COMPUTER AIDED SPEECH SHAPING

BY

SHARON L. CAIRNS

A Thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree of

MASTER OF ARTS

Department of Psychology
University of Manitoba
Winnipeg, Manitoba

ISBN 0-315-63259-3

COMPUTER AIDED SPEECH SHAPING

BY

SHARON L. CAIRNS

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

Acknowledgements

This research was supported by a University of Manitoba Fellowship (1988-1989) and a Medical Research Council Studentship (1989-1990).

The author gratefully acknowledges the guidance and support of her advisor, Dr. J. Pear, and the input of her committee members, Dr. W. Kinsner and Dr. G. Martin, University of Manitoba.

Appreciation is also extended to the author's father-in-law, R.D. Fairley, who extended considerable assistance in obtaining a hard copy of the data.

Last, but not least, the author thanks her family, particularly Stephen, for their unending patience and support.

Abstract

Examination of the shaping procedure in human subjects has been limited due to a lack of technology capable of precise measurement of behaviour and the ability to apply specified shaping parameters. This study examines the shaping procedure using a computer to shape vocal imitation of the phonemes /a/ and /e/ in three developmentally handicapped subjects. Results indicate that trial by trial responsiveness of the shaping parameters results in variable responding with no consistent trend towards the target. Maintaining a set criterion over a number of sessions resulted in greater consistency in responding and a gradual movement towards the target. Results also suggest that a reinforcement rate of approximately 50% is desirable in shaping vocal imitation in this population. Thus, small, slow step sizes with high rates of reinforcement appear to be desirable in shaping vocal imitation in developmentally handicapped subjects. Human ratings indicated that all three children showed some improvement in their ability to imitate the trained phoneme. Correlations between the computer and human raters were low and varied with bandwidth restriction, child's voice, and phoneme trained. Although the nature of this study was exploratory, it does demonstrate the potential of a more precise methodology for the study of shaping and applied speech training.

Computer Aided Speech Shaping

Shaping is a behavioral procedure that has been used to develop or train a wide variety of new behaviours in both animals and humans. For example, increasing and decreasing response force (Eisenberger, 1989; Kirkpatrick & Fowler, 1989), shaping response location (Eckerman, Hienz, Stern, & Kowlowitz, 1980), shaping response times (Alleman & Platt, 1973), shaping verbal behaviour (see below), teaching cooperative responses in children (Hingtgen & Trost, 1966), teaching pigeons to play ping pong and to cooperatively mirror each other's responses (Skinner, 1954 and 1962), training performance animals (Skinner, 1954), and teaching crows to use tools (Powell & Kelly, 1975) have been trained using shaping procedures. As defined by Martin and Pear (1988), shaping is "the development of a new behaviour by the successive reinforcement of closer approximations and the extinguishing of preceding approximations of the behaviour" (p. 69). Skinner (1953) likened the shaping of a new behaviour to the work of a clay sculptor. "At no point does anything emerge which is very different from what preceded it." This is true even though the end product is nothing like "... the original undifferentiated lump." (Skinner, 1953, p. 91). Skinner goes on to say that the effectiveness of shaping is related to the identification and utilization of the fact that complex acts are of a continuous nature. Shaping is brought about by a continual process of differential reinforcement. This is more effective than reinforcing only the target behaviour because this behaviour may never occur or may occur so rarely that there is little or no opportunity to reinforce the desired behaviour. The process of shaping also facilitates the optimal strengthening of precursor behaviours that lead up to the target behaviour.

Shaping can be used to change the quality, accuracy, quantity, intensity, timing, and the topography of behaviour. Martin and Pear (1988) suggest a number of steps to successfully shape behaviour: (1) Define the behaviour you want to end up with; (2) Define some behaviour as the beginning point. This must be a behaviour that occurs often enough to be reinforced; (3) When this initial behaviour is occurring at a high frequency, stop reinforcing the initial behaviour and begin reinforcing a slightly closer approximation of the target response; and (4) Outline the successive approximations required to get from the beginning behaviour to the target behaviour. No guidelines are offered for identifying the ideal step size. However, the steps need to be small enough to permit success but not so small that the training period becomes unnecessarily protracted. Move through the shaping steps at a pace that will result in the behaviour of each step being well established but not so well established that new approximations are unlikely to appear. If the behaviour is lost from moving too fast or taking too large of a step, one may return to an earlier approximation and define extra bridging steps. Skinner (1954) recommended that each successive step should be as small as possible to raise the frequency of reinforcement to a maximum and reduce aversive consequences to a minimum.

