stimulus waveform. The largest dynamic range is obtained with sinusoidal stimulation at low frequencies (50–200 Hz) and is typically up to 30 dB, principally because of the lower thresholds at these frequencies. That range begins to narrow at about 250 Hz, reaching 10 dB or less by about 1 kHz (Pfingst, 1984).

Figure 2 shows a schematic representation of the dynamic range for sine wave stimulation, from threshold to maximum loudness. There is considerable variability about this curve for individual patients. Once the "knee" of the threshold curve is passed at about 100 Hz, the slope increases at between 5 and 15 dB per octave (Pfingst, 1984). The dynamic ranges described by Pfingst (1984) were obtained using an animal model, and it is important to note that somewhat broader ranges (40–50 dB) at low frequencies, narrowing to 20 dB at higher frequencies, were obtained using psychophysical methods with human implant patients (Shannon, 1983a).

d. Intensity discrimination. Perhaps more important than dynamic range is the listener's ability to detect loudness differences within that range, since electronic signal processing can easily compress the range of acoustic intensities into the patient's electrical dynamic range. The loudness DL is independent of stimulation frequency and has been estimated to require between a 2.5 and a 5% intensity change (Shannon, 1983a; Simmons et al., 1979). Smaller DLs have been reported, but only rarely (Hochmair-Desoyer et al., 1983).

e. Stimulus duration. As with normal hearers, loudness increases with stimulus duration. Loudness approaches an asymptote at 100 ms and is maximum by 300 ms (Shannon, 1985).

f. Perstimulatory adaptation/fatigue. There are no truly parametric studies of threshold decay with steady-state stimulation; however, both fatigue and adaptation do occur. The duration of poststimulus fatigue (for threshold measurements) is a function of stimulus intensity, and at high stimulus intensities can last from several hundred ms to at least several minutes. Adaptation is generally less for rectangular pulses than sinusoids (Simmons, 1966; Hochmair-Desoyer et al., 1983; Shannon, 1983a).

g. Temporal discrimination. Preservation of certain types of timing information in at least some cochlear implant subjects is remarkably good. Two sequential pulses can be distinguished from one (as a "single versus double" percept) at 2-ms interstimulus intervals, thus rivaling auditory resolutions by normal listeners (Greene et al., 1983). When asked only for a "same/different" judgment, interpulse intervals as brief as 0.5 ms can be detected (Simmons, 1966). Some patients can detect 8–10 ms gaps in conventional gap
detection experiments, while others require very long silent intervals in excess of 100 ms (Hochmair-Desoyer et al., 1983; Dobie and Dillier, 1985; Moore and Glasberg, 1988; Shannon, 1989).

This has led some researchers to speculate that people with very good temporal discrimination may be better candidates for single-channel implants than those with poor time discrimination. The use of this measurement as a preimplant diagnostic requires validation, however. It also would require some residual hearing, or a preliminary electrode placement.

3. Percepts with multiple electrodes

a. Electrode interactions. Longitudinal spread of electrical charge among several electrodes stimulated simultaneously in the scala tympani is a significant but poorly resolved aspect of multielectrode stimulation. In general, bipolar stimulation between two closely spaced electrodes produces less interference than does monopolar stimulation. The principal index of such interactions has been loudness interactions, e.g., when two closely spaced electrodes, say 5 mm, are stimulated at their respective individual thresholds simultaneously, the resulting loudness is likely to be very much greater than one might expect from comparable acoustic loudness summation in a normal ear. Furthermore, such loudness changes are not always logically predictable in the same patient among different electrodes, or between patients (Shannon, 1983b).

At this writing, the adjustment of loudness for simultaneous multielectrode stimulation remains empirical. Some stimulation schemes purposefully avoid interaction by sequential, rather than simultaneous, stimulation.

b. Other percepts with multiple-electrode stimulation. Studies of speech or speech-sound discrimination have preempted studies of conventional psychoacoustics of simultaneous stimulation, except for threshold and loudness. It is not even entirely clear whether, when two electrodes are stimulated simultaneously, persons hear two separate sounds or a single fused percept.

B. Risks involved in cochlear implants

The risks of a cochlear implant can be divided into the risks associated with any similar ear surgery performed under a general anesthetic, and those risks specific to the implant. For the surgery itself, there is a 1:5000 incidence of death or serious morbidity from a general anesthetic in an otherwise healthy individual. The possibility also exists of temporary or permanent injury to the auditory or vestibular nerve, of wound infection, and, in some implant procedures, disruption of the ossicular chain. Incidence data are not available on these risks; however none of these can be considered major hazards.

Specific to implants are the following:

1. Damage to the inner ear resulting in loss of any residual hearing (this has occurred).
2. Degeneration of ganglion cells to the point that the implant becomes nonfunctional. [While the incidence is unknown, there appears to be no evidence that short-term use (3 years or less) has caused degeneration to a nonfunctional state. Experience for longer times is confined to smaller numbers of patients, but there is also no evidence of degeneration after 5–6 years of use.]
3. Temporary or permanent balance disturbances. (This has occurred in several patients temporarily after surgery, and there are patients who have experienced balance disturbances upon stimulation. Insofar as we are aware, none of these has been disabling.)
4. Failure of electrical stimulation to produce sound. (This occurs.)
5. Extrusion of the implant. (This has occurred.)
6. Foreign body reaction or wound infections, including meningitis. (This occurs.)
7. Device failure for a variety of reasons. (This occurs.)
8. Inappropriate expectations by the implant patient regarding hearing results. (This occurs.)
9. Possibility of a middle-ear infection entering the cochlea via the implant, then causing damage to the membranous labyrinth and/or meningitis. (To our knowledge, this has not occurred.)

The discussion of negative results from cochlear implants should be more open. In fact, there is very little in the medical literature, other than some individual case reports and some data on animal models (e.g., Burgio, 1986; Leake-Jones and Rebscher, 1983; Sutton, 1984; Zrunek and Burian, 1985). Thus it is impossible to provide comprehensive incidence figures for associated risks. It is reasonable to presume that some of these risks increase with significant invasions of the cochlea, i.e., an electrode(s) extending more than about 6 mm into the scala tympani, or with the use of a "hard-wired" percutaneous connector. Some risks may be decreased by so-called extracochlear stimulators, wherein the stimulating electrode is placed outside the cochlea, on the round window or round window niche. (Extracochlear devices may increase the likelihood of facial-nerve stimulation.) Whether or not these stimulators are as effective a communication aid as intracochlear devices is questioned by many practitioners. Too few patients have had extensive experience with the extracochlear devices to provide a solid, data-based evaluation of their performance (also see National Institutes of Health, 1988).

The above risks, on balance, do not appear to be serious in totally deafened adults. The question of long-term nerve degeneration associated with the use of intracochlear electrodes remains to be answered.

C. Patient selection criteria

Originally, implantation was restricted by general agreement to postlingually deafened adults who could not benefit from a hearing aid. At this time, both the 3M House and Nucleus devices are approved for experimental trials for implantation in children of various ages. The House Ear Institute has also implanted prelingually deafened adults, the opposite ears of hearing-aid users, and at least a few patients with residual acoustic response in the implanted ear.

Virtually nothing is known at the present about the actual as opposed to the published criteria for implant candi-
Most implantation groups use one or more of three general criteria:

1. An audiometric criterion—no response at the limits of the audiometer or inability to detect speech at 70 dB SPL (wearing a hearing aid; Berliner, 1985).

2. A behavioral or a hearing-aid-use criterion—test administered wearing a hearing aid and measuring discrimination of a variety of mainly speech-related sounds (Berliner, 1985; Owens et al., 1985; Tyler et al., 1984b). Some of these tests are at a suprasegmental level: word or syllable stress, number of syllables, question versus statement, etc. No uniform criteria have been adopted.

3. Inability to achieve a specified level of performance with a tactile aid is also sometimes used as a criterion for accepting a candidate for implantation. (But, good tactile performance has also been considered a positive indication for implantation, as noted in the Introduction.)

D. Acceptance by patients

An important overall measure of the efficacy of a device is its eventual acceptance for daily use. To achieve this simple but formidable goal, a prosthesis must provide benefit that is not only significant (in the statistical sense), but also meaningful or important enough to justify the effort required to maintain and use it. The size, reliability, convenience of use, maintenance expense, and even appearance all enter into a patient's decision to use a device regularly.

The 3M House single-channel device has been implanted in more patients than any other implant, and more long-term data are available on its performance. A 1984 update from this group (Brimacombe et al., 1984) shows that, excluding patients who have died or were still "in process," 176/213 postlingually deaf patients (83%) were daily users. Prelingually deaf patients (age of onset < 5) were slightly less likely to become daily users (24/35, 69%). Children (under 18) were more likely to be daily users (143/151, 95%), but they also may have been less free to choose whether to wear and use their processors (Berliner and Eihsenber, 1985). As discussed in later sections, the value of the implant for prelingually deaf persons has not been clearly demonstrated and considerably more research is needed on this population.

Patients who were not daily users usually were stimulable but chose not to wear their stimulators because of inadequate benefit or cosmetic concerns. Device failure accounted for 2.9% of all adult implants, and 2.5% of patients received no auditory sensation from electrical stimulation (Brimacombe et al., 1984).

Implant use for a group of 38 patients was high with the Nucleus device, with apparently only one patient choosing not to use it, and an average daily use of 10.4 hr/day for all 38 patients (Mecklenberg and Brimacombe, 1985). Several reports indicate daily use rates of 67%-100% for smaller groups (Dent et al., 1987; Dowell et al., 1985; Douek et al., 1983; Engelmann et al., 1981).

Clearly, the majority of cochlear-implant patients become daily device users; most will probably continue to use their devices, although long-term follow-up is available on only a fraction of the total patient pool. Device failure, particularly for more complex systems, may erode these results, but, when necessary, successful reimplantation has been accomplished (Thieleneir, 1985).

E. Patient results and performance: Postlingually deafened adults

The results presented here are a representative selection of those in the literature: they are not intended to be exhaustive. Where sufficient data are available, cited studies have been restricted to those conducted in the English language (for speech tests), using implant devices that are identical for large groups of patients. Test data are given for devices that are beyond the "custom" stage of development, since custom devices differ from subject to subject in ways that have an uncertain influence on results. Generally, although there are differences among different cochlear implant devices, the differences among patients with identical devices are at least as great as those among the devices. The information presented is intended to demonstrate the range and nature of the data obtained. We have also emphasized results from standard testing materials, rather than data from tests particular to individual laboratories.

1. Environmental sounds

The perception of environmental sounds has generally been assessed in identification tasks, using tape recordings of sounds made by common objects and animals. The Minimal Auditory Capabilities (MAC) battery (Owens et al., 1981) includes such a subtest, and similar tests have been used by many investigators. Table IV summarizes the results of several groups on these tests. Subjects are clearly able to get information about the identity of environmental sounds from their devices. This appears to be true of all the devices. Preimplant performance with a tactile or acoustic aid is shown in parentheses in this and in Tables V and VI, for those cases in which it was measured.

2. Discrimination of speech characteristics

An elementary auditory ability is the discrimination of speech from other sounds. The noise-voice test of the MAC battery is a good example of such abilities. Other similar tests assess the discrimination of different, nonlinguistic characteristics of speech sounds, such as male/female voice discrimination, speaker same/different, and spondee same/different discriminations. The results presented in Table V show that subjects are able to make some of these discriminations quite well.

3. Prosody

The perception of prosodic cues can be assessed by several measures: stressed word in the sentence, number of syllables, question/statement discrimination, etc. Each of these
TABLE IV. Perception of environmental sounds with cochlear implants (percentage correct responses) (SC = Single-channel; MC = Multiple-channel; IC = Intracochlear).

<table>
<thead>
<tr>
<th>Authors (date)</th>
<th>Device type</th>
<th>N</th>
<th>HRRC 20-item</th>
<th>Environment sounds</th>
<th>20-item closed</th>
<th>20-item open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eisenberg et al. (1983)</td>
<td>SC-IC</td>
<td>86</td>
<td>55.1</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Edgerton et al. (1983)</td>
<td>SC-IC</td>
<td>10</td>
<td>40</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Tyler et al. (1985)</td>
<td>SC-IC</td>
<td>3</td>
<td>...</td>
<td>60</td>
<td>28.3</td>
<td>28.3</td>
</tr>
<tr>
<td>Mecklenburg and Brimacombe (1985)</td>
<td>MC-IC</td>
<td>37</td>
<td>...</td>
<td>31</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Tyler et al. (1984b)</td>
<td>MC-IC</td>
<td>2</td>
<td>...</td>
<td>80</td>
<td>47.5</td>
<td>...</td>
</tr>
<tr>
<td>Dowell et al. (1985a)</td>
<td>MC-IC</td>
<td>6</td>
<td>...</td>
<td>27(13)*</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Schindler and Kessler (1987)</td>
<td>MC-IC</td>
<td>8</td>
<td>...</td>
<td>42</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

* "Modified" MAC battery tests.

