EVIDENCE OF SUCCESS
IN SPEECH AND VOICE TRAINING

Mr. Nicholas A. Orlando, M. S.
Supervising Speech Pathologist
NTID Communication Center

INTRODUCTION. In considering any program of rehabilitation for the young deaf adult, questions relating to communication become central issues. This paper reports some of the questions asked relative to speech and some of the work undertaken in efforts to gain answers. At best, the answers derived are incomplete, but they have been useful in defining the policies and program of the Speech Service Section within the Communication Center of NTID.

In order of priority, the questions posed were as follows:

1. Can significant improvement in speech be achieved in the young deaf adult?
2. If so, how much improvement can be achieved?
3. Can greater improvement in speech be attained as a result of 18 hours instruction by scheduling therapy two hours per week for 9 weeks, as opposed to scheduling one hour per week for 18 weeks?
4. Can pitch register and pitch control be improved in the profoundly deaf student?

SAMPLE. In order to approach the first few questions, 79 NTID students were selected for individual speech therapy which was designed to correct articulatory and prosodic errors. Students with voice quality and/or vocal pitch deviations of moderate or severe degree were excluded. Since little clinical effort can be given to the student with good intelligibility (profiled 5), the major consideration in sampling was to select students from the other profile categories of 1, 2, 3, or 4. This procedure was followed to sample varying levels of speech skill so that progress in therapy, if it was realized, could be studies relative to a specified pre-therapy status in oral communication, as defined by the profile intelligibility rating system.

A satisfactory distribution of students falling within respective profile ranks of: 4 (n 17), 3 (n 24), 2 (n 19), and 1 (n 16) was attained. Three additional students, profile 5 in speech intelligibility, but with some residual articulatory errors, were included for experimental purposes.

The total group was then divided using pre-therapy speech profile rating, articulation measures, high frequency cutoff in the better ear, and puretone average [1] to establish two groups for comparative studies.

Means, standard deviations, and t values presented in Table 1 reveal the two groups did not differ significantly in speech or hearing before therapy. As reported, Group I averaged 40.03% total articulation error. Group II averaged 39.90% total error. Although Group I had a slightly higher percent vowel error (37.90%) than was observed in Group II (31.05%), this difference was not found to be statistically significant. The means of puretone averages for 250, 500, and 1000 Hz were 98 dB (sd 13.40) and 95 dB (sd 14.54) respectively, thus indicating the two groups were severely and comparably impaired in hearing.

TABLE 1: Means, standard deviations and t values to evaluate differences in speech and hearing status for two groups of NTID students scheduled for individualized speech therapy. (N = 79 NTID students)

Therapy Measurement	Group I N = 40		Group II N = 49		
Articulation					
AAPS+ Total Score	mean	65.00	mean	67.00	+ .514
	sd	18.50	sd	16.00	
% Total Error	mean	40.03%	mean	39.90%	+ .328
	sd	18.57	sd	15.79	
% Vowel Error	mean	37.90%			
% Vowel Error	mean	37.90%	mean	31.05%	+ 1.526
% Consonant Error	mean	41.75%	mean	43.28%	− .375
	sd	19.71	sd	16.41	
Intelligibility					
Profile Rating	mean	2.50	mean	2.80	− 1.148
	sd	1.20	sd	1.08	
Hearing					
High Frequency Cutoff	mean	3,275 Hz	mean	3,846 Hz	− .936
	sd	2,736	sd	2,688	
Puretone Average	mean	98.15 db	mean	94.95 db	+ 1.018
	sd	13.40	sd	14.54	

+ The Arizona Articulation Proficiency Scale is a numerical scale of articulatory proficiency in which the numerical values for the sounds are related to their probable frequency of occurrence in American Speech.

Group I included 40 NTID students receiving individual speech training for one hour a week (2 half hour sessions) for 18 weeks or for two consecutive quarters. Group II included 39 students receiving training for two hours per week (two one hour sessions) for 9 weeks or four one quarter. Every student received approximately 18 hours of individual speech training. Eight experienced therapists provided training which was divided rather equally between groups. All therapists

1 To calculate a puretone average when no response could be attained at 500 or or 1000 Hz, dB values were assigned a value of 117 dB as derived from a special formula developed at NTID.

applied the same criteria in selecting targets for correction; all used amplification and followed a non-programmed but well organized course in training.

