GENERAL CONSIDERATIONS

PHONIC CONTROL OF ASSISTIVE SYSTEMS FOR REHABILITATION

P.H. Newell, Jr., A.M. Sherwood
T.A. Krouskop, and V.T. Rhyne

Summary

The loss of mobility through accident or disease greatly limits an individual's ability to pursue a productive life in society. Provision of the capability to control assistive systems both provides a degree of patient independence and the stimulation of morale which is so vital to rehabilitation. By freeing residual capabilities from the performance of necessary control functions, voice control offers a socially acceptable, natural control mode for handicapped persons.

Efforts with a simple phonic control device have demonstrated the usefulness of the concept. The first system employed resonant reed relays and a frequency meter as the sensing and feedback devices to implement four control channels for an electric wheelchair or an arm splint, with mode selection accomplished via a fifth channel. After evaluation of the feasibility of the control concept, the system was completely redesigned using all solidstate digital design. The new system provides improved reliability, lower cost and flexibility in adjusting to individual users. Suggestions for future research and development activities are presented, along with a discussion of the implications of voice control for rehabilitation.

Introduction

The application of modern medical technology and the availability of improved transportation have resulted in an increasing number of cases in which previously fatal injuries can be sucessfully treated. This, however, has the side effect of creating a larger group of temporarily and permanently disabled individuals.

This work was supported in part by the Social Rehabilitation Service, in part by the Veterans Admimistration, and in part by Texas Engineering Experiment Station.

Of particular interest to the rehabilitation community are those individuals with spinal cord injuries, due to the fact that an otherwise uncomplicated injury can result in a severe limitation of the individual's activity. At the present time there are approximately 60,000 quadriplegics in the United States alone and, of these individuals, a large number of them have severe enough involvements to make them essentially entirely dependent upon others for their daily maintenance. Aside from the obvious nursing problems this creates, the individuals involved have a tremendous morale problem, in the recognition that they will be forever dependent upon others to execute even the simplest functions of daily living.

The application of modern technological achievements to this area of rehabilitation medicine then becomes particularly important in that it provides the possibility of restoring some of the lost functionality. It is to this restoration of independent functionality and to the provision of a degree of self-sufficiency for these individuals that the rehabilitation efforts of the Texas A&M University Bioengineering Program are directed. There are a number of readily available devices on the market today which permit the individual to accomplish almost any ends he desires provided that he has the means to control these devices. Thus, the problem of critical importance in the rehabilitation effort which dominates most of this total area of endeavor is the search for the most natural and effective control means. In expolring alternatives of control, particulary for those individuals paralyzed from the neck down, it becomes obvious that the only remaining natural means which offers flexibility in terms of a wide latitude of control responses, and which at the same time is socialy acceptable, is voice control. This control has many diffrent modes of operation and a very large potential informational content. There are several other alternatives available to the individual paralyzed from the neck down such as the tongue switch and the various kinds of eye switches. However, these outputs from the physiological system can be arranged for only a rather limited number of distinct stimuli and represent a significant trade-off in terms of cost or performance. Voice control provides minimum interference with activities such as eating as compared to tongue or mouth switches, for example. Eye switches not only are expensive, but also cause distraction in the required motion of the eye away from the object of interest. The economic application of modern technology in providing voice recognition is the most promising control technique, and is the area of concern in this presentation.

Background

A number of groups and individuals have been investigating the feasibility of direct voice conversion for application in a wide number of areas ranging from sorting parcels in a post office to directly interfacing with sophisticated computers in scientific applications. The solutions to these problem situations have in general not been totally satisfactory either from a performance standpoint of from a cost standpoint. Complex programing schemes used on large analog, hybrid, or digital computers have been used with only limited success in the general recognition problem, i.e. in which the speaker, who may be almost anyone, can speak on a number of topics with a rather wide vocabulary. In general, if the machine is "trained" to respond to a limited group of persons, say 4 or 5, then it can successfully translate a fairly broad vocabulary, but mostly of disjointed speech /1/. The problems of continuous speech are quite difficult, and are being approached from both the speech recognition and the speech production standpoint /2,3/. On the other hand, there are available today, machines which can recognize only a limited vocabulary, for example, the digits from zero to nine, but will correctly translate these from the speech output of almost anyone /4/. Indeed some digital computer programs have been written which allow the computer to teach itself the vocabulary required for translation, on the basis of examination of the content of the message.

The only possible way of interfacing a large, computer speech translator to rehabilitation efforts would require a telecommunications link of some sort. The reliability of this link and the range of coverage would not be adequate for general requirements and still stay within a reasonable cost. There is at the present time, however, one commercially available system which can be trained to respond to a vocabulary of some 30 words from one individual at a time and can recognize these words with a fair degree of accuracy. However this device costs, approximately $10,000 and is about the size of a central processing unit of a mini-computer /6/. Therefore at the present, the alternatives available from these existing research programs do not seem attractive in terms of controlling assistive devices for the handicapped individual.