TABLE V. Speech discriminations with cochlear implants (percentage correct responses) (SC = single-channel; MC = Multiple-channel; IC = Intracochlear).

<table>
<thead>
<tr>
<th>Authors (date)</th>
<th>Device type</th>
<th>N</th>
<th>MAC noise/voice</th>
<th>Male/ Female</th>
<th>Spondee same/different</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edgerton et al. (1983)</td>
<td>SC-IC</td>
<td>11</td>
<td>70</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Tyler et al. (1985)</td>
<td>SC-IC</td>
<td>3</td>
<td>58.7&quot;</td>
<td>71.7</td>
<td>75</td>
</tr>
<tr>
<td>Mecklenburg and Brimacombe (1985)</td>
<td>MC-IC</td>
<td>33</td>
<td>80'</td>
<td>83.5</td>
<td>80</td>
</tr>
<tr>
<td>Tyler et al. (1984b)</td>
<td>MC-IC</td>
<td>38</td>
<td>...</td>
<td>85(66)</td>
<td>...</td>
</tr>
<tr>
<td>Dowell et al. (1985a)</td>
<td>MC-IC</td>
<td>2</td>
<td>95&quot;</td>
<td>95</td>
<td>...</td>
</tr>
<tr>
<td>Schindler and Kessler (1987)</td>
<td>MC-IC</td>
<td>6</td>
<td>96(69)&quot;</td>
<td>83(68)</td>
<td>88(70)</td>
</tr>
<tr>
<td>Chance</td>
<td></td>
<td></td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

* Augmented test 40 items.

* "Modified" MAC tests.

TABLE VI. Perception of prosody cues with cochlear implants (percentage correct responses) (SC = Single-channel; MC = Multiple-channel; IC = Intracochlear).

<table>
<thead>
<tr>
<th>Authors (Date)</th>
<th>Device type</th>
<th>N</th>
<th>Word</th>
<th>Stress</th>
<th>MTS test</th>
<th>MAC battery tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eisenberg et al. (1983)</td>
<td>SC-IC</td>
<td>86</td>
<td>34.5</td>
<td>73.9</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Edgerton et al. (1983)</td>
<td>SC-IC</td>
<td>12</td>
<td>34.5</td>
<td>73.9</td>
<td>65</td>
<td>...</td>
</tr>
<tr>
<td>Mecklenburg and Brimacombe (1985)</td>
<td>MC-IC</td>
<td>37</td>
<td>34.5</td>
<td>73.9</td>
<td>68(51)</td>
<td>...</td>
</tr>
<tr>
<td>Tyler et al. (1984b)</td>
<td>MC-IC</td>
<td>2</td>
<td>34.5</td>
<td>73.9</td>
<td>57.5</td>
<td>50</td>
</tr>
<tr>
<td>Dowell and Orth (1985)</td>
<td>MC-IC</td>
<td>2</td>
<td>34.5</td>
<td>73.9</td>
<td>85</td>
<td>...</td>
</tr>
<tr>
<td>Schindler and Kessler (1987)</td>
<td>MC-IC</td>
<td>6</td>
<td>34.5</td>
<td>73.9</td>
<td>46(49)*</td>
<td>80(47)</td>
</tr>
<tr>
<td>Chance</td>
<td></td>
<td></td>
<td>34.5</td>
<td>73.9</td>
<td>80</td>
<td>70</td>
</tr>
</tbody>
</table>

* "Modified" MAC tests.
measures has been used for one or more implants. Table VI gives results on several prosodic tests applied to different devices. Many more data are available in the literature, but are based on tests in languages other than English, or tests only used within one laboratory, and so cannot be compared.

4. Consonant and vowel confusions

An elemental test of the speech-discrimination ability of an implant subject is identification of consonants and vowels, or at least correct classification of these speech sounds into categories. This type of testing has not been as extensively performed as many of the other forms of speech testing, such as word and sentence tests.

The specific consonants and vowels, and means of testing their identification, have differed widely between testing sites and devices tested. Some test a small set (4–5) of stimuli, while others use extensive sets (up to 19 or so). In a few cases, overall percentage correct scores are reported, but raw response matrices are not presented. Some results were collected in native languages other than English, or in dialects that differ from American English. As a result of these and other differences, a detailed analysis of the results from individual studies is difficult and interpretation may be influenced by one's theoretical position. Nonetheless, several generalizations can be made about the results.

Generally, independent of implant type or the particular device used, the information that allows speech discriminations tends to be related to the envelope of the signal, rather than its spectrum or its phonetic characteristics. In agreement with this generalization, results with hearing-impaired and with normal-hearing listeners demonstrate that more information than has generally been recognized may be available in the speech envelope, and that this information may be of particular value when combined with visual cues (Erber, 1972; Van Tasell et al., 1987).

For consonants, voicing is probably the major cue available to implant subjects, though even it is used imperfectly (Tyler, 1984b). Many subjects are also able to make some manner distinctions, i.e., to distinguish between nasals, stops, and fricatives; these discriminations are probably based primarily on characteristics of the time-intensity envelope of the signal and, to a lesser extent, on low-frequency information. In a few cases, discrimination may be made between consonants differing in place of articulation. This discrimination probably results from the implant subject's sensitivity to differences in the waveform envelope, rather than to spectral correlates of the place distinction. This is an important interpretative point, since place discrimination is generally assumed to be based on sensitivity to the signal's spectral characteristics.

Vowel discrimination has been tested even less frequently than consonant discrimination. Primarily, subjects are sensitive to differences in vowel duration and to the frequency of the first formant (Rosen et al., 1985). First-formant frequency can help distinguish some vowels. Because the first formant is usually a low frequency, discrimination of first-format frequency is probably mediated by processing of the waveform in the time domain in a manner similar to that of the fundamental frequency.

Cochlear implants have the promise of being a useful aid to speechreading, since many articulatory features of phonemes, such as voicing, are reflected in sound, but are not visible. Subjects generally show improvements with speech-reading of consonants or vowels in nonsense syllables when stimulation is added to visual cues; however, the improvements are not always substantial (e.g., Tyler et al., 1984b, 1985). This is particularly true, of course, of individuals who were already good speechreaders prior to receiving a cochlear implant. Rosen et al. (1981), showed that the considerable amount of information that is coded in voice pitch, combined with its relative invisibility, makes the voice pitch an important aid to speechreading.

Performance of the most successful multichannel implant patients is exemplified by subject "SS" on whom a very complete series of (listen-only) speech tests are reported by Dorman et al. (1988). This listener was said to be the "best, or equal to the best, of 100 listeners using the Symbion device." He was able to discriminate as well as most normal listeners on phonemic continua that could be resolved by purely temporal cues ("split-slit") and did very well for stimuli that differed in the frequency of the first formant. SS identified monosyllabic words at 62 percent correct, with a range of 44–78 percent correct over sessions.

For further details of the data collected, interested readers are referred to original sources for the original consonant and vowel discrimination data.2

5. Word identification

Most cochlear-implant patients are not able to recognize words without being able to see the face of the talker. Many patients are able to recognize some spondaic words (e.g., baseball, hot dog), but this may be due to a combination of gross time-intensity information and memory for a limited number of spondee words with which they are familiar. It is difficult to interpret the recognition of numbers for the same reasons.

Some of the most successful patients are able to recognize some unfamiliar single-syllable words in isolation. Performance of greater than 25 percent correct word recognition has been reported in 5 out of 12 subjects with the Vienna device (Hochmair-Desoyer et al., 1985); 1 out of 15 with the UCSF/Storz device (although two others scored 24%) (Schindler and Kessler, 1987); and 5 out of 7 patients with the Melbourne device on the Boothroyd list (but 0 out of 7 on the NU-6 list) (Dowell et al., 1984). Banfai et al. (1984a) reported that, in an undifferentiated group of 22 prelingually and 13 postlingually deafened patients, the average score on an open-set monosyllabic word test was 17 percent correct. Eddington and Orth (1985) reported on one patient who scored 36 percent correct on the NU-6 list.

6. Sentence recognition

When comparing visual testing to auditory-plus-visual testing, nearly all patients show an improvement in word recognition in sentences with their implants. The amount of
enhancement depends, in part, on their baseline lipreading skill, which varies enormously across individuals. While some of the most successful individuals can recognize some words in sentences without visual cues, that remains the exception rather than the rule. The following discussion is therefore limited to the results of auditory-plus-visual testing.

The recognition of words in sentences is typically better than the recognition of isolated words, presumably due to the addition of grammatical and contextual cues. Several groups, particularly those groups who reported patients recognizing words in isolation, have noted a few subjects who could recognize more than 25% of the words in sentences (e.g., Dowell et al., 1985a). The obtained values depend on the number of presentations of each test sentence, the internal predictability of the words in the sentences, and the manner of scoring (key words, every word, meaning).

Dowell et al. (1985a) reported that one patient out of six scored 38% on the CID everyday sentence test. The other patients got only two or three words correct. Tyler et al. (1984b) tested two additional patients with the Melbourne device; one obtained 33 percent correct recognition of words in sentences with a familiar speaker, and both scored about 40 percent correct when a picture of an object mentioned in the sentence was presented as a cue. Burian (1984) noted that 5 out of 14 postlingually deafened patients scored higher than 25 percent correct on sentences. Hochmair-Desoyer et al. (1985), reporting on what are presumably at least some of the same patients, reported that 7 out of 12 patients scored above 25 percent correct recognition of words in sentences.

7. Tracking

De Filippo and Scott (1978) described a measure of proficiency in the reception of connected discourse, which has come to be known as tracking. In this method, the speaker (tester) reads aloud from a printed text, and the listener (subject) must repeat the message (a word, phrase, or sentence, depending on the level of training) verbatim. According to the original description of the method, errors must be resolved using a hierarchy of strategies (repeat the word or phrase, paraphrase or use synonyms, spell it, etc.), but only the verbal channel is to be used, with no recourse to purely visual input (charades, signing, finger-spelling, etc.). The speaker proceeds when all previous words have been correctly repeated. The subject’s score is reported in terms of words per minute (wpm) correctly communicated. Normal subjects can perform at about 110 wpm (i.e., half the normal reading-aloud rate, since both tester and subject have to repeat all words); 70 wpm has been proposed as an approximate threshold of social adequacy (Levitt, 1985).

Perhaps the greatest advantages of this test are its simplicity and face validity. Only a trained tester, some reading material, and a stopwatch are required. Not only can this test be applied by all groups doing cochlear implant work, but it is also applicable to other sensory aid or substitution systems, such as tactile aids. Tracking has unusual face validity as a practical measure of communicative benefit, since it uses connected discourse. Compared with element-processing tests like those of word and phoneme recognition (Dent et al., 1987), the tracking task is easier because of contextual cues but may be more difficult because of the requirement to process at high speeds (up to 10 phonemes/s in normal speech). A final advantage of the tracking method is that it can also be used as a form of therapy—both in rehabilitation classes and as homework.

However, there are certain difficulties in using the tracking method to analyze and compare results across groups (Tye-Murray and Tyler, 1988). Materials have not been standardized and obviously will vary in difficulty and resulting wpm scores. Because the primary goal is to assess the contribution of the implant as an aid to speechreading, the comparison between scores obtained by vision only and those obtained by vision plus stimulation is of greater importance than the absolute levels of either score. There is considerable variability in tracking scores for individual 5–10 min sessions, even after accounting for obvious sources of variability, such as degree of learning, different speakers, and different materials. Some of this variance appears to result from abrupt changes in the strategies used by speaker and the receiver. There appears to be no well-accepted convention for dealing with these problems, but most authors (admirably) present tracking results on a trial-by-trial basis.