PROCEDURE. To evaluate the results of therapy, the Arizona Articulation Test and the W-22 Word Lists [2] were tape recorded before and after therapy. Words were recorded at five second intervals, using high fidelity equipment and a constant mouth to microphone distance. Pre and post therapy articulation scores were derived by playing back tape recordings of the Articulation Test for independent phonetic analyses by two speech pathologists. In playback, pre and post therapy recordings were randomized. Inter-judgment agreement in the phonetic analyses were acceptable. The incidence of error varied 2% or less in 74% of the students analyzed.

Scores of the two auditors were then averaged to determine the final articulation measures for each student. In reporting results, the Arizona Articulation Proficiency Scale total score and percentages of error were determined.

The Arizona Articulation Proficiency Scale was initially selected for use at NTID because it provided numerical values for phonemes weighted relative to their occurrence in spoken English. Although this measure may be more closely related to speech intelligibility than other articulation measures, the AAPS score, of itself, did not provide other information desired. For this reason, incidence of error for varied phonetic groupings was determined from the Arizona Test and expressed in percentage.

Word intelligibility was determined by playback of the W-22 lists for five auditors. In recording the four W-22 word lists, word lists and word order were varied to reduce the effect of familiarity of material upon word recognition. A word was scored as incorrect if it was not the phonetic equivalent of the intended utterance. The results of the five auditors were then averaged to determine the percentage of words correctly understood for each student.

RESULTS. *Can significant improvement be achieved? Does the variable of scheduling influence articulation improvement?* Responses to therapy were evaluated in terms of the Arizona Articulation Proficiency Scale (AAPS) scores, percentage reduction in articulation error, and percent improvement in word intelligibility. Results, summarized in Table 2, show that Group I, receiving therapy one hour per week for 18 weeks, averaged a 10% reduction in total error: Group II, receiving therapy two hours per week for 9 weeks, averaged 9.18% reduction in total error. The

---

2 *Four W-22 Word Lists were used to assess word intelligibility, simply because the same lists were being used in the Audiological Service Section to assess hearing discrimination at the time of the study. It was originally proposed that speech intelligibility measures for each student would be studied relative to the student's performance in auditory discrimination as measured by the W-22 lists.*

only significant difference between scores for the two groups pertained to percent reduction in vowel error. Group I, receiving therapy for 18 weeks, attained a greater gain (12.68%) than that attained by Group II (3.03%), receiving the same number of hours within 9 weeks.

TABLE 2: Means, standard deviations and t values to evaluate differences in response to therapy as a function of scheduling; Group I received 1 hour per week for 18 weeks; Group II received 2 hours per week for 9 weeks.

Measurement	Difference (Pre-Post Therapy Measures)				
	Group I N = 40		Group II N = 39		
Articulation					
AAPS Score	mean	10.46	mean	7.08	1.665
	sd	11.32	sd	5.78	
% Total Error	mean	10.00%	mean	9.18%	.418
% Total Error	mean	10.00%	mean	9.18%	.418
	sd	10.99	sd	5.53	
% Vowel Error	mean	12.68%	mean	3.03%	3.618**
	sd	14.08	sd	9.00	
% Consonant Error	mean	10.70%	sd	11.54%	−.371
	sd	12.06	sd	7.40	
Intelligibility					
Word	mean	7.80%	mean	4.45%[+]	1.583
% Words Correct	sd	8.75	sd	7.87	

**Significant at .01 level
+ N = 26 for this measure

This observation suggests that the more concentrated 18 hours of therapy schedule may have unwisely over-emphasized consonant correction per se, and did not provide enough time to stabilize the target within varied vowel contexts. As a result, vowel error was not reduced as much as desired and word intelligibility did not improve (mean 4.54%) as much as was observed in Group I (mean 7.80%) receiving the same number of hours over a longer period of time. Although the difference in word intelligibility was not statistically significant, a tendency for the better performance in vowel articulation to be identified with greater gain in intelligibility is noted.