The negative results of these general efforts do not, however, indicate negative result for the special case of the hadnicapped individual. This particular problem allows a great deal of simplification of the general problem, in that only one individual will operate each device and only a limited vocabulary is required. Thus, it is feasible to explore further the potentialities of this type of control mechanism.

As a result of experiences gained in working with a rehabilitation clinic at the University of Alabama Medical School, the first author recognized the potential for development in this area and conceived the idea of the phonic controller which responds to the pitch of a hummed signal. This design concept was implemented in a feasibility prototype using resonant reed relays as the frequency selectors /7/. This device had five distinct input frequencies corresponding to four control functions for either a wheelchair or a hand splint, with the fifth function being used to change the mode of operation from the wheelchair to the hand splint. Because the resonant reed relays are inherently very narrow bandwidth, typically on the order of one hertz, three relays were put in parallel, each tuned to a slightly different frequency to extend the overall bandwidth to approximately 3 to 4 hertz. Along with this controller, an inexpensive frequency meter was developed, which allowed visual feedback to indicate what frequency was being hummed in relation to the control bands. Thus the operator could readily see whether he needed to raise or lower the pitch of his voice to select the desired function.

However, once the control bands are set by adjusting the resonant frequencies of the respective relays, it is quite difficult to change the band frequencies. This initial device was free of false positives in the sense of a response to the voice signal rather than the hummed signal, because of the resonant nature of selection device. The first unit used a carbon throat microphone of the conventional type and was also provided with a shoulder switch for emergency cut-off. Power for the device came from the system being controlled, i.e. when used with the wheelchair, the operating power came from the wheelchair batteries.

This early model did demonstrate the feasibility of the concept in that the wheelchair could be driven forward, turned left or right and driven in reverse, simply by controlling the pitch of the hummed signal. However, the inability to adjust the frequency to individual users and potential problems in maintenance of the original device led to the decision to redesign the unit using a digital logic approach rather than the resonant reed relay approach to detect the pitch of the hummed signal.

Current Design

Capitalizing on the experience gained with the original unit, a completely redesigned phonic controller was developed at Texas A&M University. This new design depends upon a digital counter and a comparator circuit rather than a resonant device to extract the frequency information. The input device for the digital unit consists of a dynamic throat microphone, which gives good protection from external interference and excellent sensitivity to hummed signals. The microphone output is coupled to a bandpass amplifier and Schmitt trigger combination (see Figure 1). The output of the Schmitt trigger goes to the digital logic which in turn drives the feedback devices and relays which control the motors.