Table VII summarizes the speech-tracking data obtained by several cochlear-implant groups. Where a certain group has published more than one paper including tracking data, the most recent has been chosen. A disappointing omission from the table is the Vienna group, which has reported no tracking data so far. Groups with all types of devices—single- or multichannel, intra or extracochlear—found that their systems provided substantial aid to speechreading for most or all of their subjects. When this was tested statistically (e.g., Robbins et al., 1985; Dent et al., 1985), the gains were clearly significant (i.e., large enough that they were probably not due to chance alone). There is considerable variability across subjects; one group (Robbins et al., 1985) found that prior success with a hearing aid was correlated with ability to benefit from the implant. Inspection of the data suggests that some of the variability might also be reduced by stratifying subjects according to vision-only (VO) scores) speechreading skills: the largest percentage gains are made by subjects who are fair-to-good speechreaders, while the largest gains in wpm seem to accrue to subjects who are fair-to-good speechreaders (VO = 10–30 wpm). Some subjects even show a decrement in VO performance over time, as they come to rely less on speechreading. (Other explanations are possible: for example, both the tester and the patient may become frustrated with the VO condition and may not try as hard as before). Clearly, it could be misleading to present only a sound-plus-vision to vision-only ratio [(S + V)/VO] at asymptote as a measure of benefit, when this ratio can be artificially inflated by a reduction in VO performance. None of the subjects reported to date has achieved the 70 wpm level, but this must be tempered by the recognition that tracking rates for normal subjects (110 wpm) do not apply fairly to the testing of deaf subjects, who must look up from the written material, before speaking, to permit speechreading. Under these conditions, even normal subjects track at only about 75 wpm (Robbins et al., 1985).
It should also be mentioned that the range of wpm for vision plus cochlear stimulation in Table VII includes the values reported for vibrotactile aids plus speechreading, as described in the following chapter.

Comparison across groups shows that the ranges of all groups overlap, in both raw scores and percentage gain. High test–retest variability was apparent in all series, as was a substantial learning effect, for most subjects. The dramatic effect of different reading material has been demonstrated by Diller et al. (1983) and could account for some differences among different groups. Somewhat more troubling are the procedural differences across groups: some permitted the use of writing and signing (Robbins et al., 1985) or pointing to phoneme charts (Tyler et al., 1984b). Thus it would be premature to conclude that any reliable differences have been demonstrated among different cochlear-implant systems on the basis of the tracking data. It is hoped that investigators continue to present tracking data in detail (performance by trials) and that procedural details should be reported explicitly. In particular, reading materials should be standardized, perhaps with a library of readings of similar difficulty. Comparisons among patients based on this test may be difficult, even with a standardized protocol.

F. Patient results and performance: Prelingually deafened adults

As remarked earlier, many implant centers do not consider prelingually deafened adults for cochlear implantation. The results that are available show clearly that the performance of prelingually deaf adults does not match that of postlingually deaf adults, even to the extent of being less likely to become daily users of the device (see Sec. II D).

Discussed here are some of the few reports that explicitly separate prelingual data from postlingual data. Other reports may include prelingual and postlingual data, but in a way that cannot easily be separated. These reports should be interpreted carefully, since they are based on small numbers of patients, with unknown and diverse etiologies of deafness; all are based on results from non-English-speaking patients.

Burian (1984) reported on 10 prelingually deafened patients using the Vienna implant. Five patients scored at chance (not defined) on a vowel recognition task (test not described), and five scored between 75 and 87 percent correct. Six patients were unable to score above chance on a consonant recognition task, and the other four scored between 69 and 81 percent correct. These scores were generally poorer than those obtained from postlingually deafened patients tested by the Vienna group, but there was some overlap. Burian also noted that 6 of his 10 prelingually deafened patients could recognize some sentences (task not described), with scores ranging from 13 to 45 percent correct.

Chouard et al. (1983) present results from 15 prelingually deafened patients. Using a 20-word test (tested 6 months following implant), only three patients showed an improvement of more than 15% in the sound-plus-visions condition compared with the vision-only condition.

Dillier and Spillman (1984) reported on two prelingually deafened patients who tracked at 12.1 and 17.8 wpm in the vision-only condition, which increased to 13.2 and 23.3 wpm in the sound-plus-visions condition.

G. Multiple- versus single-electrode stimulation

The potential for speech perception using multiple- versus single-electrode stimulation is still unresolved in the minds of many investigators. Excellent results have been obtained with some patients with both types of stimulation. It is not yet clear what proportion of patients implanted with each device achieve significant levels of open-set word recognition. This issue is further complicated because such a wide range of performance is observed with most implants. The most recent reports available suggest that adult patients with the multichannel Symbion and Nucleus implants generally outperform patients with the single-channel 3M-House implant (Gantz et al., 1987). However, comparable performance has been achieved by some patients with the single-channel 3M/Vienna and Duran/Cologne implants (Banfai et al., 1984b). It should also be mentioned that multichannel processing does not always yield even good performance. Multichannel devices have been manufactured that have not produced better results than is typical of single-channel devices in any patients (Tyler et al., 1989a, b). There have been several studies comparing single- and multi-electrode stimul-
lation in the same subject with opposing results. However, in most studies there is a serious problem in experimental design: a patient with considerable experience with one stimulus mode (or "code") is unlikely to perform optimally during a relatively brief exposure to another. Until this issue is addressed, conflicting results will probably continue to be reported.

H. Patient results and performance: Children

There is no issue in cochlear implants more controversial than the implantation of children (Simmons, 1985; Tyler et al., 1987; Popelka and Gittleman, 1984). Many of the points of disagreement can only be alluded to here, but they include: the possibility of implanting a very young child who would have performed better with a hearing aid in the implanted ear than with the implant; unknown effects of long-term (lifelong) electrical stimulation; possible unreliability of implanted devices and the need for repeated surgery; uncertainty regarding the proper educational environment for implanted children (manual versus oral, or total-communication programs); and the chances of damaging a cochlea that might have been implanted with a greatly improved device at some future time.

Though cochlear implants have already gained acceptance for selected adult patients, there is a consensus that implantation criteria should be more conservative for children (see National Institutes of Health, 1988). Put another way, many clinicians and scientists agree that a higher benefit-risk ratio is required to implant children. In opposition to arguments against implanting children is the demonstrated importance of early experience on the development of language. It is possible that early electrical stimulation of the auditory system may be an important aid to auditory system development, to the acquisition of linguistic skills, or both.

Questions such as these need to be addressed by studies of the effects of implants in those children who do receive them. The difficulty with such studies is that they do not produce clear or quick answers because of: the difficulty of testing children's performance; the necessary longitudinal nature of the studies; the diversity of educational settings of the children; the difficulty of long-term follow-up; and the small number and relative heterogeneity of the samples of children available for this research.

Assessment of the benefits in those children who have been implanted has been hampered by two additional problems. One is that changes in speech-production skills, language, and educational achievement, while being desirable benefits of the implant, are influenced by many variables and this complicates the attribution of changes to the implant (Tyler et al., 1987). Another problem relates to the tests that have been chosen by implant researchers and clinicians to measure benefit. Unfortunately, the validity of two of the most commonly used tests, the Discrimination After Training test, and the Test of Auditory Comprehension, has been questioned. The former may encourage the teaching of test items, and the latter is heavily dependent on cognitive maturation.

In this section, only a small amount of data has been represented, since there are few published studies. More studies are now being conducted and should be published within the next few years. The categories of benefit discussed here relate best to adult users of cochlear implants. Anecdotal reports of diminished hyperactivity and similar positive effects of the implant have not yet been quantified and are not considered here. Although many researchers and parents have high hopes for the effectiveness of cochlear implants in children, the absence of appropriate research makes it as yet impossible to know whether these expectations are valid. There are currently two investigational devices being implanted in children in the United States: the 3M/House and the Nucleus devices (Berliner and Eisenberg, 1987; Clark et al., 1987). Implantation of children with the 3M/House device began in 1980, and there are now over 270 children that have been implanted. (Manufacture of the 3M/House device had been discontinued at the time of printing of this article.) The Nucleus children's implant program began in 1987. Early tests of children's abilities with implants did not show maximum scores on elemental closed-set tests such as the MTS test. More recently, however, open-set discrimination testing (on a highly selected group of children) has been reported in which 9 of 10 children with the 3M/House implant showed some open-set word recognition and 10 of 19 showed open-set comprehension on items from the Glendonald auditory screening procedure, administered without visual cues (Berliner and Eisenberg, 1987). Recent testing by Moog and Geers (1988) has confirmed, though primarily on closed-set tests, that at least some children with cochlear implants perform remarkably well. The natural development of children's sound and language capabilities means that assessing the value of the implant to deaf children requires long-term study. The results reported here represent only the beginning of those that should be obtained.

1. Environmental sounds

Thielemeier et al. (1985) indicated that the average child (N = 32) tested 1-year following implant could perform at level 1 of the test of auditory comprehension. At this level of the test, the subject must discriminate between linguistic versus human nonlinguistic versus environmental sounds. This does not imply that the children could recognize environmental sounds, since gross time-intensity cues could be used to distinguish speech from nonspeech stimuli. Although the specific number is not stated, some children could not perform this task following implant. Popelka and Gittleman (1984) reported data for one 8-year-old boy who had received the 3M/House implant and scored at chance on this task. Berliner and Eisenberg (1987) reported that one of their better implanted postlinguistic children (age 15 years) was able to score 95 percent correct on the House Ear Institute environmental sound test.

2. Stress/prosody

Thielemeier et al. (1985) reported that the average score of 30 children tested at 1 year following implant was at level 6.8 on the Discrimination After Training test (Thielemeier, 1984). The levels of this test progress on a hierarchy of primarily prosodic discriminations. The mean level for 106 children before implant was 2.2. This suggests that the children may be making progress in using the stimulation pro-
vided by the implant, although the increasing age of the children would also tend to result in improvement, and no control data on children with comparable hearing losses are available for this test. The data vary over the entire range of the test (levels 0 through 12), and it is not known how many children could not perform these tasks. Popelka and Gittelman (1984) reported that the 8-year-old boy they tested scored at chance on the MAC question/statement test. Preoperative testing on the same child with conventional amplification showed above-chance performance.

3. Speech discrimination

Popelka and Gittelman (1984) reported that the 8-year-old boy they tested performed at chance on the MAC spondee same/different test.

4. Vowel/consonant recognition

No data are available from children on their ability to recognize vowels or consonants.

5. Word recognition

Berliner and Eisenberg (1987) reported that ten of the children they considered their better overall performers scored from 0 to 75 percent correct, with a mean of 39 percent correct, on word recognition using the 12 words from the Glendonald auditory screening procedure. The child tested by Popelka and Gittelman (1984) scored 0 percent on the MAC word-recognition test and 22 percent correct on the MTS word-recognition test.

6. Sentence recognition

Berliner and Eisenberg (1987) tested 19 of their better children on the Glendonald sentence comprehension test, for which the child must respond to (but not repeat) ten questions. In all, ten of the children were able to answer at least one question correctly. The average score was 16 percent correct (1.6 questions answered correctly). The 8-year-old boy tested by Popelka and Gittelman (1984) showed no benefit in the speechreading of words in sentences on the MAC speechreading test.

7. Tracking

No tracking data have been published on implanted children.

I. Cost

The cost of cochlear implants includes that of the device itself, which ranges from approximately 5,000 to 12,000 dollars, depending on the number of channels and the complexity of the processor, plus the cost of the surgery, related medical services, and hospitalization. To these are added the costs of postoperative follow-up and adjustment of the processor. These costs are not extreme for medical treatments, but they are considerably greater than those of other types of speech-perception aids. One of the difficulties for prospective patients may be the problem of assessing the cost-benefit ratio for alternative systems, in consideration of the considerable differences in cost.

The Appendix lists the major characteristics of five of the most widely used implant devices available in the United States. (Note: Costs and other data on commercial devices in this report have not been updated since its original completion in 1988.)

III. SENSORY SUBSTITUTION: VISUAL AND TACTUAL METHODS

In this section, we discuss alternatives to the presentation of acoustic or electrical transforms of speech to the auditory system. Some of these nonauditory alternatives are considered natural, in the sense that they do not require transformation devices, while others, here termed synthetic, do. Before discussing specific nonauditory forms of speech reception, we consider the relative merits of auditory versus nonauditory representations of speech, with particular attention to the critical issue of the age of the onset of deafness. We then describe a variety of natural communication methods currently in use within the deaf (or deafblind) communities. The next section reviews results obtained with synthetic systems. Consistent with the rest of this report, we focus on systems for speech reception rather than speech training. Also consistent with the character of most research on synthetic systems, we focus on the reception of speech rather than of environmental sounds. Even systems that are marginal for aiding speech reception can enable the user to detect and identify a wide range of environmental sounds. Finally, we consider issues related to clinical applications of sensory substitution aids.

A. Auditory versus nonauditory speech perception

Both conventional hearing aids and cochlear implants present information to the hearing-impaired individual by stimulating the impaired auditory system. An alternate strategy is to display the incoming information to a different sense modality—i.e., sensory substitution. The logical support for this strategy increases with the magnitude of the hearing loss, the extent to which the impairment involves the auditory nerve and more central portions of the auditory system, and the degree to which the onset of the impairment preceded language learning. We ignore here the idea that audition is biologically necessary for speech decoding, because results on visual and tactual speechreading demonstrate that speech can be understood when it is received through other sensory systems. This view does not deny, of course, that the auditory modality may be (and probably is) the sensory channel best matched to the speech code. It remains possible that other channels may be sufficiently adaptable to serve as effective alternatives.