How much improvement can be achieved? The relative speech improvement, expressed in percent reduction in total articulation error, resulting from 18 hours of individualized speech instruction averaged approximately 10% and did not vary as a function of difference in scheduling. All other factors excluded, the amount of improvement seems generally related to the number of hours provided in training.

On the basis of these findings, a two hour per week therapy is recommended, provided the speech staff can fulfill its responsibilities in meeting the varied needs of NTID students. Doubling the weekly hours in therapy with the same clinical staff, of necessity, reduces the

number of students in therapy during any one quarter. This disadvantage may be partially overcome by selecting students carefully and by limiting therapy in most instances to two consecutive quarters, to be followed by group therapy designed to stabilize newly acquired speech skills. Limiting individual therapy to two consecutive quarters is, in most instances, a necessity at NTID for two basic reasons:

1. The average duration of stay on campus is two years;
2. NTID students have pressing needs for communication training in other areas, i.e., auditory training, speechreading, etc.

Since no significant differences in percent total error were revealed as a function of the scheduling of 18 hours of therapy, data for both groups were combined to study the modifications in articulation achieved and to relate achievement to pre-therapy speech status.

MODIFICATIONS IN ARTICULATION AS A RESULT OF THERAPY-CLINICAL IMPLICATIONS. In Table 3, the pre-therapy and post-therapy articulation scores are reported. Total scores on the AAPS, and reduction in error expressed in percentage, show that improvement in speech articulation significant at the .01 level was achieved in all phonetic groupings after 18 hours of individualized training.

In overview, approximately 9% reduction in total error was realized. Consonant error reduction (mean 11.12%) was better than vowel error reduction (6.81%). This may again reflect too much emphasis on consonant correction without enough time given to practicing consonants in varied vowel contexts within functional word groupings to attain good prosody, and smooth transition to preserve the perceptual identity of both phonemes in continuous speech production.

In this regard it should be noted that vowel error before therapy averaged 33.72%. Consonant error (mean 42.51%) was less than 10% higher. These findings again stress the importance of training for consonant improvement using a variety of vowel contexts in carefully structured speech practice materials. If materials are designed for the purpose of sequentially varying the target within varied vowel context, and within relevant language units (functional word groupings), training will be more effective and successful carryover into conversational speech facilitated. As mentioned previously, greater improvement in vowel articulation was identified with greater gain in word intelligibility.

Since consonant targets for correction were individually determined, a varied pattern in improvement relative to phonetic grouping would be anticipated and was realized. The percentage reduction in error for plosives (mean 9.00%), nasals (10.57%) and glides (mean 9.77%) tended to be similar, as graphically presented in Figure 1. Fricative error reduction (mean 14.04%) tended to be higher. This is probably because the /s/ is commonly defective and has a high incidence of occurrence in conversational English. For these reasons, /s/ is fre-

TABLE 3: Means, standard deviations and t values for measures of articulation secured before and after individualized speech training. (N=79 NTID students)

Measurement		Pre Therapy	Post Therapy	t
AAPS+				
Total Score	mean	66.12	74.91	+8.556**
	sd	17.37	14.72	
Percent Error				
Total	mean	39.09	30.37	-6.762**
	sd	17.29	14.71	
Vowel	mean	33.72	26.91	-3.980**
	sd	20.09	17.16	
Consonant	mean	42.51	31.39	-9.897**
	sd	17.94	14.95	
Plosive	mean	30.62	21.62	-6.426**
	sd	20.17	15.88	
Fricative	mean	48.15	34.11	-8.197**
	sd	19.19	16.46	
Affricate	mean	79.35	71.34	-3.157**
	sd	27.10	29.69	-3.157**
Nasal	mean	32.90	22.33	-5.854**
	sd	25.21	16.79	-5.854**
Glide	mean	46.05	36.28	-3.781**
	sd	27.47	21.94	
Total Voiced	mean	49.73	39.52	-5.864**
	sd	19.85	16.52	
Total Voiceless	mean	36.68	26.25	-6.023**
	sd	19.55	16.56	
Blend	mean	63.18	44.01	-6.971**
	sd	26.92	24.66	

+ The Arizona Articulation Proficiency Scale is a numerical scale of articulatory proficiency in which the numerical values for the sounds are related to their probable frequency of occurrence in American Speech.
** t 2.660 significant at .01 level.

quently selected as a target. As a correlary, less gain was attained for the affricates (8.01%). Affricates, which are extremely difficult for most deaf students, are not emphasized in therapy because of low frequency of occurrence in conversational English.

