Adapting to Changed Hearing: The Potential Role of Formal Training

DOI: 10.3766/jaaa.21.9.6

Arthur Boothroyd*

Abstract

Changed hearing occurs when sensorineural loss is acquired or increases, when hearing aids or cochlear implants are first acquired, when hearing aids are reprogrammed, and when cochlear implants are remapped. The changes affect speech perception—a process in which decisions about a talker’s language output are made on the basis of sensory and contextual evidence, using knowledge and skill. The importance of spoken communication dictates speedy and optimal adaptation to changed hearing. Adaptation is a process in which the individual acquires new knowledge and modifies skill. Formal training provides the listener with the opportunity to enhance both knowledge and skill by spending time on speech perception tasks without the demands, constraints, uncertainties, and risks associated with everyday communication. Benefits of such training have been demonstrated in terms of improvement on trained tasks and talkers, generalization to untrained tasks and talkers, improvements in self-perceived competence, and reduction of self-perceived handicap. So far, however, we lack information on which aspects of training are responsible for benefit, which aspects of perception are changed, how individual differences interact with the foregoing, and whether these benefits translate into significantly increased participation and quality of life.

Key Word: Auditory rehabilitation

Abbreviations: CAST = Computer-Assisted Speech Training; LACE = Listening and Communication Enhancement

INTRODUCTION

For present purposes, changed hearing refers to changes accompanying the acquisition or increase of sensorineural hearing loss, the acquisition or reprogramming of hearing aids, and the acquisition or remapping of cochlear implants. Changed hearing can have two main effects. First is a change in the relationship between sound stimuli (what goes into the ear) and sound sensations (what the listener experiences). Second is a change in the amount of information made available by more peripheral auditory mechanisms to more central auditory and linguistic decision-making mechanisms. The immediate change in information can be negative (when a hearing loss is acquired, for example), or positive (when hearing aids or cochlear implants are acquired). Even in the second case, however, the change from preloss to postfitting will be negative.

For most persons with hearing loss, the primary impact of changed hearing is on speech perception. Because of the importance of spoken language communication in everyday life, adaptation to changed hearing needs to be as effective as possible, and it needs to occur as soon as possible. During this process, one can hypothesize two kinds of learning related to the two kinds of change outlined above. The first is learning new relationships between sound sensations and the language patterns they represent. The second is modifying perceptual skills to deal with reduced information. The goals of this article are:

1. To provide a conceptual model for speech perception as a basis for discussing training and its potential effects.

*Professor Emeritus, Doctoral program in Speech and Hearing Sciences, City University of New York.

Arthur Boothroyd, 2550 Brant Street, San Diego, CA 92101; Phone and Fax: 619-231-7948; E-mail: aboothroyd@cox.net

Preparation funded in part by grant #H133E080006 from the Office of Education, National Institute of Disability Rehabilitation and Research (NIDRR), to the Rehabilitation Engineering Research Center of Gallaudet University.

The contents do not represent the opinions of the Office of Education.

Endorsement by the U.S. Federal Government should not be assumed.
2. To review options for formal training and evidence for its benefits.
3. To outline gaps in current knowledge and opportunities for future research.

Speech Perception

Figure 1 provides a conceptual model of speech perception. Communication begins when a talker’s language patterns (phonemes, words, sentences, intonation, etc.) are converted to speech movement patterns which generate sound patterns. Depending on the observer’s auditory capacity (and, in face-to-face communication, the observer’s visual capacity), some of information in the sound and movement patterns becomes available as sensory evidence. Information about the identity of a language pattern is present in the context. This context is both general (the world, people, and the language) and specific (the immediate environment, the people involved, and the language being used). Depending on the observer’s world, social, and language knowledge (and, to a certain extent, the observer’s sensory capacity), a portion of the external information in the context is converted to internal contextual evidence. Perceptual processing requires the making of perceptual decisions about the most likely origin of the sensory evidence. Options for the language patterns about which decisions are to be made are drawn from knowledge. Contextual evidence helps determine these options and influences the observer’s confidence in the decisions. The effectiveness of perceptual processing requires perceptual skill. Once an option being considered is accepted, it becomes a perceived language pattern. Perceived language provides additional context for subsequent decisions. At this point, success can be defined as correspondence between the observer’s perceived language patterns and the talker’s actual language patterns. This gives us the luxury of quantifying performance in terms of percent correct recognition of such things as phonemes, words, or other language patterns. Of course, true success involves accurate perception of the talker’s meaning and intent, but this topic is beyond the scope of the present article. Moreover, the ability to infer meaning and intent from perceived language patterns is not directly affected by changed hearing.