The visual sense, like the auditory sense but unlike the tactual sense, has evolved to detect and discriminate signals generated at a distance. The visual sense also appears to be superior to the tactual sense with respect to information-processing capacity. Compared with audition, however, the visual sense is poorly matched to rapidly changing stimuli.
In addition, it is normally more heavily used than the tactual sense. Thus it is not clear a priori which of the two senses is most appropriate for exploitation as a hearing substitute. For those individuals who are blind as well as deaf, of course, there is no choice: only the tactual sense is available (since smell and taste are clearly inadequate because of their inability to follow rapidly changing stimuli).

In general, the communication methods that employ nonauditory displays can be subdivided into "synthetic" and "natural." The synthetic methods make use of acoustic inputs and require a device to transform the acoustic energy into a visual or tactual stimulus. These methods are products of the scientific research community and, for the most part, are still in the experimental stage. The natural methods make use of nonacoustic inputs and do not require any transformation device. Most of these methods were born of necessity within the deaf community and have been used extensively for many years. Whereas the synthetic methods operate on acoustic signals, the natural methods operate on nonacoustic signals associated with the acoustic signals (e.g., the lipreading signals associated with talking) or nonacoustic signals that are specially designed for deaf individuals (e.g., the signs in sign language). The former type of natural methods, e.g., lipreading, will be referred to as natural-general and the latter type (e.g., signing), which requires special knowledge or behavior on the part of the sender, as natural-special.

The synthetic methods have greater ultimate potential as substitutes for hearing because they, like hearing, directly sense the acoustical environment. The natural methods are important, however, because they are the methods actually being used by the deaf population (and the methods in which subjects have received long-term training) and because they provide important information on the capacity of the visual and tactual senses as substitutes for hearing. In addition, one of these methods, namely visual speechreading (lipreading), is usually assumed to be available to the user when considering any type of speech-perception aid. Independent of whether the aid is a tactile aid, a conventional hearing aid, or a cochlear-implant aid, when the hearing loss is profound the aid is usually envisioned as an aid to lipreading.

The relative attractiveness of the visual versus the tactual channel is different for natural methods and synthetic methods. Whereas in the natural domain vision has the advantages of not requiring direct physical contact and of functioning at a distance, in the synthetic domain (where the input is acoustic for both visual and tactual systems), this advantage disappears. This distinction is reflected in the fact that natural visual methods are used by essentially all deaf individuals who are not blind, while research on synthetic methods has mostly included work on tactual displays.

It should also be noted that the research groups that have been involved in development of synthetic systems, i.e., sensory-substitution aids, differ in certain ways from those involved in the development of cochlear implants. Whereas the latter effort has been largely an effort of the medical/clinical community and has received extensive financial and industrial support, the former is primarily the work of university-based investigators and has the characteristics, typically, of more academically oriented research. One of these academic characteristics is unwillingness to attempt full-scale clinical trials on impaired individuals until very late in the development of new devices.

8. Postlingual versus prelingual deafness

The relative usefulness of sensory-substitution aids versus cochlear implants may depend to a great extent on whether the patient is postlingually or prelingually deaf. Because of the great importance of this issue, it is discussed before reviewing the properties of sensory substitution aids.

1. Postlingually deaf people

For patients who previously had a sense of hearing and learned language before the onset of deafness (postlingual deafness), the cochlear implant has the advantage of providing the patient with the important psychological benefit of "hearing again" (i.e., of experiencing some form of auditory sensation). Thus, even if a sensory-substitution aid resulted in equivalent communication performance, it might not provide equivalent subjective satisfaction. Furthermore, for postlingually deafened patients, the amount of learning/training required to fully exploit a sensory-substitution aid is likely to be greater than that required to fully exploit a cochlear implant. With the sensory-substitution aid, one must learn not only to distinguish among the sensations produced by the aid, but also to correctly match these sensations with external acoustic events. Although there is undoubtedly some transference across senses, and some learning of this sort must also take place with the implant, for this class of patients the amount of learning that is required for optimal performance is likely to be less with the implant. To date, most patients treated with cochlear implants fall into the postlingual category.

2. People deaf since birth

For patients who have been deaf since birth, the above advantages of cochlear implants may be diminished. A patient to whom a sensory-substitution aid is applied at birth, like a patient to whom a cochlear implant is applied at birth may, develop a sense of hearing that meets scientifically meaningful criteria for that sensory modality. Not only will both patients "hear" according to an objective, behavioral definition of hearing (with a discriminative capacity that depends on the resolving power of the prosthetic device and the stimulated sensory apparatus), but both patients will experience the impinging acoustic energy in terms of externalized object/events rather than feelings in or on their own bodies. Just as a normal-hearing person does not localize an acoustic stimulus in his or her ear (and becomes conscious of the role played by the ear in sensing the acoustic environment only when the relation between the ear and this environment is altered), it seems very likely that the person who has grown up with a tactile aid probably will not localize an acoustic stimulus on the body surface stimulated by the tactile aid. It is also possible that the advantage of cochlear implants with respect to the above-mentioned learning problem might disappear. For both the sensory-substitution aid and the cochlear implant, the learning required to construct
the appropriate perceptual space will be substantial (as it is in the development of the normally hearing child). The ages at which clinical application of the various types of aids is appropriate are decreasing and now appear to include young children. Further decreases, leading to the inclusion of infants as candidates for treatment, will depend on further development of the aids and of the methods for early detection of deafness, as well as the experience gained with children.

C. The natural methods

The principal natural methods that are in use are speechreading, cued speech, fingerspelling, and signing.

Speechreading methods are natural-general methods, visual or tactual, in which the stimuli read by the deaf individual are generated by the actions of the face associated with speech production. Visual speechreading (lipreading) involves viewing the face of the talker and is used to some extent by all individuals when communication is difficult. Tactile speechreading (Tadoma or the vibration method) involves feeling the face of the talker and is used by a small number of individuals who are blind as well as deaf.

Cued speech and fingerspelling are natural-special methods in which the sender employs special behavior but still operates within the language used by the normal-hearing community (e.g., English). In cued speech, hand signals are used to supplement the facial signals used in speechreading. In fingerspelling, hand signals are used to represent letters of the alphabet. Fingerspelling, of which there are many variants, is used extensively in both the visual and tactual modes (in the tactual case, the sender's hand is placed in the hand or hands of the receiver). Cued speech, which, like fingerspelling, was developed for the visual mode, is rather new and has not yet been used in the tactual mode.

Signing is a natural-special method that involves a special language. Again, it was created for the visual mode; however, it is also used extensively in the tactual mode (with the receiver's hands being used to monitor the configurations and positions of the sender's hands). Various combinations of signing and fingerspelling (or, more generally, "manual communication") are used by a substantial portion of the world's deaf population. The use of such methods is especially likely if the onset of deafness is early and the hearing-impaired person belongs to a community in which the use of such methods is common. Although sign language and its use has been a major area of philosophical, educational, and political controversy (as well as a focal point of linguistic research), this discussion is confined to issues related to the successful recognition of spoken language.

1. Speechreading

a. Visual. Apart from signing, visual speechreading or lipreading is the main natural sensory-substitution method used by hearing-impaired people. Many synthetic methods are regarded (and designed) as supplements to lipreading. Lipreading performance tends to vary widely across subjects. Whereas some individuals receive only marginal help from lipreading (even after substantial formal or informal training), others are able to follow continuous speech almost error free. Although general intelligence and language competence are obviously relevant factors, and some progress has been made in the attempt to delineate some of the more subtle sources of this variation (e.g., De Filippo, 1980; Williams, 1982; Ronnberg et al., 1982; Ronnberg and Lyxell, 1986; Lyxell and Ronnberg, 1986), the ability to predict individual performance is still far from perfect.

Aside from information conveyed through facial expressions, most of the information received through visual speechreading concerns identification of speech segments; the suprasegmental aspects of speech, such as stress, number of syllables, and intonation contour, are generally not well received (Risberg and Lubker, 1978). Identification of consonants and vowels in isolated nonsense syllables by hearing-impaired subjects shows average scores of roughly 60% correct, in tests for which chance performance is less than 10% correct (Heider and Heider, 1940; Walden et al., 1977; De Filippo, 1980). Confusions among consonants correspond mainly to errors in the detection of voicing or the discrimination of manner of articulation, whereas among vowels the confusions relate to similarity of lip shape. Identification of monosyllabic words drawn from large or open sets (i.e., sets in which the number of response alternatives is indefinitely large) by profoundly deaf subjects show scores in the range 30–60% correct (e.g., Nicholls and Ling, 1982; Erber, 1972; De Filippo, 1984). Comprehension of sentences, as measured by key-words correct or by speed of tracking (De Filippo and Scott, 1978), varies widely. Key-words correct varies over the whole range 0%-100% for hearing-impaired subjects (Johnson, 1976; De Filippo, 1980; Fujikawa and Owens, 1978, 1979). Tracking rates as high as 50–65 wpm have been observed (Grant, 1980; De Filippo, 1984; De Filippo and Scott, 1978; Sparks et al., 1978a, 1978b), as well as much lower scores (e.g., Brooks, 1984).

b. Tactual. In tactual speechreading, speech is received by placing a hand on the face of the talker and monitoring the mechanical actions of the face associated with speech production (jaw movement, lip movement, oral airflow, and laryngeal vibration). This method, often called Tadoma (after Tad and Oma, the first children to use the method) has recently been the subject of considerable research interest (e.g., see Norton et al., 1977; Reed et al., 1982a, 1985a); however, its current use among deaf and deaf-blind individuals is very limited (Schultz et al., 1984).

Results on nine deaf-blind adults who use Tadoma as their primary means of communication, and whose impairments were prelingual, show average percent–correct scores of roughly 60% for identification of 24 consonants; 45% for identification of 15 vowels; 40% for identification of monosyllabic words (open-set); and 65% for key words in CID sentences spoken slowly (Reed et al., 1985a). These results are similar to those obtained with normal subjects listening to speech in a background of white noise with a signal-to-noise ratio in the range 0–6 dB (Miller et al., 1951). Three of the more skilled deaf-blind subjects who have been studied were able to achieve rates of 30–40 words/min (roughly one-third the normal auditory rate) on the speech tracking task (De Filippo and Scott, 1978) using grade-school reading materials as text. As would be expected, most of the errors...
that occur in speech reception via Tadoma can be explained by its failure to make available information about tongue position.

Results on normal subjects with simulated deafness and blindness (Reed et al., 1978, 1982a, b) indicate that the deafblind subjects have no special tactile sensitivity: performance of the two groups is roughly equivalent on tests of basic tactile discrimination abilities. These results also suggest that learning Tadoma (for someone who already has language knowledge) is roughly comparable in difficulty to learning a complex foreign language.

In the same manner that lipreading demonstrates the adequacy of the visual system for receiving spoken speech, Tadoma demonstrates the adequacy of the tactual system for receiving spoken speech. As will be seen in subsequent sections, the performance achieved with these methods provides merely a lower bound; performance can be improved (in the visual case, and probably also in the tactual case) by augmenting the speechreading stimulus with other stimuli.

c. Visual and tactual combined. At least one individual, who is deaf but not blind, achieves excellent results by supplementing visual speechreading information with tactual information obtained by placing a hand on the shoulder and neck of the talker (Plant and Spens, 1986). The improvements obtained with this method, which the subject has used for 40 years, over visual speech reading alone are 88% versus 44% for consonant identification, 66% versus 33% for open-set word identification, 98% versus 85% for scores in CID sentences, and 63 versus 42 wpm in continuous discourse tracking. The results obtained when the hand is removed from the shoulder and stimulated instead with an experimental artificial tactile aid are intermediate between these two extremes. However, the total exposure time for each of the two aids tested was less than 4 h.

2. Cued speech

In cued speech, hand configurations are used to disambiguate stimuli that are confused in speechreading (Cornett, 1967). Eight hand configurations are used to code consonants and four locations to code vowels, and the hand cues are presented in synchrony with the spoken cue. Cued speech was designed to supplement visual speechreading and to date has been used only in the visual sense. The extent to which these same hand configurations can be used to improve tactual speechreading is only now being studied (Reed et al., 1987b).

The effectiveness of cued speech as a supplement to lipreading is illustrated by the study of Nicholls and Ling (1982) on 18 children with profound losses who were trained in this method for at least 4 years. On the average, identification performance with nonsense syllables improved from 30 or 36 percent correct with lipreading or cueing alone to 84 percent correct with the two combined. Similarly, for identification of the final word in low-predictability sentences, the scores were approximately 25% (lipreading alone), 43% (cueing alone), and 97% (lipreading plus cueing). As stated by Nicholls and Ling (1982:267-268):

The subjects' responses under the combined condition were outstandingly and uniformly good and merit considerable attention. The system enabled all of the subjects to receive precise phonemic and linguistic information both at a syllabic level and in running speech. Speech reception at an equally high level of accuracy by profoundly and totally hearing-impaired children has not been previously reported. The children's average scores ... are within the range of normal-hearing listeners' reception of similar materials through audition.