Consonant blends generally are not emphasized in the speech training of most deaf students at NTID. This decision is based upon the premise that an acceptable blend should not become a target for correction unless the student can produce most consonants in cvc utterances intelligibly. Specifically, a blend is not considered target unless the student can produce both elements of the blend in cvc

utterances with minimal distortion. Despite this policy, a significant gain in the articulation of blends was achieved as a result of training. This finding increases confidence that emphasis upon cv, cvc, and vc utterances in functional word groups is correct. In overview, improving articulatory precision of individual phonemes in varied word context apparently improves the student's ability to produce consonant blends appropriately.
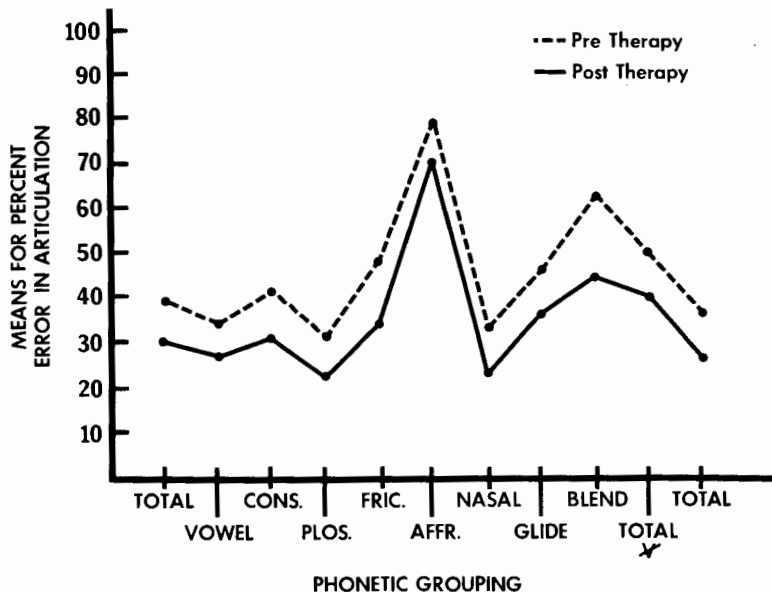

Figure 1. Pre and post therapy means graphed for articulation errors catagorized by various phonetic groupings (N = 79 NTID students)

The deaf student's problem in confusing voiced-voiceless pairs is also confirmed in Table 3. Before therapy, error on voiced consonants (mean 49.73%) tended to be higher than errors on the voiceless counterparts (mean 36.68%). After therapy, voiced and voiceless errors were both reduced 10%; however, the higher incidence of error on the voiced phonemes persisted.

The apparent higher incidence of error of voiced phonemes confirms the observation that some deaf students do not understand the concept of voicing; differentiate voice-voiceless features auditorily, and sustain phonation consistently between voiced elements in cvc utterances. The cessation of voicing between phonemes destroys the prosody of the pattern as well as the acoustical transitions important to the perceptual identity of both phonemes.

Since this liability relating to voicing affects many phonemes and seriously effects intelligibility, the pattern of errors should be carefully evaluated as it relates specifically to the distinctive feature of voicing before beginning therapy. To illustrate, if several voiced-voiceless sub-

stitutions occur within the plosives and/or fricative category, considerable effort is indicated to explain voicing and develop tactile and/or auditory differentiation of voicing before beginning speech therapy per se.