The model offered here identifies four major contributors to perceptual processing and decision making: sensory evidence, contextual evidence, knowledge, and skill. Sensory evidence is limited by hearing loss and the invisibility of most speech movements, but it can be optimized with hearing aids and cochlear implants, assistive listening devices, and control of the visual and acoustic environment. Contextual evidence can be controlled in the design and implementation of training protocols. But knowledge and skill are what the trainee brings to the decision-making task and are, therefore, potential targets for learning and adaptation.

Of the many aspects of knowledge illustrated in Figure 2, only one is of immediate concern, knowledge of the relationships between language patterns and sensory evidence. Changed hearing may require the observer to learn new or modified relationships. Part of this learning can involve more general adjustment to changed relationships between sound stimuli and world, social, and linguistic context.
sound sensations, usually discussed under the heading of acclimatization (Arlinger et al, 1996; Philibert et al, 2005). When there is no change in the total amount of information available in the sensory evidence, the goal of formal instruction may be limited to increasing the speed and effectiveness with which these new relationships are learned.

This brings us to the topic of speech-perception skill, which involves many factors, as illustrated in Figure 3. Some of these factors are more related to the process. Others are more related to the person. Examples of process-related components include attention, retrieval, decision-making, balancing accuracy and speed, balancing contextual and sensory evidence, and making full use of working (or short-term) memory. The role of attention is to optimize sensory evidence and the allocation of resources to the task at hand. Possible interpretations of the evidence must be retrieved rapidly from knowledge, while using context to select those most likely to be correct. Decision making requires acceptance of a possible interpretation when its estimated likelihood of being correct, after the sensory evidence is taken into account, exceeds some criterion. The decision-making process must balance speed and accuracy. This balance is influenced by that between contextual and sensory evidence (less use of sensory evidence may increase speed but also increase the chance of error). The fact that language and speech patterns are distributed over time requires effective use of working memory to store sound sensations and partially recognized utterances pending arrival of more evidence and/or completion of decision making.

Examples of person-related components include flexibility, risk-tolerance, confidence, and the ability to perform two or more tasks at the same time. Flexibility, as used here, implies the ability to shift the relationships among speed, accuracy, and acceptance criteria according to the demands of the communication situation. Individuals differ in their tolerance for risk of misunderstanding, communication breakdown, and the resulting frustration or embarrassment. They also differ in terms of their confidence in entering a communication situation and belief in their ability to deal with it in spite of their hearing difficulties. And effective communication requires multitasking; in conversation, we need to know what was said, what it meant, why the talker said it, and how we will respond—all at the same time.

Note the possible interactions between process-related and person-related factors. The ability to multitask, for example, is almost certainly related to perceptual speed and working memory, while risk tolerance is probably related to the acceptance criteria for perceptual decisions. It is reasonable to assume that significant improvements in the process-related factors listed here will have a positive influence on person-related factors and vice versa.

The need to balance speed and accuracy is critical to speech perception because speed is controlled not by the observer but by the talker. In this respect, speech perception differs from reading, in which speed is controlled by the reader. Nevertheless, reading competence also requires a balance between contextual and sensory evidence in order to optimize speed while maintaining an acceptably low chance of error. To accomplish this, the competent reader uses a series of saccadic movements in which the eye takes in only enough sensory evidence to confirm context-based expectations (Rayner, 1998). The movement stops momentarily when this process fails, and the reader reaches a point where he needs to take full account of the available sensory evidence. The author believes that the competent speech perceiver does something similar—optimizing speed by using only enough of the available sensory evidence to confirm context-based expectations. Because there is surplus information (redundancy) in the speech signal, effective perception and communication are possible in spite of such things as language short cuts, imperfect articulation, unfamiliar accent, acoustic distortion, reverberation, and masking noise.