3. Fingerspelling

Fingerspelling is a system in which words are spelled out letter-by-letter using handshapes that correspond to the letters. It is sometimes used as a complete communication system by itself and sometimes as a supplement to American Sign Language (e.g., to communicate proper nouns for which signs have not been developed). Unlike cued speech but like signing, fingerspelling is used in the tactual mode as well as the visual mode.

At a typical transmission rate of five letters/s (roughly one-third the normal speaking rate), trained subjects can receive fingerspelling visually with negligible error rates. For example, for open-set identification of words in isolation, scores usually fall in the range 90-100 percent correct (e.g., Zakia and Haber, 1971; Thomson, 1984). At a rate of 15 letters/s, however, scores are reduced below 50 percent correct (Thomson, 1984). Roughly speaking, the function describing the dependence of percentage-correct on rate of presentation for visual fingerspelling appears to be similar to the equivalent function for windowed reading (reading through a hole or window in a mask that is moved across the text) with a window width of two-four letters (Thomson, 1984). The errors made in visual fingerspelling generally occur among letters represented by similar handshapes, finger direction, and finger identity (e.g., see Hawes and Danhauer, 1980).

In a recently initiated study of the tactual reception of fingerspelling (Reed et al., 1986, 1987a), five deaf-blind subjects achieved perfect scores on conversational sentences presented at rates of five letters/s and were able to achieve scores on the tracking test (De Filippo and Scott, 1978) of roughly 30 words/min (a result similar to that obtained by experienced Tadoma users).

4. Signing

The major varieties of sign language (used by roughly half a million people in the United States) are often described as falling on a scale that extends from manually coded or signed English (SE) to Pidgin Sign English (PSE) to American Sign Language (ASL). ASL is a natural and complete sign language; it not only has a grammar and vocabulary that is entirely distinct from English, but it employs communication techniques that differ from those used in any spoken language. SE makes use of manual signs to represent the English language; such systems are often devised
and used in educational settings. PSE is a blending of SE and ASL; it borrows heavily from the grammars and vocabularies of both SE and ASL. The learning of ASL or PSE, even with sophisticated instruction, is (like Tadoma) comparable to the learning of a complex foreign language. Apparently, those acquiring ASL after the age of roughly 14 years never achieve the skill of a native signer (Mayberry and Fischer, 1983), roughly comparable to the learning of a second spoken language after childhood.

Research on ASL has demonstrated that it has a vocabulary size and grammatical richness comparable to that of the world's spoken languages. Furthermore, it can be used to communicate ideas at roughly the same rate as ordinary speech. In the studies by Bellugi and her associates (Bellugi and Fischer, 1972; Klima and Bellugi, 1979), although signing rates were found to be roughly one-half speaking rates (2.4 signs/s compared with 5 words/s), the rates were found to be roughly equal when compared in terms of "propositions" (approximately 0.8 propositions/s). According to Bellugi and Fischer, the economy of sign-language grammar, the use of parallel channels, and the compacting of information by modulating signs, tends to compensate for the slowness of sign formation relative to word formation. Research on tactile reception of signing by the deaf-blind (Reed et al., 1986) shows that this method has roughly the same efficiency as Tadoma. For highly trained subjects, scores for the two methods decrease from approximately 80 to 65 percent correct (in key words or key signs) as the production rate increases from 0.3 to 0.6 propositions/s.

D. Synthetic systems

Synthetic systems, either visual or tactual, vary with respect to the level of information reduction (i.e., the number of classificatory decisions made) associated with automatic processing prior to stimulation of the subject. At one extreme, the display presents essentially the whole "raw" signal to the subject. Spectral displays, in which the short-term spectrum of the acoustical signal is presented to the subject, constitute the main example at this extreme. At the other extreme, one can imagine a synthetic system that contains a speech-recognition device at the front end and an output display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject). Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject. Although no such system is yet available for the deaf population, limited versions (e.g., in display of text to the subject.

Intermediate between these two extremes are synthetic systems, many of which have been studied in the past, that extract and display a specified subset of speech characteristics, such as whether the speech is voiced or unvoiced, whether it contains frication noise, or how the fundamental frequency of the voiced portion varies. The questions of whether and how to preprocess speech signals through the use of "smart front ends" incorporated in synthetic systems depend on the relative ability of the subject's perceptual-cognitive system to extract the essential elements of complex signals versus the scientist's ability to incorporate the right choices into the system design (and the engineer's ability to realize these choices effectively). The results of research performed to date do not provide unequivocal answers to these questions.

A major limitation of most synthetic systems is the lack of perceptual richness in the display. It is well known that in order for a display to be effective in transmitting complex information to the human observer, it must include variation over a large number of independent perceptual dimensions (e.g., Miller, 1956). On the whole and due in part to technical difficulties, most synthetic displays that have been tested are impoverished compared with the natural displays (e.g., the face) considered in the previous section.

Note also that all synthetic sensory-substitution systems (except for not-yet-realized systems in which the display consists of visually presented text) require that the subject learn a new perceptual code. Not only does this learning (or training) requirement constitute a fundamental limitation on the clinical applicability of such systems, but it also greatly complicates and slows down the research. In order to fully explore a given system, it is necessary to develop a portable, wearable version and to examine performance with the system over a substantial time period (Watson, 1980).

1. Visual

Visual spectral displays of speech have long been studied in connection with visible speech and the reading of speech spectrograms. Tests with the visible speech translator (Potter et al., 1947; Stark et al., 1968; House et al., 1968; Houde, 1980), a real-time sound spectrographic display, showed that three adults (two normal-hearing and one deaf) were able to acquire a vocabulary of 300-800 words after 100-200 h of training. Later studies, however, indicated that this type of system might be more appropriate as a speech training aid or as a supplemental aid (Kopp and Kopp, 1963; House et al., 1968; Houde, 1980, 1982) rather than as a self-sufficient substitute for hearing.

Recent research on the reading of speech spectrograms (Zue and Cole, 1979a, 1979b; Cole et al., 1979, 1980; Zue, 1981; Greene et al., 1984) indicates that subjects can read speech spectrograms with high accuracy given sufficient training and no time pressure during the reading test. In the study by Greene et al. (1984), naive normal-hearing observers required roughly 20 h to learn to identify 50 monosyllabic words with 100% accuracy when each word was represented by a single utterance and displayed for 6 s. Furthermore, less than 1 h of additional training was required to obtain scores in excess of 90% when the same words were represented by a new set of utterances from either the same speaker or two new speakers. When the three talkers were combined in a mixed test, performance (without additional training) dropped to 81%. Finally, the amount of training was inadequate to identify new words; the score with novel words dropped to 6% (although one-third of the phonemes in the new word list were identified correctly). As in most other investigations concerned with reading speech spectrograms, the dependence of performance on temporal factors (stimulus duration, response time) was not studied. The durations of the visual stimuli were roughly a factor of 10 greater than the durations of the natural acoustic stimuli.
Variations of the straightforward spectrographic approach, directly concerned with impaired listeners but oriented toward speech-production training rather than speech reception, have also been studied (e.g., see the reviews by Pickett, 1968, 1969; Levitt, 1973, 1985). In accordance with the speech-training goal, these systems tend to incorporate processing algorithms and displays that permit the observer to focus on selected characteristics of the speech signal, and they do not operate in real time.

In a project designed to explore the use of speech reception of articulatory features complementary to those obtainable through residual low-frequency hearing, a real-time visual display of place and manner information (hand extracted from the speech waveforms) was tested with and without low-frequency auditory information (Goldberg, 1972). Reception of speech segments for both normal and hearing-impaired listeners was found to improve with the visual display (e.g., recognition of 48 consonant-vowel-consonant [CVC] stimuli changed from 40% to 75% when the visual display was added). Significant transfer of learning (one of the main goals of the study) was also observed. How the results would have been affected if lipreading had been included is unknown.

Two speech-feature-extraction systems that were designed for practical use in speech reception are the Upton Eyeglass and the Autocuer. The Upton Eyeglass (Upton, 1968) contains a signal processing scheme (analyzer) and a set of miniature lights (display) that are designed to convey to the user characteristics of the incoming sound, such as voicing or friction, that are difficult to lipread. The analyzer consists of filters and logic circuits that classify the sound, and the display consists of a light-emitting diode array in the form of a block figure 8 that is superimposed (by the use of a mirror) on the face of the talker. Although little systematic evaluation of this device has been completed, some results are available (Pickett et al., 1974; Gengel, 1976, 1982). An improvement of roughly 16% over lipreading alone was exhibited for monosyllabic words by the poorer lipreaders in a 13% when it was combined with both lipreading and amplification (Gengel, 1982). These improvements are similar to those achieved by Upton himself with the same display, but smaller than those achieved (25%–30%) with a modified display that employs color as well as spatial coding.

The Autocuer (Cornett et al., 1982) is similar in that it too makes use of LEDs (miniature lights) and an optical system mounted on eyeglasses to project near the mouth of the talker, an image that displays the results of a classification analysis of the incoming speech sound. Two important differences, however, are the modern technology incorporated into the autocuer (i.e., microcomputer, integrated circuits) and the use of cued speech (which is known to be successful) as a starting point for the system design. Unfortunately, no adequate evaluation of the system has yet been performed. Pilot studies indicate that significant improvements in lipreading may be possible; however, the reliability of the automatic speech-feature-extraction system, the ability to understand continuous discourse at normal rates with the display even without any processing errors, and the influence of processing errors on speech-reception performance all await careful quantitative determination.

2. Tactual

General reviews of research on the use of the tactual sense as a substitute for hearing are available in Kirman (1973), Reed et al. (1982), and Sherrick (1984). In essentially all cases, the tactual displays are designed to be applied to the skin and to stimulate cutaneous receptors. (Throughout the rest of this discussion, we refer to such displays as tactile). The possibility of developing synthetic tactual systems that make use of kinesthetic receptors (those that provide information about relative movement of parts of the body) for communication purposes is only now being explored (Loomis and Lederman, 1986). The tactual displays are usually based on homogeneous arrays of vibrators or electrocutaneous stimulato's. The array may consist of only a single element or include scores of elements arranged in a rectangular matrix. Body sites to which tactual displays have been applied include the fingertip, hand, wrist, forearm, collarbone, thigh, stomach, and recently the pinna. Different body sites have different spatial resolution and different-sized cortical representations. The extent to which different sites can be equated by appropriate scaling of element size and interelement distances in the array is not yet clear. For example, the effect of body site on temporal resolution (of great importance for speech reception) has not yet been determined.

Also poorly understood but potentially of great importance is the developmental plasticity of the tactile portions of the central nervous system. It may be important to provide the substitute device or aid early in life in order that the special neural circuitry required for optimal use of the device can develop.

As with the visual sense, substantial tactile research has been conducted on spectral displays (Reed et al., 1982a). Since the tactile sense is relatively insensitive to the frequency composition of the stimulating waveform, spectral displays in the tactile sense use a frequency-to-place transformation: the outputs of the filters used to achieve the spectral decomposition are applied to different regions of skin. In a very rough sense, this frequency-to-place transformation is similar to that performed by the cochlea in the ear.

a. Spectral displays. Spectral displays of speech stimuli have been examined in a wide variety of studies (see the references cited in Reed et al. (1982a); the experimental survey by Spens (1980); and the work by Greene et al. (1983); Brooks and Frost (1983); Spens (1984); Brooks (1984); Brooks et al. (1985); Craig et al. (1985); Blamey and Clark (1985); Brooks et al. (1986a,b); Potts and Weisenberger (1987); and Weisenberger (1987)). Aside from variations in the body site and in the type of stimulation (mechanical...
versus electrocutaneous), the factors that vary among the different studies include the number and geometric arrangement (linear or planar) of the stimulators, the type of encoding scheme employed (e.g., how signal amplitude is represented), and a variety of parameters related to details of the signal processing (choice of filters, time constants, sampling rates, etc.). In most linear displays, acoustic signal amplitude is coded as intensity of the tactile stimulus. In planar arrays, it is coded either as stimulus intensity or as place along the second dimension (the first dimension being used for frequency). When signal amplitude is coded as stimulus intensity in planar arrays, the second dimension is often used as “time sweep” the pattern of stimulation across the skin to simulate the motion associated with natural tactile exploration.