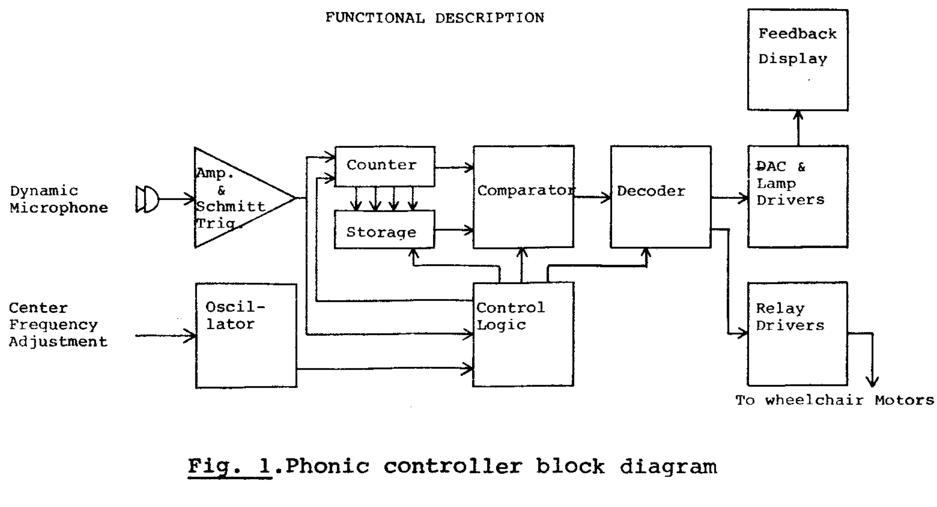

Fig. 1.Phonic controller block diagram

The sequence of operation begins with the reception of a hummed signal in the throat microphone and the subsequent amplification and squaring by the amplifier - Schmitt trigger combination. Following the Schimtt trigger is a two decade counter which totals the number of zero crossings of the processed hummed signal per set period of time. The counting period is controlled by an oscillator which allows it to count for approximately one quarter of a second before stoping the count and comparing the most significant four bits of the counter and the four most significant bits of the previously counted and stored value. If the counter value and the stored value are identical for three sucessive counting periods, which indicates that a hummed signal, not a speech signal is present, the control logic allows the decoding circuit to respond to the number stored. This number is stored as a binary coded decimal, which thus can indicate up to 10 separate states. Of these 10 states, only 6 are used to control the wheelchair motors. The number 3 corresponds to a hard LEFT turn, the number 4 to a LEFT and FORWARD, the number 5 to FORWARD, the number 6 to RIGHT and FORWARD, the number 7 ti hard RIGHT, and the number 9 corresponds to REVERSE, with the number 8 reserved as a guard band (Fig. 2). The control logic also contains a section which disables the decoding circuits if the input signal is stopped for more than two-hundredths of a second. This prevents false positives following a desired sequence of commands, such as when an individual begins talking immediately after stopping humming.

An important part of this system, as of the initial system, is the feedback employed. The feedback system again is comprised of a meter which displays the relative frequency of the hummed signal. There are also four lights which indicate the selected function. Although the internal design of this feedback system is quite different from the original feasibility prototype design, the functioning of the two units is almost identical, except that, in the current design, the indicated frequency corresponds to the selected function by changing along with changes in the oscillator frequency. By observing the motion of the needle on the frequency meter, one can readily gain proficiency in the operation of the controller and the controlled device.

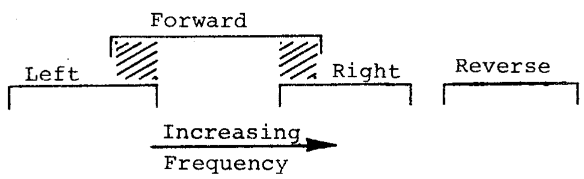

Fig. 2. Control bands for the tamu phonic controller

Since the oscillator frequency is externally adjustable by means of a potentiometer, and since the oscillator frequency determines not only the respons time of the unit, but also the frequency which corresponds to for example, "5", or FORWARD, it is simply a meter of changing the potentiometer setting to adjust the frequency range to suit different individuals. For example, by increasing the oscillator frequency and thereby shortening the counting period, the external frequency corresponding to the FORWARD command can be raised to accomodate a female rather than a male user. There is, however, one negative aspect to this means of changing the frequency. The response time is also changed when the center frequency changes, and since the response time is a critical parameter for ease of operation, it might become necessary to make internal wiring changes to accomodate some extremes of frequency ranges which might be desired. However, within the normal range of frequencies utilized, this presents no real problem. The present form of the implementation, namely the use of a wirewrap panel, permits any modifications in design to be easily made.

Operational Features

Once the unit is mounted on the wheelchair, its operation is easily intergrated into the routine handling of the patient. To activate the controller, one simply positions the microphone on the throat, and turns on the both the wheelchair and the controller unit. The controller is now ready to use. As long as the user does not sustain a tone for the required period of time the unit will not operate, thus preventing false starts while talking. To drive the wheelchair forward, one simply humms the correct tone, holding it for the required three-quarters of a second to initiate action, and then sustaining the tone for as long as one desires to continue in the same direction. However once the system is activated, it responds within approximately one-quarter of a second to a change in the control input. The exact length of the time to respond to a change depends upon both the oscillator frequency and upon the relative time within the counting cycle at which the change occurred.

Since the comparator and decoder circuits are based on the most significant bits, the control signal need only be maintained within a relatively broad band of approximately 15 to 20 hertz to maintain the same control function. As long as the pitch of the humming is maintained within the operating range, the operator can change frequencies and thereby steer the chair in the desired direction quickly and effectively. Since the wheelchair is driven by two independent motors, there is a general tendency for the chair to drift to the right or left when (supposedly) being driven straight ahead. However, because of the ease of phonically controlling the direction of motion, this presents no real problem, an minor course corrections can be readily and effectively made. With the existing system it is also possible to maneuver in close, crowded spaces such as one might find in an office area. Due to the effectivness of the microphone location and characteristics and to the amount of amplification present, only a very low-level hummed signal need be present to activate the device. This level of sounds has not been found to be objectionable in any application to date.

At the present time two units are being evaluated by the Veterans Administration Hospital in New York City. It is anticipated that from these clinical evaluations further refinements will be suggested and will be implemented in further models of the device. Some of the anticipated problem areas lie in the area of the physiological constraints on the user. For example, for a quadriplegic with so severe an involvement of his extremites as to require this type of device, the operation of the controller might be excessively fatiguing, or might require more breath control than the individual has. The optimum type and placement of the feedback system must also be considered.