Effective perception and communication are also possible in the presence of the hearing changes and loss of information accompanying moderate cochlear damage. This accounts, in part, for the delay between the onset of hearing loss and the search for solutions. Severe
cochlear damage, however, can reduce the available sensory evidence below that needed to keep up with the talker while maintaining an acceptably low chance of error. Some of the auditory evidence can be restored with hearing aids or cochlear implants, and vision provides a valuable complement to hearing in face-to-face communication. But, for many individuals with hearing loss, speech perception remains a challenge even after they have accepted sensory assistance. Moreover, the loss of information caused by the hearing loss makes them less tolerant of the loss of information caused by other sources of interference such as those listed in the previous paragraph. Some improvement may occur as the individual adapts to changed relationships between language patterns and sensory evidence. The residual information loss, however, may also require the individual to modify perceptual strategy so as to maintain effective speech perception in spite of reduced sensory evidence. Formal instruction to address the modification of perceptual strategy probably calls for a different and more extensive protocol than that required for learning new stimulus-sensation relationships. It is not clear, however, how much change of perceptual skill is possible, which aspects of skill might be malleable, and how individual characteristics might affect both.

We do, however, have existence proof of the possibility of maintaining effective speech perception in spite of very limited sensory information. This comes in the form of highly competent lipreaders. It is generally believed that such individuals are unusually effective users of context to compensate for limited sensory evidence—a hypothesis supported by the fact that even the most competent fail when the topic of a conversation changes. It is also possible, however, that these individuals rank highly on the other process-related and person-related factors listed above. At the time of writing, and in spite of a century of research, we lack information on exactly what makes a highly competent lipreader or whether formal training can result in high levels of competence in deafened adults who lack the necessary aptitude.

Some readers may have noted the Bayesian logic underlying the earlier description of decision making. Thomas Bayes was an 18th century cleric and mathematician with an interest in chance and probability (Bellhouse, 2004). Bayesian theory accounts for the fact that recognition of an object or event depends not only on the available sensory evidence but also on the observer’s context-based expectations. The observer starts with a possible interpretation of the evidence (a hypothesis). Based on expectations, the observer already has a sense of the likelihood that this interpretation is correct, regardless of the evidence (an a priori probability). When the evidence is taken into account, the observer’s estimate of this likelihood may increase. If it reaches a high enough value (an acceptance crite-

ration), the interpretation is accepted and the object or event is recognized.

Bayes’s contribution to this topic was to recognize the importance of the observer’s knowledge of the chances of certain causes producing certain effects. He demonstrated how this knowledge can be used in reverse to interpret evidence in terms of its probable cause. This process is embodied in Bayes’s equation (see Appendix A). This equation, applied, for example to word recognition, is a mathematical version of the following statement: “My confidence that what I am hearing comes from the word I expect increases in the same proportion as my belief that this word, rather than some other word, would produce what I am hearing.”

Application of Bayesian theory to speech perception is not intended to suggest that the human brain carries tables of probabilities and probability distributions. Nevertheless, the implications are in keeping with observed behaviors. It has long been known, for example, that when a word is expected, either from context or because it is a common word, it can be recognized at a lower signal-to-noise ratio (Miller et al, 1951; Rosenzweig and Postman, 1957; Broadbent, 1967). Considerable research has been done on “priming,” in which a listener’s expectations are manipulated with predictable results on perceptual outcome. Phenomena such as categorical perception, auditory-visual speech perception (Massaro, 1987), and the McGurk effect (McGurk and MacDonald, 1976) can be modeled using Bayesian statistics, and there has been extensive work on the application of Bayesian modeling to reading (Norris, 2006).

**Formal Training**

A lot of effort has gone into the study of adult learning. Most of this work has focused on the acquisition of new motor skills or explicit knowledge (Knowles et al, 2005). Less attention has been paid to the acquisition of new perceptual skills. One of the principal conclusions from studies of adult learning, however, is applicable to the present discussion. The motivation for learning comes from the learner and is often in response to a specific life-changing event—in the present case, an unacceptable loss of communicative function. Note, also, that adults with acquired hearing loss do not seek formal speech-perception training in order to raise their word recognition scores by a small but statistically significant amount. Their goal is the restoration of effective communication, full participation, and quality of life.