Evaluations at the segmental level with the tactile display alone suggest that, unlike the results for Tadoma, identification performance for vowels is superior to that for consonants. For example, with the MESA (an electrocutaneous system using 288 planar electrodes), recognition was essentially perfect for a set of 8 vowels, but only 50% for a set of 8 plosives and nasals and 70% for a set of 9 fricatives (Sparks et al., 1978). Studies of word-recognition training with tactile displays indicate a word-acquisition rate (without lipreading) of roughly 1–3 words/h, with profoundly deaf subjects performing less well than artificially deafened subjects. One deaf subject (using a 23-channel vibratory system of Engelmann and Rosov, 1975) scored 80 percent correct on 165 words after 170 hrs of training, while two others (using the 16-channel vibratory system of Brooks, 1984) scored 80 percent correct on 50 words after 25 h. One highly motivated artificially deafened subject learned a tactile vocabulary of 250 words in 81 h (Brooks, 1984; Brooks et al., 1985). This result is similar to the best results obtained for a normal-hearing subject with the visible speech translator (Potter et al., 1947): 300 words after 91 h. Results on the effect of adding the tactile display to lipreading for one artificially deafened subject after 200 h of experience with the display (Brooks, 1984) show an improvement of 69% vs 39% for identification of open-set isolated words (Brooks et al., 1986a), 80% vs 58% for words in open-set sentences presented twice on each trial (Brooks et al., 1986b), and 50 vs 15 wpm on the tracking of continuous discourse (Brooks et al., 1986b). These results are in sharp contrast to some previous results on tracking with artificially deafened subjects (Sparks et al., 1979) that suggest that the improvement provided by an aid (over lipreading alone) disappears after roughly 10–15 h of experience with the tracking task (or, alternatively, when the tracking score approaches 50 wpm). In addition to verifying the general value of certain tactile aids for speech reception, recent work by Potts and Weisenberger (1987) and Weisenberger (1987) indicates that two tactile aids that produce equivalent benefits at the level of speech segment identification do not necessarily produce equivalent benefits at the level of continuous discourse tracking.

Unlike cochlear implants, for which the size, mass, and power consumption per stimulating element or channel are extremely small, in tactile aids (both vibratory and electrocutaneous) they are substantial. Most of the work on spectral displays has employed systems with sufficiently many channels (e.g., more than 10) to rule out applicability as a full-time wearable aid. In terms of clinical applicability, there is an essential need for improved transducers and/or consideration of systems in which the cutaneous nervous system is stimulated more directly (e.g., see the discussion of the transducer problem in Sherrick, 1984). Prior to 1985, only one many-stimulator spectral system even approached applicability as a fully wearable aid (Saunders et al., 1976). Now, however, additional wearable multichannel systems are appearing (e.g., Blamey and Clark, 1985; Cholewiak and Sherrick, 1986; D. Franklin, personal communication, 1986). This trend is likely to accelerate as the efficiency of vibratory transducers increases (Spens, 1985; Franklin, 1985).

b. Speech feature and fundamental frequency displays.
Additional research on tactile displays has been concerned with systems that extract special features of the speech signal and/or that involve few enough channels to make feasible their use as full-time speech-reception aids.

A number of studies have focused on the use of multichannel systems to convey fundamental frequency (Willemain and Lee, 1972; Stratton, 1974; Spens, 1975; Grant, 1980; Boothroyd, 1985; Boothroyd and Hnath, 1986). Single-channel displays of fundamental frequency have also been considered (e.g., Rothenberg and Molitor, 1979; Boothroyd and Hnath, 1986); however, the poor frequency resolution of the skin strongly limits the amount of information that can be conveyed in this manner. (We believe that the results of Boothroyd and Hnath, which show no advantage for their multichannel system over their single-channel system, are atypical.) Further studies have emphasized displays in which two or three channels are used to display the energy in distinct frequency bands and/or articularatory information derived from direct pick-ups on the talker (Scott and De Filippo, 1976; De Filippo and Scott, 1978; Scott et al., 1982; De Filippo, 1984). A novel feature of these studies is the use in some cases of a combination of vibratory and electrocutaneous stimulation to increase the richness of the display. In general, the scores obtained with the displays and lipreading combined are significantly better than the scores obtained with lipreading alone. (Most of these experimental aids were tested using unimpaired persons as subjects.)

Studies have also been conducted with single-channel systems, a few of which are now commercially available as wearable aids. Although some single-channel systems display the audio signal directly to the skin, most of the modern ones take account of the poor sensitivity of the skin at the higher frequencies and display a low-frequency tone of constant frequency with an amplitude that is modulated by the audio envelope (e.g., Beguesse, 1976; Spens and Plant, 1983; Franklin, 1984). This processing scheme destroys detailed spectral information, but ensures that the time-intensity contour and prosodic characteristics of the speech signal can be perceived.

A variety of case studies describing performance with single-channel vibrotactile aids are now available (Plant, 1979; Sheehy and Hansen, 1983; Friel-Patti and Roeser,
Goldstein and Proctor, 1985; Geers, 1986). Of particular interest are reports concerning the effects of single-channel aids on the development of speech and language in prelingually deaf children. Proctor and Goldstein (1983) and Geers (1986) each introduced such an aid to a prelingually deaf child (approximately 3 years of age) and reported increased spontaneous vocalization and rapid growth of receptive vocabulary. Results on two slightly older children (Goldstein and Proctor, 1985) indicated significant improvements on a test of auditory language comprehension (Carrow, 1973) after 11 months of experience with the aid. Friel-Patti and Roeser (1983) evaluated the effects of a single-channel aid on the communication skills of four prelingually deaf children whose average age was 4 years. The amount of vocalization plus signing initiated by the children increased over the period and the aid was worn and appeared to decrease following its removal. In general, these results are consistent with the earlier study of Goldstein and Stark (1976), who demonstrated (using a vibrotactile array) an increase in the production of consonant-vowel utterances by prelingually deaf children aged 2 to 4 years who received speech-production training with the tactile aid compared with a control group who received identical training without the aid.

c. Single-channel vibrotactile systems compared with cochlear implants. Further data on single-channel systems, primarily for postlingually deaf adults, is becoming available in connection with cochlear-implant evaluations. Increasingly, investigators are attempting to compare the reception performance achieved with the implant to performance achieved with a tactile aid (e.g., Carney, 1984; Tyler et al., 1984b; Blaney et al., 1985; Miyamoto et al., 1987). In the study by Tyler et al., which compared cochlear implants (both single and multichannel) with a single-channel vibrotactile aid applied to eight normal-hearing subjects, the best results achieved with the tactile aid were roughly equivalent to the best results achieved with any of the implants (averaged over the ten tests performed). Except for certain prosodic tests, however, the worst results obtained with the tactile aid were inferior to the worst obtained with the implants. When the comparison between vibrotactile aid and implant was made on the implanted patients (using a test of syllabic-stress patterns, a male versus female test, and a four-choice spondee test), the results with the tactile aid were inferior. However, whereas the subjects had at least 3 months experience with their implants, they had only 1 h of training with the single-channel tactile aid.

In the study by Blaney et al., which compared the benefits of a hand-held bone vibrator with those of a conventional high-powered hearing aid in postlingually deaf adults who were prospective cochlear implant patients, the results showed roughly equivalent positive effects for closed-set speech tests without lipreading, but no positive effects for open-set tests or tests with lipreading. Subjects who subsequently received implants (multichannel) demonstrated significantly improved performance with the implant, including performance on open-set tests and tests with lipreading. In the study by Carney (1984), the results of comparing a single-channel implant with a single-channel tactile aid applied to normal-hearing subjects showed no dramatic differences, although other differences between the young normal subjects and the implant wearers may have influenced this result. In the study by Miyamoto et al. (1987), concerned with the use of tactile aids in the evaluation procedure for cochlear implant candidacy, a number of wearable tactile aids (one or two channels) were compared with implants (single channel). Although many tests produced roughly equivalent results for the two types of prosthetics, some showed distinctly superior results for the implants. The authors concluded their report by observing that “... the question is not which device, cochlear implant or tactile device is better, but which device is more appropriate for which patient.” An attempt to summarize comparative information on implants and tactile aids (multichannel as well as single channel) has been made by Pickett and McFarland (1985). A discussion of alternatives to cochlear implants that includes comments on ethics, cost, and clinical suitability is available in Martin (1983).

In a thoughtful and clinically relevant overview of results obtained with deaf children, Moog and Geers (1986) examined the benefits of single-channel vibrotactile aids and cochlear implants relative to performance obtained with conventional hearing aids. Based on the study of ten children (ages 2–13) who showed no ability to discriminate sounds on the basis of spectral information and had very poor speech reception through their hearing aids, Moog and Geers concluded that the benefits of the two types of single-channel devices were roughly equivalent, that these benefits were most clearly evident in “getting young children started in learning spoken language” (see also Richardson, 1986), and that substantially improved results would be obtained with multichannel devices that provide spectral information.

d. “Binaural” vibrotactile aid. In another unusual study, Weisenberger et al. (1987) explored the performance of a “binaural” vibrotactile aid in which different vibratory stimuli were applied to the skin of each ear canal by vibrating earmolds. (There is evidence that responses by profoundly deaf individuals to stimuli produced by ordinary hearing aids at very high levels result from excitation of the tactile system—e.g., Nober, 1967; however, in the present study, the system was designed for tactile stimulation at all levels.) It was thought that the use of the same site for the tactile aid as for a conventional hearing aid might increase the acceptability of the tactile device as an aid to hearing and also facilitate the use of hybrid systems involving both auditory and tactile stimulation. Both earmolds in the binaural aid tested provided crude spectral information by vibrating at 80 Hz in response to low-frequency acoustic energy and at 300 Hz in response to high-frequency acoustic energy. Tests on a few normal-hearing and impaired listeners, concerned with sound source localization, environmental sound identification, and syllable rhythm and stress identification, showed mixed results. The performance for the latter two tasks was superior to that for the localization task. It is uncertain, however, whether performance was superior to that which would have been achieved with conventional single-channel
tactile aids (e.g., most such aids permit one to identify common environmental sounds).

e. A vibrotactile "super subject." A case study that is unique in that the tactile aid was worn by a profoundly deaf adult for roughly 13 years has been reported by Cholewiak and Sherrick (1986). The aid employed two channels and was tested using the tracking method (De Filippo and Scott, 1978). The aid improved the subject's lipreading rate from 44 to 58 wpm in Russian, 14 to 26 wpm in English, and 6 to 12 wpm in Hebrew. Although these results apply to only one (rather special) subject, they do demonstrate that a relatively simple tactile aid can increase lipreading rates as well as provide useful environmental information, provided the subject is sufficiently well trained.

E. Clinical treatment by sensory-substitution aids

Although the number of sensory-substitution aids receiving field use is increasing rapidly, the amount of clinical data available is still small. Thus discussion of clinical treatment by such aids must necessarily involve substantial extrapolation. In addition, as in the research area, attention in the clinical area has focused on tactile rather than visual aids.

In general, clinical treatment by a sensory-substitution aid should be considered only in cases of profound bilateral hearing loss. For cases in which the loss is less severe, conventional hearing aids are likely to prove more helpful.

For those with profound bilateral losses, for whom the perception of normally produced speech cannot be significantly enhanced by means of conventional acoustic hearing aids, the choices are either a sensory-substitution aid or a cochlear implant. Although high-powered conventional hearing aids can also lead to above-threshold sensations in such listeners, these sensations appear to result primarily from stimulation of the tactile sense (thus conceptually transforming the aids to inadvertent and inefficient sensory-substitution aids). The option of using no aid at all is generally a poor one since, at the very least, a sensory-substitution aid provides a general awareness of environmental sounds. As noted in the introduction to this report, this discussion is addressed to those who believe that the use of manual methods as the sole means of communication is inadequate and that the awareness of sounds and the understanding of orally produced speech is important. To the extent that these assumptions are rejected by a given deaf individual, the entire discussion may be found irrelevant.

Among the general factors that are relevant to comparisons between sensory-substitution aids and cochlear implants are the following.

(a) The scientific and professional community concerned with sensory-substitution aids has a different character from that concerned with cochlear implants: the former is dominated by academicians and nonmedical clinicians, whereas the latter includes, in addition, a very strong medical component. Many of the factors listed below are directly related to this difference in the developers and advocates of the two types of aids.

(b) The number of patients treated clinically with sensory-substitution aids now appears to be of the order of 1000—not much less than the number treated with cochlear implants. Of the sensory-substitution aids now in the field, almost all stimulate the tactual sense and almost all employ merely one or two channels. In contrast to the patient population for cochlear implants, the majority of people using sensory-substitution aids are children.

(c) The level of funding for research and development (from both government and nongovernment sources) is quite low for sensory-substitution aids. For example, it appears that the total direct costs paid by the National Institutes of Health (NIH) for work on these aids in 1987-1988 was less than one-third that paid by NIH for work on cochlear implants.