ERROR TYPE — CLINICAL IMPLICATIONS. Further analysis of data for 40 students (Group I) was undertaken to study error type and improvement in articulation relative to error type. Before therapy, errors identified as distortions (mean 14.77) exceeded by twice the number identified as substitutions (mean 6.52). Omissions were much less frequent (mean 1.92). The latter finding is favorable in the sense that most NTID deaf students seem to have a fair concept of consonant production. They were at least attempting most phonemes, albeit many times a distorted or substitute version of the sound was achieved.

The high incidence of distortions is favorable in another way. It should be remembered that distortions, counted in articulation testing, may not be severe enough to destroy the perceptual identity of intelligibility of the phoneme in speech context. That is, many phonemes may be distorted and tolerated perceptually without losing intelligibility[3]. This seems especially true if the listener has visual access to the speech gestures of the deaf speaker.

The incidence of error type before and after therapy is presented graphically in Figure 2. After therapy, distortion errors averaged 10.17, substitutions averaged 5.72, and omissions averaged 1.45. Reduction in distortions as a result of training (mean 4.60) far exceeded the reduction attained for substitutions (mean .80) and omissions (mean .47). This finding, which would be anticipated in a hearing population, indicates that distortions in the speech of the hearing impaired population at NTID can be more easily corrected than other error types. This finding is also consistent with data accumulated at NTID which shows many distortion errors noted in diagnosis can be corrected with stimulation. Compositively, these observations are favorable relative to prognosis for speech improvement, especially since distortions have been found to be the most common error type.

INDIVIDUAL VARIATION IN RESPONSE TO THERAPY. The comparatively large standard deviations reported for the difference between pre and post therapy speech measures (Table 2) attests to great variations in individual response to therapy. To illustrate, changes in word intelligibility after therapy for the 40 students in Group I ranged from -7.6% to +32.0%, with a mean gain of +7.8%, sd 8.75. The distribution of relative gain or regression in word intelligibility as a consequence of the 18 hours of therapy offered over an 18 weeks interval of time, is reported in Table 4. In overview, the findings are positive. Eighty-five percent of the students improved. Sixteen students

3 *This difference between articulation and intelligibility testing may partially explain the fact that proportionately greater gain was achieved in articulation (10%) than was indicated by intelligibility assessments (7%).*

58

(40%) showed a good gain exceeding 10%; eighteen students (45%) gained 10% or less, but six students (15%) regressed in word intelligibility as a result of the training.
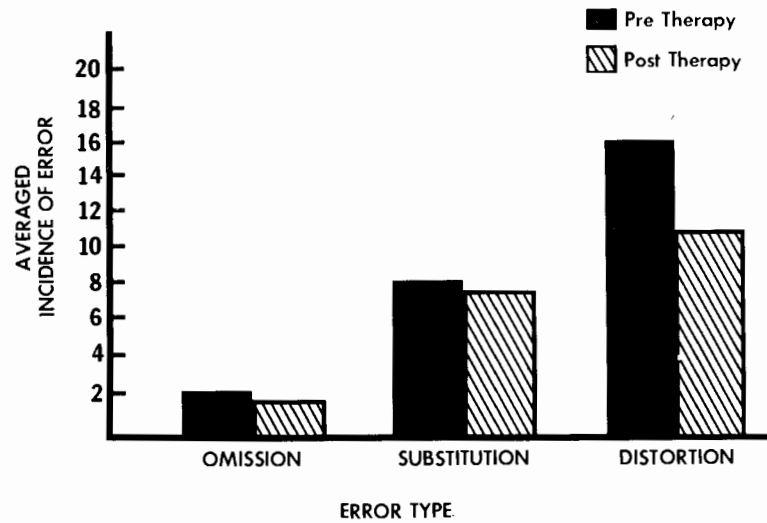

Figure 2. Averaged incidence of error type before and after speech training (N = 79 NTID students)

TABLE 4: Incidence of relative gain or regression in word intelligibility after receiving 18 hours of individualized speech therapy. N=40 NTID students)

Range of Scores in Percentage for Word Intelligibility	Number of Students	Percentage of Students	
+30.1—35%	1	2.5	
+25.1—30%	2	5.0	
+20.1—25%	0	0	40%
+15.1—20%	2	5.0	
+10.1—15%	11	27.5	
+ 5.1—10%	9	22.5	45%
+ 1.0— 5%	9	22.5	
0%	0	0	
− .1— 5%	3	7.5	15%
− 5.1—10%	3	7.5	

To examine variation in response on the basis of pre-therapy speech status, data for 79 students was regrouped on the basis of speech intelligibility profile categories. Means and standard deviations

for percent reduction in articulation error for each profile grouping are reported in Table 5, and graphed in Figure 3.