Assuming the learner is motivated, what are the requirements for a program of formal training? This author suggests:

1. The process should be unthreatening. In other words, the negative consequences of failure should be minimized. As an example, the learner should not feel that

2. Group training with a clinician. While group training may involve less tailoring to individual needs, it does reduce cost, and there are potential psychosocial benefits from interactions within a peer group.

3. Computer-assisted self-instruction. Recognizing the barriers of cost and the absence of third-party coverage, several rehabilitation researchers have explored the possibilities of self-instruction. Available options at the time of writing include:

Options for Formal Training

There are several options for formal training. For example:

1. Individual training with a clinician. Advantages are the possibility of tailoring to individual needs, personal counseling in terms of confidence, risk tolerance, and assertiveness, together with coaching on control of the visual and acoustic environment, and on confirmation and repair strategies. There are, however, some drawbacks. Individual training is costly, the total time-on-task is low, and the variety of inputs is minimal.

2. Group training with a clinician. While group training may involve less tailoring to individual needs, it does reduce cost, and there are potential psychosocial benefits from interactions within a peer group.

3. Computer-assisted self-instruction. Recognizing the barriers of cost and the absence of third-party coverage, several rehabilitation researchers have explored the possibilities of self-instruction. Available options at the time of writing include:

- Adaptation to Changed Hearing/Boothroyd
a. Listening and Communication Enhancement (LACE) from Neurotone (http://www.neurotone.com/index.html). This program is one of the most popular, and it meets many of the criteria listed above (Sweetow and Henderson Sabes, 2006). Where the availability or use of a computer is a problem, it has been adapted for use with a DVD player and television set.

b. Sound and Way Beyond is based on the CAST (Computer-Assisted Speech Training) program (Fu and Galvin, 2007) and is available from the Cochlear Corporation (http://www.cochlearamericas.com/Support/2427.asp).

c. eARena is a package developed and available from Siemens through audiologists. It provides both instruction and practice, the latter including not only speech but also environmental sounds (https://hearing.siemens.com/en/04-products/18-earena/01-details/earena-details.jsp).

d. Seeing and Hearing Speech is available from Sensimetrics (http://www.seeingspeech.com/). As the name suggests, this program is not limited to hearing but also provides visual and auditory-visual learning opportunities.

e. The author has developed several programs for Computer-Assisted Speech Perception self-training (CASPER) under a grant to the Rehabilitation Engineering Research Center at Gallaudet University (http://www.hearingresearch.org/). These programs provide auditory, visual, or auditory-visual learning opportunities at the vowel and consonant level (CasperCon) and at the sentence level (CasperSent) (Boothroyd, 2008). Currently under development is an auditory-only program (AudioCasper) that provides training in the context of short stories. In this last program, stories are presented one sentence at a time, and the trainee is asked to repeat each sentence before being shown the text. He or she then self-scores by clicking on the words that were heard correctly. If there were errors, the sentence is heard and repeated again before moving on to the next one. All of the Casper programs provide the option of replacing self-scoring by scoring with the assistance of a clinician or significant other.

f. A recent addition to computer-assisted programs is Read My Quips from Advanced Hearing Concepts (http://www.advancedhearingconcepts.com).

Evidence of Effectiveness

As a contribution to a special issue of the Journal of the American Academy of Audiology on evidence-based practice, Palmer and Sweetow reported on a systematic review of the research on one-on-one auditory training with hearing-impaired adults (Sweetow and Palmer, 2005). In the six studies that met their acceptance criteria, there was evidence of improvement on speech-perception tasks used in training and of carryover to other talkers and tasks. More recently, Humes and colleagues demonstrated performance improvements as a result of training on the recognition of frequently occurring words, with carryover to sentence perception and to a novel talker (Humes et al, 2009).