(d) Treatment by sensory-substitution aids is much less expensive than treatment by cochlear implants. The former is on the order of 1000 dollars, whereas the latter is likely to exceed 15000 dollars. It should be noted, however, that both types of treatment are increasingly being covered by third-party payments.

(e) Treatment by sensory-substitution aids, unlike treatment by cochlear implants, involves no surgery and is completely reversible. Thus such treatment is potentially applicable to a broader class of patients.

(f) For both sensory-substitution aids and cochlear implants, there is great variation in perceptual performance across subjects. Consequently there is considerable overlap of the distributions of performance for these two types of aid. For both types, there exist subjects who receive essentially no benefit. However, to date, the best performance with cochlear implants is superior to the best performance exhibited with sensory-substitution aids.

(g) The potential performance of sensory-substitution aids relative to cochlear implants depends on individual characteristics of the patients considered, including especially age at onset of deafness, but also other physical and psychological variables.

F. A proposed approach to making decisions about sensory-substitution aids

In general, an ideal clinical path for adults and older children (age greater than 12 years) with profound bilateral losses might be structured as follows.

Initially, the patient should be treated with a sensory-substitution aid on an extended trial basis. After conclusion of this trial period, the patient's experience with the sensory-substitution aid should be compared with expected performance with a cochlear implant. "Expected performance" here means that predicted from available current statistics on cochlear-implant performance and characteristics of the patient. On the basis of this comparison, plus the other major relevant factors (e.g., in the medical and financial domains), the patient should then decide whether to be implanted. If implantation is chosen, the sensory-substitution aid could be abandoned or maintained as a supplement, depending on the additional benefit provided. (This ideal but painstaking approach is rarely taken in actual practice.)

Assistance in the process of selecting a sensory-substitution aid for the initial trial period can be obtained by contacting relevant research groups, clinicians, and manufacturers/distributors of sensory-substitution aids; most individuals in
the medical community are not currently well-informed in this area. Termination of the initial trial period will occur when the implant decision is made or earlier, perhaps, if the patient judges the aid to be worthless. Ideally, the trial period should be continued as long as the increased exposure and associated opportunities to learn the aid are leading to improved performance. In practice, however, a period of 3–6 months may be more appropriate. During this period, tests can be performed with this aid that are identical to tests used for evaluation of cochlear implants so that direct comparisons can be made, and so that the patient can understand the significance of the cochlear-implant tests. In all cases, the trial period should include a training and testing program designed by an appropriately experienced therapist.

The ideal clinical path for younger children (age less than 12 years) could be similar to that described above for adults, but would require more parental input. Before this ideal path can be specified with confidence, however, additional data on the performance of both cochlear implants and sensory-substitution aids in these younger patients is required.

Commercial firms that are producing wearable tactile aids are listed in the Appendix.

IV. CONCLUSIONS AND RECOMMENDATIONS

The questions the working group was asked to address were, most simply, "What speech-perception aid is best for whom, and what help can one expect from the most appropriate aid?" Individual chapters have treated in considerable detail each of the three major types of aids. This chapter focuses on the classes of impaired listeners for whom the various aids appear to be most appropriate, and on the kinds of help each class of listeners may expect. In a final section, the working group has identified a number of questions, both basic and applied, that need to be answered before more successful speech-perception aids can be developed.

A. Who is a candidate for what aid?

1. Conventional electroacoustic aids versus cochlear implants and sensory-substitution aids

This question deserves a data-based answer, at least a statistical or probabilistic one. There are, however, few studies that have systematically tried more than one type of aid on individual patients. Differences in the ability to identify a talker's intended message are often great among individuals with audiometrically identical hearing losses. It is often difficult to determine whether good performance reflects a superior aid or a superior listener. The data that are available suggest that a significant amount of the variation in performance is a function of the individual rather than of the aid or of the care with which the aid is fitted or adjusted. Nevertheless, the single variable that most strongly determines a listener's ability to perceive speech accurately through an electroacoustic hearing aid or, correspondingly, that establishes the need to consider other types of aids (cochlear implants or sensory-substitution aids) is the portion of the speech spectrum (at the output of the acoustic aid) that the listener can successfully process. This variable, in turn, is determined primarily by the degree and type of hearing loss. (In this chapter all references to hearing loss refer to sensorineural loss. Conductive hearing loss can generally be treated successfully either through surgery, medication, or by sufficient acoustic amplification to compensate for the loss.)

The working group concludes that individuals whose speech-frequency thresholds show mild-to-severe hearing loss (25–90 dB) will receive more usable speech-waveform detail through a conventional acoustic aid than from either of the two classes of aid that transform sound into nonacoustic stimulation. We stress that this means that the appropriate aid for the overwhelming majority of hearing impaired persons is the conventional electroacoustic aid, and that will continue to be true unless very substantial improvements are made in the other types of aids.

For those with speech-frequency losses in excess of 115 dB, the working group concludes that such people will derive little help from acoustic amplification. (Note, however, that a conventional aid, if visible to the talker, may cause the talker to speak distinctly in full view of the impaired listener and thereby to facilitate lipreading. Such modifications in the behavior of the talker can provide substantial improvements in communication.) The person who receives no significant help from acoustic amplification is clearly a candidate for a cochlear implant or for one of the sensory-substitution aids. By "no significant help" is meant that electroacoustic amplification fails even to usefully enhance the speechreading (lip reading) of normal conversation. Recommendations for individuals with losses in the region 90–115 dB, a region in which the listener may receive small but significant benefits from acoustic amplification, are, as one would expect, more controversial and are highly dependent on the needs and expectations of the individual impaired listener and on the person making the recommendation.

2. Cochlear implants versus sensory-substitution aids

As indicated above, there is general agreement about the limitations of conventional electroacoustic aids for cases of profound hearing loss. There is not such general agreement, however, on the relative merits of the different possible treatments for such cases. In addition to the fundamental choice between oral and manual communication (see the comment in the Introduction), there are many areas of disagreement concerning the relative merits of cochlear implants and sensory-substitution aids within the oral approach. Unfortunately, too few patients have been systematically trained and tested with more than one class of aid to make available the controlled data base required to address this question directly. Generally speaking, physicians and scientists participating in implant programs consider individuals with profound losses to be candidates for implants, barring some complicating physical or psychological condition. However, other scientists, particularly those whose experience includes research on sensory-substitution aids, are not convinced that the implant should be the only form of aid recommended to the profoundly deaf. As noted in the Introduction, some groups have considered superior performance after brief exposure with a vibrotactile aid an indicator that the patient is...
capable of coping with a new code and therefore should be implanted, whereas others have asked why the same observation might not be seen as an indicator that work with the vibrotactile aid should be continued. Some of the issues relevant to this argument are discussed in the following paragraphs. Unless stated otherwise, it is assumed that the patient is a postlingually deaf adult.

a. Objective performance. A few cochlear implant users (less than 5%-10%, depending on the implant type) have been remarkably successful, even to the point of being able to conduct communicatively useful telephone conversations capable of coping with a new code and therefore should be implanted, whereas others have asked why the same observation might not be seen as an indicator that work with the vibrotactile aid should be continued. Some of the issues relevant to this argument are discussed in the following paragraphs. Unless stated otherwise, it is assumed that the patient is a postlingually deaf adult.

b. Subjective experience and patient acceptance. There is no doubt, on the basis of the reports of postlingually deafened individuals, that electrical stimulation of the cochlea yields auditory experience, i.e., the sensation of sound. (This result is consistent with Johannes Müller's 1826 "law of specific nerve energies," which states that sensory quality is specific to the nerve that is stimulated rather than to the form of energy taken by the stimulus. ) It is also clear that for postlingually deafened individuals who have had only limited training and experience with sensory-substitution aids, stimulation of the visual or tactile sense by means of such aids does not lead to a sensation of sound. This difference in the two classes of aids can obviously play a major role in determining the relative degree of user satisfaction: quite apart from objective performance, the cochlear implant can return a sense of hearing to the postlingually deafened patient.

The percentage of satisfied, often enthusiastic, users reported by implant teams is impressive. This finding is so common in groups that have implanted large numbers of postlingually deafened adults that there seems little likelihood that it can be explained entirely as experimenter bias on the part of the implant team. Moreover, it is easy to understand how such satisfaction exerts a powerful influence on physicians attempting to advise individual patients. Although satisfaction may be influenced by the patient's perception of the implant team's enthusiasm or belief that such an expensive procedure must necessarily be of considerable value, and although these psychological variables are very difficult to assess, satisfaction has clearly been an important factor in the success of implant programs. Independent of the underlying causes, comparable user satisfaction has not been reported for sensory-substitution aids.

c. Deaf children and prelingually deafened adults. Many of the above comments must be modified or supplemented when considering prelingually or postlingually deafened children or prelingually deafened adults. We list below the special factors to be considered in these cases.

(a) Less is known about the relative effectiveness of cochlear implants and sensory-substitution aids for children than for postlingually deafened adults. This is due in part to the smaller number of children implanted and in part to the increased difficulty in the interpretation of performance data obtained from children during years of rapid physical and intellectual change.

(b) Special caution is required in treating children with procedures that are irreversible. Not only is it more difficult to assess the degree of deafness in children, but also the changes associated with children's growth, as well as the rapid evolution of aid technology, favor the use of such treatments as extracochlear implants (Fourcin et al., 1979; Douek et al., 1983) and sensory-substitution aids, which minimize damage to the system. In this connection, it is noted that over 75% of the tactile aids sold by the leading manufacturer of such aids are used by children.

(c) It is important that children be provided with the most effective aid available during critical periods of learning and neural development associated with speech production and speech understanding. Although it is generally agreed that the aid should have sufficient resolving power to enable the user to distinguish and identify the various speech sounds, there is uncertainty about the importance of stimulating the auditory system as opposed to other sensory modalities during these critical periods. For example, although some animal studies are consistent with the notion that stimulation of the auditory system is essential, there exist adult humans (e.g., Tadoma users) who have substantial communication skills and a high degree of linguistic competence despite the complete absence of both auditory and visual stimulation beyond the first 18 months of life. Clearly, this important issue will be clarified as data become available on the results of applying both cochlear implants and sensory-substitution aids to very young children and infants. It remains to be seen whether early application of sensory-substi-
tution aids as well as implants can lead to a true replacement for hearing (in the sense described near the end of Sec. III).

(d) The question of whether to use intracochlear implants in young children cannot be answered unequivocally. The most positive view is that this is the single treatment most likely to result in successful oral-aural speech production and perception. The most negative view is that implants are unjustified when there are safer, much less expensive alternatives available, and no clear evidence has yet proven these alternatives to be less effective aids to speech development. Working group members with clinical or research experience with cochlear implants generally support the FDA decision to permit small-scale studies of children implanted with intracochlear implants.

(e) To the extent that a given profoundly deafened adult can be regarded as having no significant previous auditory experience, the advantage of treatment by cochlear implant is greatly reduced. Not only may such an individual be totally uninterested in a speech-communication aid of any kind (because of an allegiance to the world of manual communication), but since no sense of hearing was ever present, the notion of restoring this form of sensory experience has no meaning. In other words, one of the main advantages of cochlear implants over sensory-substitution aids for postlingually deafened adults may not apply to prelingually deafened adults.

(f) Patients who are doing well with a tactile aid and continue to show improvements should probably not be implanted until a sufficient trial is completed. There is no clear reason to implant a patient who achieves successful communication with a tactile aid.

(g) Specific recommendations for elderly patients are contained in the recent report of the CHABA Working Group on Speech Understanding and the Aging (1988).

d. The research-treatment context. Although laboratory studies indicate overlap between performance with cochlear implants and performance with vibrotactile aids, and although vibrotactile aids cost much less (by roughly a factor of 10) and involve no surgery, they have not been commercially available until the last few years. Consequently, the pool of patients with substantial experience is much smaller for vibrotactile aids than for the cochlear implants. Numerous reasons have been proposed to explain the rapid growth of cochlear implants versus alternative devices. One explanation that is frequently proposed is that cochlear implants were developed within the medical community as a means of helping patients for whom no other currently available treatment was successful, whereas sensory-substitution aids have been mainly a product of academic research, the strongest orientation of which is toward the discovery of general scientific principles. Research scientists have neither the credentials required to test new devices on patients nor, apparently, very strong inclinations to collaborate with others in doing so before all possible shortcomings have been eliminated. Had this been the orientation of the medical community, it seems likely that no patients would yet have been implanted. Differences in the goals and strategies of the clinician versus those of the scientist are difficult to document in any objective manner. The working group is convinced, however, that those differences, possibly as strongly as differences in effectiveness, have contributed to the rapid increase in the number of patients implanted and the relatively sluggish progress of alternative solutions, especially vibrotactile and visual aids.