TABLE 5: Means and standard deviations for percent reduction of articulation error achieved after eighteen hours of speech training, with students grouped on the basis of pre-therapy profile ratings. (N = 79 NTID students)

Speech Profile Rating	Number	Reduction in Articulation Error	
1	16	mean sd	11.56% 11.01
2	19	mean sd	10.95% 6.67
3	24	mean sd	9.63% 9.83
4	17	mean sd	7.29% 6.70
5	3	mean	3.33%

A tendency for reduction in error to be greater in the lower profile groupings is noted. That is, students in the lower intelligibility groups before therapy seemed to show somewhat greater reduction in articulation error as a result of training.
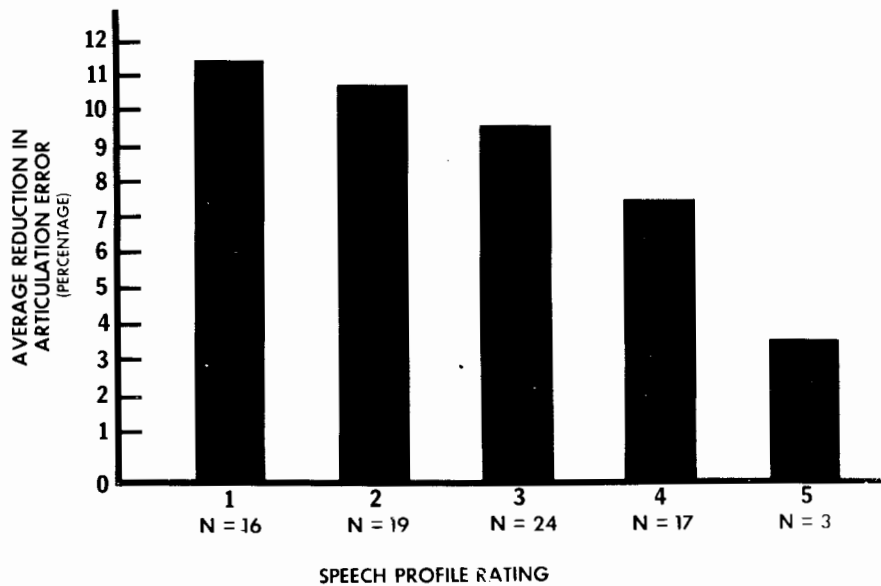

SPEECH PROFILE RATING

Figure 3. Average reduction in articulation error achieved after therapy with data grouped on the basis of pre therapy speech profile rating (Total N = 79 NTID students)

This observation encouraged further efforts to examine more closely the relationship between pre-therapy speech status, defined by *percentage error,* and reduction in error achieved as a result of therapy. A statistically significant positive correlation ($r = +.464$) was obtained indicating again that students with a higher incidence of error before therapy tend to show greater reduction in error as a result of therapy providing they have sufficient motivation and have learned the basic coordination of respiration and phonation to sustain and control vocal pitch within an appropriate range. Although students with faulty pitch were not included in these studies, the reader is reminded that unfortunately many profoundly deaf students with a high incidence of articulatory errors do not meet the stated criterion for success in therapy.

The logical question posed for these students is as follows: *Can pitch register and pitch control be improved in the profoundly deaf student?* Thirteen percent of the students entering NTID have pitch registers grossly inappropriate for age and sex, with an associated inability to control pitch. In general, this group of students is characterized by profound hearing loss, faulty coordination of respiration and phonation, and an inability to manage the aerodynamics of the breath stream of intelligible speech production.

With this group of students in mind, a program of therapy to improve pitch register and control in deaf speakers was developed and evaluated after students received two hours of therapy each week for one academic quarter (18 hours).