Hawkins has published a review of group intervention with a focus on counseling and self-perceived benefit. Once again, there were reports of positive outcomes, at least in the short term (Hawkins, 2005). In an interesting study by Chisolv and colleagues, trainees who received group counseling following hearing aid fitting showed somewhat more improvement on some self-assessed communication measures than those who did not (Chisolv et al, 2004). The latter group, however, showed improvements one year later, eliminating the group difference. This appears to be a case where intervention may have speeded adaptation without affecting ultimate performance.

Computer-based training programs have demonstrated positive outcomes. Sweetow and Henderson Sabes, for example, reported on a multicenter crossover study of LACE, involving 65 adult hearing-aid users. Performance improvements, attributable to training, were noted for tasks involving speech in babble, speech with talker competition, speeded speech, short-term memory, and word identification from context only. There were also statistically significant performance improvements on untrained tasks and reductions of self-perceived handicap (Sweetow and Henderson Sabes, 2006).

Fu and colleagues have reported positive outcomes from training with the CAST program, using both cochlear implant recipients (Fu and Galvin, 2007) and hearing participants listening to cochlear implant simulations (Nogaki et al, 2007). Both groups can be considered to be dealing with changes in the relationship between language events and sensory evidence. Less positive were the results of a study by Stacey and colleagues in which 11 adult cochlear implant users trained on a computer-based program for 3 wk, 5 day/wk, 1 hr/day. Although there was a significant improvement in consonant identification, there was no evidence of carryover to word recognition in sentences and little evidence of self-reported benefits (Stacey et al, 2010). Participants in this study, however, had a minimum of 3 yr prior experience with their implants.
so they had already had ample opportunity to learn new relationships between language patterns and sensory evidence.

Data on the use of recorded books for adaptation to changed hearing are not available at the time of writing, but an ambitious observer-blinded study is underway by Chisolm and colleagues in which the potential benefit of this approach is being compared with that of the LACE program as a supplement to hearing-aid fitting (http://www.clinicaltrial.gov/ct2/show/NCT00727337?term=Chisolm&rank=1).

The Importance of Time-on-Task

While it is clear that performance can improve with training, it is seldom clear what aspect of training has the most impact. It is reasonable to assume, however, that time-on-task is important. Results from the work of the author and colleagues have supported this assumption. For the author, the most striking example came from an unpublished study of lipreading training, using an earlier version of the CasperSent software. This study used the CUNY topic-related sentence sets, recorded on video laser disk and presented under computer control via a television monitor without sound. Participants were asked to repeat each sentence, and the words correct were scored by the experimenter. The purpose was to determine the effects of feedback on learning rate. Three groups of five participants were trained. Participants in Group 1 simply tried to repeat each sentence and were given no feedback on performance. Participants in Group 2 tried to repeat each sentence and were then shown the text before moving on to the next sentence. Participants in Group 3 used a protocol that had been developed with the specific aim of facilitating learning. After the first attempt to repeat a sentence, trainees were shown only the correct words, with dashes indicating the words not yet recognized. They saw the sentence spoken again and were given a second or third chance to repeat it. If there were still words unrecognized after the third attempt, they were shown the full text and given a last chance to observe the sentence being spoken before moving on. This approach was intended to enhance the use of context while focusing attention on relationships between language patterns and sensory evidence. An initial plot of the mean percentage of words recognized, shown in Figure 4, appeared to show a higher learning rate for Group 3.

It was quickly realized, however, that Group 3 spent more time on each sentence than did the other two groups. When the growth functions were plotted as a function of time-on-task (see Figure 5), there were no obvious differences of learning rate between the groups. A logical interpretation was that the primary virtue of the complex procedure used with Group 3 lay not in the details of the training protocol but in the more efficient use of recorded speech materials. The ultimate performance of the two groups receiving trial-by-trial feedback was higher, on average, than that of the group receiving no feedback, but within-group variability was too high to support generalization to the means of the population represented by these three samples.

The significance of time-on-task also helped explain the findings of a study of speech-perception training in late-deafened adults receiving multichannel

Figure 4. Words recognized in everyday sentences, by lipreading alone, as a function of sentence set. Data points are means for groups of five normally hearing participants with no previous experience on this task. Each sentence set contained 12 sentences of varying length and topic. Curves are least-squares fits of exponential growth functions to the data.