Future Activity

There are many possibilities for further development in the general area of vioce control. Several of these will be pursued initially on the basis of their significance and ease of accomplishment. One of these is the provision of a toggle mode of operation. At the present time, in order to provide maximum safety of operation, the chair only operates while the operator continues to hum. This obviously indicates that when the operator runs out of breath, the chair will stop. One way to surmount this problem would be to provide the chair with a toggle mode, that is, a specific signal would cause the chair to turn on and continue to go forward until another signal is received which either steered the chair to the left or the right or caused the chair to stop. Preliminary design work on this mode of operation has begun, primarily in evaluation of the safety hazards.

Closely related to the toggle mode is to control the speed. For close-in maneuvering, the chair ideally should be driven in low speed, because of the problems of overshooting due to inertial effects. However, when traveling over relatively long distances down a hall or between buildings, it is desirable to move at the fastest possible speed for that chair. Thus a means of speed control would be desirable in the chair operation. This could be acheived in several ways, probably the simplest of which would be to design the controller such that a sustained forward direction motion for more than a fixed interval of time, such as 10 seconds, would cause the chair to go high speed mode, while all short range, i.e. those maneuvers immediately following the initiation of chair operation, would be accomplished at low speeds.

The question of safety and reliability of operation must also be considered. However, to a large extent, these problems are more dependent upon the controlled device than they are on the controller. For example wheelchair braking is rarely provided on electric chairs, on the theory that the motors can be used. This is not necessarily true in all cases, since commonly encountered ramp slopes are too steep to allow effective use of existing wheelchair designs. A careful study of the entire wheelchair system is therefore indicated.

The final short range objective for this controller involves the ability to change modes such as was present in the original feasibility prototype. This can be readily accomplished and would allow the controller to be interfaced to other devices, such as a typewriter, television, telephone, or other similar environmental control device. The mode control could also be used to aid in integration the phonic controller into a larger control system such as might be used in a vehicular control environment. Although each of these devices presents unique challenges in interfacing, they share a mutual difficulty in physically interconnecting with control unit which presumably is mounted on the wheelchair. Thus a common interconnection scheme is needed along with specific interfacing logic.

As a parallel activity, the development of a small vocabulary word controller is being pursued at the present time. Even though the general problem is quite difficult, as mentioned above in the background section, it is clear that a sufficient reduction of the requirements will make it feasible to design a word controller. For example, there are a number of quite inexpensive controllers on the market which respond simply to the number of syllables uttered. By sufficient restrictions on the frequency of the inputs or the complexity of the control vocabulary, it is obvious that a simple discrimination procedure can be devised to "recognize" these highly constrained "words". Thus the question becomes one of optimizing in the trade-off between complexity and resultant costs on the one hand and simplicity and resultant loss of functionality on the other. In a careful consideration of alternative control schemes for a single user, a limited word control device will therefore be included.

Conclusion

The phonic control work to date has demonstrated the feasibility of voice control both in terms of operational effectivness and of low costs. Currently these units sell for approximately $400 each, with a segnificant reduction in cost possible with higher volume production. This control scheme is unobtrusive, very acceptable socially, easy to learn, presents minimum interference to activities, and much potential for future development. With further sophistication of the command sequences voice control would allow the quadriplegic to engagé in such activities as eating which are not permitted by devices which involve the use of his mouth, chin, or tongue and similarly would not be distracting as are the devices which require him to shift his vision.

It is felt that this device and the potential for develpment inherent in it represents a significant start towards the goal of providing a quadriplegic with the necessary tools to become selfsufficient in his daily life, and indeed with the possibility of regaining some measure of self-support.

References

/1/ Bobrow, D.G. and Klatt, D.H., "A Limited Speech Recognition System", *Proc Fall Joint Computer Conference*, pp. 305-318, 1968.

/2/ Pierce, J.R., "Whither Speech Recognition", *J. Acoust. Soc. Am.* 46 (4), pp. 1049-1051, 1969.

/3/ Noll, A.M., "Whither Speech Production", *J. Acoust. Soc. Am.* 47 (5), pp. 1612-1617, 1970.

/4/ Scarr, R.W.A., "World Recognition Machine", *Proc. IEE* 117 (1), pp. 203-212, Jan., 1970.

/5/ Clapper, G.L., "Automatic Word Recognition", *IEEE Spectrum* 8 (4), pp. 57-69, Aug., 1971.

/6/ *Scope Electronics*, 1860 Michael Faraday Drive, Reston, Va. 22070, "Voice Command System".

/7/ Newell, P.H.,Jr., and Barr, J.S., "Voice-Operated Powered Devices", *Mech. Engr.*, pp. 25-27, Nov., 1970.