To identify 10 students most appropriate for training, speech samples of over 150 students were reviewed. The initial identification of students with high pitch register and lack of pitch control was con. firmed independently by two speech pathologists after live speech and voice evaluations. Audiological data revealed all students had profound hearing loss. Puretone averages for the better ear ranged from 92 dB to 116 dB, with a mean of 107 dB. Loss at 250 Hz ranged from 75 to 110 dB, with a mean of 84 dB.

The program of therapy formulated emphasized maximum use of tactile, kinesthetic and proprioceptive feedback, without neglecting the visual and auditory systems. The general objectives for the student in training were:

1. To become aware of the vertical shifting of the larynx during swallowing and during phonation of vowels at high and low pitch levels;
2. To become aware of changes in the height and tension of the larynx as vocal pitch changed;
3. To voluntarily produce low and high pitch levels during sustained vowel phonation and during syllable repetition;
4. To stabilize self-monitoring of vocal pitch during production of syllables, words and phrases;
5. To increase the duration of relaxed phonation.

To evaluate the effectiveness of the training program, speech

materials were recorded before and after therapy. The Rainbow Passage was recorded for perceptual ratings. In addition, ten sentences randomly selected from the 600 Clarke Intelligibility Test Sentences were recorded to assess intelligibility. All recordings were made under satisfactory conditions, using high fidelity equipment.

Five experienced listeners independently rated the Rainbow Passage for pitch register and control. Pitch was rated on a seven point scale. A rating of 4 indicates pitch appropriate for age and sex, while 7 indicates pitch much below optimum, and 1, much above optimum.

Pre and post therapy ratings for each subject are reported in Table 6. As reported, all subjects were substantially above normal pitch before therapy. After training, 4 students had normal pitch and 4 had pitch slightly above or slightly below normal. Two students did not improve perceptually. Both of these students were unable to voluntarily shift pitch. On the basis of this observation, it is recommended that voice therapy be discontinued if the speaker cannot learn to shift pitch voluntarily.

The averaged ratings of pitch control revealed all subjects achieved control with no pitch breaks observed after therapy. Although subjects did not achieve normal intonation, all could control the fundamental—within a limited speaking range.

The measures of intelligibility derived by auditing the pre and post therapy records of the Clarke Sentences revealed word intelligibility within sentence context averaged 28% correct before therapy and 40% correct after therapy. Thus, an overall improvement of 12% resulted from the training program which did not specifically involve articulation training. This finding strongly suggests that respiratory and phonatory control should be established in deaf speakers before training in articulation is initiated.

TABLE 6: Averaged ratings for pitch register and pitch before and after eighteen hours therapy. (N= 10 NTID students)

Subject	Averaged Rating of Pitch* (Rainbow Passage)			Averaged Rating of Pitch Control** (Rainbow Passage)		
	pre	post	diff.	pre	post	diff.
0	2.0	1.8	.2	3.4	3.8	.4
2	2.2	2.2	0	2.8	4.0	1.2
3	1.6	2.8	1.2	4.0	4.0	0
4	2.6	4.0	1.4	2.0	4.4	2.4
5	2.2	3.0	.8	4.0	3.8	-.2
6	1.0	2.8	1.8	3.8	4.0	.2
7	2.0	2.8	.8	2.0	4.0	2.0
8	1.0	4.4	3.4	1.2	4.0	2.8
9	1.0	4.2	3.2	3.2	3.8	.6
10	2.4	4.4	2.0	3.4	3.8	.4
mean	1.8	3.3	1.5	3.0	4.0	1.0

*Pitch Register Rating Scale
1 Much above optimum
2 Moderately above optimum
3 Slightly above optimum
4 Appropriate for age and sex
5 Slightly below optimum
6 Moderately below optimum
7 Much below optimum

**Pitch Control Rating Scale
2 Noticeable breaks or fluctions of large   magnitude
2 Noticeabe breaks or fluctuations
   of large magnitude
3 Noticeable breaks or fluctuations
   of small magnitude
4 Flat within limted speaking range
5 Normal-satisfactory moduation of pitch

63