Figure 5. The upper graph compares the exponential growth functions from Figure 4 and suggests a higher learning rate for Group 3. But when replotted in terms of estimated time-on-task, as in the lower graph, the learning rates for the three groups are very similar.
cochlear implants (Boothroyd, 1987; Boothroyd et al, 1987). Trainees received no formal training for 1 mo after initial activation of the implant. Two 1 mo periods of formal training were then provided (roughly 4 hr/wk), separated by a 1 mo no-treatment period. One training period was at the vowel and consonant level. The other was at the sentence level. The order was counterbalanced across trainees. Testing and training were done by lipreading, hearing, and the two combined. The most striking finding was that the largest improvements occurred during the first month without formal training, presumably as a result of time spent on everyday communication. After that, the only change that could be attributed to formal training was during sentence-level training and was limited to the auditory-alone condition. Our conclusion was that these individuals spent a lot of time on the task of auditory-visual speech perception during everyday communication, and the increase provided by formal training was relatively small. But opportunities for auditory-alone practice were limited and, for this activity, formal training provided a significant addition to total time-on-task.

Another example of the role of time-on-task comes from the study of Nogaki et al, mentioned earlier, in which normally hearing listeners were trained on the perception of noise-vocoder cochlear-implant simulation with shifted spectral envelopes. In this case, it was found that the amount of learning depended only on the number of training sessions and not on the way training sessions were distributed over time (Nogaki et al, 2007). This is not to say, of course, that distributed training is unnecessary. Depending on the complexity of the task and the need for use and retention of new skills, there are obvious benefits to distributing training over several sessions rather than compressing it into one.

The attention given here to time-on-task may seem redundant. It is common knowledge that skills require practice and nonuse leads to nonpossession. What is important in the present context, however, is defining the task. Unlike motor skills, where the nature of the task is self-evident, perception is an internal process. We are all capable of “hearing without listening.” In fact, some of the maladaptive procedures adopted in response to deteriorating hearing may have the false appearance of being on task. Examples include tuning out, smiling and nodding, and nonstop talking (to avoid listening). When spending time on formal training aimed at improving speech perception, it is important to optimize allocation of time to the task outlined in Figure 1, which might be defined as “listening with intent to understand.”

To summarize, there is ample evidence that formal training can enhance speech perception performance, regardless of the context of its provision. In some studies, the performance improvement is demonstrated only for the talker and materials used in training. There are data, however, to support the conclusion that performance benefits can generalize to other talkers and other materials. These benefits may well depend on the extent to which the addition of formal training to the demands of everyday communication increases time-on-task by a significant amount. There is also some evidence to show that training combined with counseling and instruction can increase self-perceived competence and reduce self-perceived handicap.

Research Needs

Although the research literature provides evidence of improved performance following formal training, several important questions are unanswered:

1. When performance improvements are attributable to training, what aspect of speech perception has changed? Are the benefits attributable only to learning new relationships between language patterns and sensory evidence, or has there been a change in perceptual strategy? If the latter, are there effects on such things as selective attention, retrieval, processing speed, acceptance criteria, or the use of context? How much of the benefit is attributable to such things as increased confidence and greater risk tolerance? Answers to these questions will require training studies that go beyond the demonstration of performance increase and probe details of the perceptual process. A simple option, for example, is to assess the use of context by manipulating sentence redundancy. Indeed this was the goal of Kalikow and colleagues when they developed the Speech Perception In Noise (SPIN) test (Kalikow et al, 1977). In this connection the k-factor transform (see Appendix B) can provide a useful index of the use of context (Boothroyd, 1978; Boothroyd and Nittrouer, 1988; Nittrouer and Boothroyd, 1990). The use of these and other techniques could have value, not only in describing outcome but also in planning intervention.