B. Expectations and hopes for the future

1. Improved devices

The working group believes that, for each class of aids that has been considered, substantial improvements in speech-reception performance can be achieved by improving the design of the aids. Focusing first on conventional acoustic aids, for example, we note the intense current research and development activity in the area of adaptive background noise reduction. Based on preliminary results in this area, we believe it highly probable that hearing aids with significantly improved noise reduction will be commercially available in the near future. Furthermore, since these noise-reduction schemes are essentially independent of the form of stimulation, once it has been demonstrated that they are useful for hearing aids, they undoubtedly will be incorporated into cochlear implants and sensory-substitution aids as well.

Another major thrust, relevant to cochlear implants and sensory-substitution aids, concerns the development and evaluation of multichannel systems. Theoretical considerations, as well as preliminary experimental results, indicate that aid performance can be substantially improved by partitioning the frequency spectrum into a number of frequency bands and stimulating distinct regions (in the cochlea or on the skin) with the signals derived from the various frequency channels. It is unlikely, however, that the following prediction made by Parkins and Anderson in 1983 (a prediction stated for cochlear implants, but also relevant to sensory-substitution aids) will be completely satisfied unless a significant increment in research effort occurs in the next few years (Parkins and Anderson, 1983:530):

We believe that it is possible to develop a cochlear prosthesis that will provide excellent speech discrimination for a significant number of implanted patients without use of visual cues and that such a device will be available in the next ten years.

In addition to multichannel systems based on frequency decomposition, multichannel systems based on other parameters are being considered for systems that stimulate the tactile/kinesthetic senses (inspired by the results on Tadoma, as described in Sec. III). In these systems, not only are the speech signals decomposed into distinct channels on bases other than frequency content, but the signals are also presented using displays that are perceptually richer than the homogeneous vibrator arrays normally used with the frequency-decomposition systems. Whether practical systems of this type can be developed for clinical application and the extent to which such practical systems actually lead to improved speech reception remains to be seen.

Research is also being accelerated on the use of automatic speech-recognition systems for speech-reception aids. If one regards the aid primarily as a supplement to lipread-
ing, and therefore designs the speech-recognition system merely to disambiguate visually confusable speech, then the automatic speech-recognition problem is vastly simplified. In principle, the output of the speech-recognition system can be coded for the visual system (as in the Autocuer system mentioned in Sec. III), for the tactile/kinesthetic system, or for the auditory system. We expect that speech-reception aids that incorporate advanced automatic speech-recognition systems as part of the preprocessing, like those that incorporate adaptive noise-reduction systems, will be commercially available within the next 5–10 years.

2. Improved knowledge of human capabilities

Advances in technology promise the availability of computerlike hearing aids (and cochlear implants and sensory-substitution aids) in the very near future. However, despite this ability to build almost any form of processor, the optimal forms to build remain very poorly defined. This problem, which exists whether the final transduction ends in acoustic, electrical, or mechanical waveforms, results primarily from our inadequate knowledge of human sensory and cognitive capabilities. Although there are undoubtedly many areas in which increased knowledge would be useful, it is clear that this problem is greatly complicated by our lack of knowledge concerning (1) the limits of human abilities to learn alternatives to the natural speech code and (2) the neurodevelopmental consequences of long-term use of such alternate codes. Clarification of these issues requires a major research program in which there are systematic studies of perceptual and neurophysiological development for a variety of types of aids applied to a variety of types of patients over a variety of time periods at a variety of ages. The systems studied should include hearing aids that involve major transformations of the speech waveform (e.g., by frequency lowering) and extracochlear electrical stimulation aids, as well as cochlear implants and sensory-substitution aids; the individuals tested should include cases of prelingual as well as postlingual deafness in both adults and children; and the course of perceptual learning should be tracked over periods of years. Through such a program these fundamental issues can be well enough understood to design optimal aids for each of the several classes of hearing-impaired listeners.

C. Overall conclusions

The overall conclusions of the working group are threefold. First, substantial advances have been made both in the development of improved aids and in the scientific foundations of aid development. Much has been learned about the possibilities and limits of acoustic amplification, about the phenomena associated with electrical stimulation of the cochlea, and about the potential of other senses to convey speech information. The aids now available offer not only a great variety of acoustic amplification systems (in very small, cosmetically acceptable packages), but also, through electrical stimulation of the cochlea or through stimulation of other senses, the possibility of significant help for cases of profound deafness.

Second, it is important that the levels of speech communication that can be expected with the current devices (both the improvements and the limitations on these improvements) be fully appreciated. This is particularly true in the case of both implants and sensory-substitution aids for individuals with postlingual profound deafness. These persons must be clearly informed that the most likely positive outcome of their use of these aids is that they will experience less difficulty in lipreading. Through the use of a well-fitted, properly adjusted, conventional acoustic hearing aid, persons whose losses range from clinically minimal to moderate (losses of 30–50 dB) can expect to understand speech nearly as well as do persons with normal hearing, at least in the absence of high levels of extraneous noise. Persons who use traditional aids and who have greater loss must expect less success in understanding speech, with performance gradually reducing to no significant enhancement of lipreading as losses reach 95–110 dB. Performance with lipreading alone can vary from almost perfect ability to hold a conversation (so long as there are no abrupt changes of topic or talker) to recognition of no more than an occasional word, depending on training, motivation, and, apparently, innate aptitude. For those whose hearing loss is in the range for which traditional acoustic aids may yield no significant help to lipreading (95–110 dB), and definitely for those whose loss exceeds this range, one of the alternative aids will probably provide some useful increment to lipreading alone. Most major cochlear-implant groups are, to the best of our knowledge, making serious efforts to see that patients’ expectations are realistic. Nevertheless, it is emphasized that great care should be taken to help prospective patients understand the difference between an aid to lipreading and a neural prosthesis that will return their hearing. While the performance reported for a few stars would seem to justify the latter description, the more modest expectation remains valid for the majority of patients. Also troublesome, is that star performance on speech-recognition tasks does not appear to be predictable from other measures of auditory processing.

Exactly what help the average patient can expect from sensory-substitution devices is even less clear than for cochlear implants, in part because they have not been commercially available until the past few years. In laboratory studies, these devices have yielded increments in speech reception above that for lipreading, similar to those achieved by average cochlear-implant users. In addition, as stated previously, these devices are much less expensive (by as much as a factor of 10) and involve no surgery. No individual, however, has demonstrated star performance with such a device.

The working group supports the conclusion that each class of aids has demonstrated a sufficiently viable clinical potential that additional research on all classes of aids is justified. We feel that the membership of the working group is sufficiently representative that this report accurately reflects both the consensus and the disagreements that might be discovered by polling most of those currently active in research on aids to speech communication. We offer this report as a balanced statement of the current scientific understanding of a very complex problem.
The third and final conclusion is that there are many reasons to have high hopes for the future. Current research results on multichannel systems and on noise-reduction schemes (as well as on automatic speech recognition), together with the demonstration that the tactile/kinesthetic senses are adequate for receiving speech provided the user is thoroughly trained at a sufficiently early age and the reports of star implant users, all suggest that major improvements are within our grasp. If the research effort on this problem continues at its present rate, the next decade will yield aids that are substantially more effective than those now available. This is an important conclusion, because for many hearing-impaired persons only a modest increment in the current effectiveness of aids would be required for them to become socially adequate communicators (in the aural-oral mode). Unlike Parkins and Anderson, however, we do not feel confident enough to predict a truly successful auditory prosthesis for the profoundly deaf in a decade.

ACKNOWLEDGMENTS

Partial support for preparation of this report was provided by NIH Grant DC 00250 from the National Institute for Deafness and Other Communication Disorders to Indiana University. We wish to thank Milton Whitcomb, Study Director of CHABA, for helpful discussions and Deborah Giddens for assisting in the preparation of the report.

APPENDIX: COMMERCIAL AVAILABILITY

This Appendix lists the major characteristics of five of the most widely used implant devices in the United States as well as commercial firms that are producing wearable tactile aids, in 1986–1987.

1. Implant devices

Symbion/Ineraid: Four-electrode analog bandpass filtered, simultaneous stimulation, monopolar intracochlear scala tympani (22 mm), percutaneous connection to electrodes.

Nucleus: 21-channel digital sequential stimulation with speech feature extraction coding (stimulation rate is determined by voicing rate). The peak energy is extracted in 10-ms epochs and used to determine which of the 21 electrode pairs will be stimulated (according to a predetermined "low to high" electrode, pitch ranking); a single electrode is stimulated at a time. Intracochlear scala tympani (25 mm), demodulated, implanted electronics.

3M House: Single channel amplitude-modulated 16-kHz carrier, passive implanted electronics without demodulation analog stimulation limited to 340–2700 Hz, intracochlear (6 mm). (No longer available in 1991, but service is provided.)

3M Vienna: Single electrode analog, with wideband equalization, implanted electronics with carrier demodulation, round window implant site.

UCSF/Storz: Four-electrode analog band-pass filtered, bipolar stimulation, scala tympani insertion (22–25 mm), active implanted electronics. (Note: in 1988 this device was not in production but may be manufactured by another company.)

Table AI provides summary information about implant devices; more detailed statements from manufacturers have been published in ASHA, 27: 27-34, 1985.

2. Tactile aids

Commercial firms that are producing wearable, portable, tactile aids include: AB Special Instruments (Minivib 3), Stockholm, Sweden; Audiological Engineering (Tactaid), Cambridge, Mass.; Coulter Associates (Portapitch), Washington D.C.; Siemens Hearing Instruments (Minifonator) Union, N.J.; Tacticon Corp. (Tacticon), San Rafael, Calif.; Telex Communications (Televibe), Minneapolis, Minn. We do not know of any wearable, portable, visual aid that is available commercially, although this situation may change in the not-too-distant future with commercial introduction of the Autocuer (R. Beadles, personal communication, 1986).

Although the primary subject of this section is the conventional hearing aid, defined as an acoustic amplifier with variable gain and frequency response, several other types of aid are also discussed. In general, we have attempted to include here all types of aids other than electrical cochlear stimulation and sensory-substitution aids.

Further detail on consonant (C) and vowel (V) confusions by cochlear implant subjects can be found in the following original sources: 3M House (C/V: Edgerton et al., 1983a; C: Edgerton et al., 1983b; C/V: Tyler et al., 1985); 3M Vienna (C: Hochmair-Desoyer et al., 1980; V: Hochmair-Desoyer et al., 1981; C: Hochmair and Hochmair-Desoyer, 1983; C/V: Tyler et al., 1985); Nucleus (C: Dowell et al., 1982; C/V: Tyler et al., 1984b; C: Dowell et al., 1985).


Speech-perception aids


Stevens, S. S. (1975). Psychophysics: Introduction to its Perceptual, Neu-
WORKING GROUP ON COMMUNICATION AIDS FOR THE HEARING-IMPAIRED

CHARLES S. WATSON (Chair), Department of Speech and Hearing Sciences, University of Indiana
ROBERT A. DOBIE, Department of Otolaryngology, University of Washington
NATHANIEL DURLACH, Research Laboratory of Electronics, Massachusetts Institute of Technology
LARRY E. HUMES, Department of Speech and Hearing Sciences, University of Indiana
HARRY LEVITT, Program in Speech, City University of New York
JAMES D. MILLER, Central Institute for the Deaf, St. Louis
CARL E. SHERRICK, Department of Psychology, Princeton University
F. BLAIR SIMMONS, Department of Otolaryngology, Stanford University Medical Center
GERALD A. STUDEBAKER, Memphis Speech and Hearing Center, Memphis State University
RICHARD S. TYLER, Department of Speech and Hearing, University of Iowa
GREGORY P. WIDIN, 3M, St. Paul, Minnesota

COMMITTEE ON HEARING, BIOACOUSTICS, AND BIOMECHANICS (CHABA)

JOHN H. MILLS (Chair), Department of Otolaryngology, Medical University of South Carolina
ROBYN COX, Speech and Hearing Center, Memphis State University
ROBERT A. DOBIE, Department of Otolaryngology, University of Washington
WILLIAM J. GALLOWAY, Tarzana, California
C. DANIEL GEISLER, Department of Neurophysiology, University of Wisconsin
IRA J. HIRSH, Central Institute for the Deaf, St. Louis
LARRY E. HUMES, Department of Speech and Hearing Sciences, University of Indiana
DENNIS McFADDEN, Department of Psychology, University of Texas
ROBERT D. SORKIN, Department of Psychology, University of Florida
CONSTANTINE E. TRAHIOTIS, Department of Otolaryngology, University of Connecticut Health Center
HENNING E. VON GIERKE, Yellow Springs, Ohio
MILTON A. WHITCOMB, Study Director
ARYLSS K. WIGGINS, Administrative Secretary