2. When performance improvements are attributable to training, what aspects of the training protocol were responsible? Training protocols are generally designed on the basis of some theoretical construct, with plausible reasons why the training should have the desired effect. But observation of the effect does not guarantee validity of the construct. It is always possible that the results were actually caused by a single feature of the protocol, or even a feature of which the designer was unaware. It is possible, for example, that the details of training are less important than the fact that it significantly increases time-on-task relative to everyday communication alone. It is incumbent on the researcher to follow the demonstration of positive outcomes with experiments designed to determine which aspects of the training were responsible.
3. How do individual characteristics interact with training and outcome? People with hearing loss differ in terms of auditory capacity, perceptual strategy, knowledge, personality, communication needs, and reaction to communication difficulty or failure. We lack information on how knowledge of these differences might affect candidacy for formal training or the choice of training goals and procedures. Such considerations suggest the need for more single-subject designs in this work (Hersen and Barlow, 1976; Kratochwill and Levin, 1992).

4. Did the changes influence participation and quality of life? The demonstration of small but statistically significant improvements of group mean performance on trained and untrained tasks is not sufficient to justify expenditure of time and money on formal training. The ultimate goal is bring the individual trainee to an acceptable (to him or her) level of communicative function in the real world, thereby increasing participation and quality of life. Self-reported benefits have an obvious role here, but their value would be enhanced by more objective measures.

There is a clear need for research that addresses these questions.

NOTE

1. We cannot recognize something outside of our knowledge. Indeed, the word recognition literally means “knowing again.”

References


APPENDIX A. Bayesian Modeling

Bayes's equation provides a quantitative model for perceptual decision making:

$$e \cdot pH = pH \left( \frac{H \cdot pe}{pe} \right)$$

(A1)

Where:

e is the sensory evidence (what you heard and/or saw),
H is the hypothesis being tested (the word you think might have been responsible for what you heard),
$\text{pe}$ is your estimate of the chances of what you heard being present regardless of whether $H$ is true,
$\text{Hpe}$ is your estimate of the chances of what you heard being present if $H$ is true,
$\text{pH}$ is your estimate of the chances that $H$ is true before you considered the evidence, and
$\text{epH}$ is your revised estimate of the chances that $H$ is true after you considered the evidence.

If $\text{epH}$ reaches or exceeds some acceptance criterion, you accept $H$. In other words, you decide that what you thought caused what you heard was the actual cause.

Note that $H$, $\text{pH}$, $\text{pe}$, and $\text{Hpe}$ are aspects of knowledge. But retrieving this knowledge, using it, and applying appropriate acceptance criteria are aspects of skill, as illustrated in Figure A1.

The Bayesian model used here to support the description of perceptual decision making leaves open the question of how perceptual hypotheses are first retrieved from knowledge. Presumably, this process involves contextual evidence, suprasegmental features, and the initial identification of stressed vowels, but the details are far from clear. The topic, however, is relevant to the issue of formal training and is in need of research. Note, also, that the description offered here is for a single decision process. The perception of continuous speech is obviously multilayered and involves many overlapping decisions about language patterns of varying complexity.

Figure A1
APPENDIX B. Quantifying the Use of Linguistic Context

The influence of linguistic context on the recognition of language patterns lends itself to quantitative modeling. Consider, for example, the recognition of words with and without sentence context. To a first approximation:

\[ p_{ws} = 1 - (1 - p_{wi})^k \]

Where:
- \( p_{ws} \) is the probability of recognizing words in sentence context
- \( p_{wi} \) is the probability of recognizing words in isolation
- \( k \) is a constant that accounts for the effects of both the redundancy in the sentence context and the listener’s ability to use it.

The k-factor is a dimensionless quantity representing the amount by which a listener uses context to increase the information contained in individual words. It is mathematically equivalent to the proficiency factor of articulation index theory. In the present application it tells us the amount by which the context multiplies the available information. The value of k can be determined by measuring word recognition in isolation and in sentences, under difficult listening conditions. The k-factor is the ratio of the logarithms of the two error probabilities

\[ k = \frac{\log(1 - p_{ws})}{\log(1 - p_{wi})} \]

Studies of normally hearing adults listening to sentences of known topic, under conditions of low-pass filtering, have given k-factors ranging from 4 to 10. These data show that the effect of sentence context on word recognition can be considerable but individuals may differ in the extent to which they take advantage of sentence context (Boothroyd, 2002).