Although shaping is a widely practised technique for teaching new behaviours, there have been few systematic examinations of the procedure. Alleman and Platt (1973) stated that the application of shaping is a vague art form which is often dependent on the trial and error skills of the technician. Part of the problem is related to vaguely defined response dimensions and shaping parameters which prevent replication thus limiting the growth of science. Alleman and Platt advocate the use of percentile reinforcement schedules which specify a percentage of values which must be exceeded to produce reinforcement while controlling for the probability of reinforcement. This schedule allows for contingencies to be defined in terms of current behaviour which decreases the possibility of the behaviour losing contact with the contingencies. While using this method to shape response rates in laboratory animals, they found a large step size to be more efficient than a smaller step size. Eckerman et al. (1980) also found that large rapid shaping steps maximized shaping of response location in pigeons. These authors suggest that extinction of a previous step leads to increased variability in responding which increases the probability of behaviour in the desired direction. In a later examination of percentile schedules, Davis and Platt (1983) found that the shaping of response location occurred whether a fixed criterion schedule or a targeted percentile schedule was used.

A recent study by Midgley, Lea, and Kirby (1989) demonstrated that shaping rats to deposit a ball bearing down a hole using a shaping algorithm was more consistent than hand shaping and resulted in higher levels of reinforcement, making periods of extinction less likely. These authors stipulated a set of rules as to which responses would be reinforced and when they would be reinforced according to a specified algorithm. They found that using this algorithm resulted in within session movement both up and down the hierarchy of shaping steps. The authors suggest that a simpler algorithm than the one they used might be better but at the same time warned that a straight percentile reinforcement would not be wholly satisfactory due to the frequency with which time based backtracking was required. Although this work is replicable and sheds some light on the shaping process, it is still limited in its ability to inform us about shaping parameters such as step size due to the use of a heterogeneous behaviour sequence rather than a response continuum. However, in this regard, this study chose shaping steps that required the occurrence of all lower level steps which is an improvement over heterogeneous behaviour sequences which do not necessitate all previously learned behaviours to be performed with each new approximation of the target behaviour.

To bring the shaping procedure into the realm of science, one needs to define a set of behaviours which can be precisely measured by technology on an interval continuum, have movement along that continuum controlled by a specific set of rules, and conduct the procedure in a manner that is replicable. One study meeting these criteria was a pigeon study in which the birds were trained to contact an arbitrarily defined spherical region (Pear & Legris, 1987). In this study, precise procedural specification of the shaping procedure was conducted with a computer controlled system which continually tracked the pigeon's head.

Martin and Pear (1988) suggest that normal acquisition of language occurs through a shaping process with a progression from babbling to baby talk to words in the child's native language. A simplified example of this process would be for parents (or caregivers) to reinforce the sounds "mmm" and "daa". Over a period of time these sounds would be placed on extinction and only "ma-ma" and "da-da" would be reinforced. Finally, only "mommy" and "daddy" would be reinforced and all earlier approximations placed on extinction.

Shaping is a common approach to teaching language to speech deficient individuals with vocal imitation being the first step. To shape speech, one begins by reinforcing a vocalization. Gradually, the utterance must more closely resemble the target sound modelled by the teacher until the student is only reinforced for producing the desired sound. This process is repeated with other basic vocal sounds and when a number of them have been learned, they are combined to form words. Once verbal imitative behaviour has been trained, functional and spontaneous speech across a variety of settings and persons can be taught (Garcia & DeHaven, 1974).

Early speech shaping studies used large step sizes which were dependent on human judgement. For example, Hingtgen and Trost (1966) shaped vocal responses in early childhood schizophrenics with the following steps: (1) make any sound including humming, coughing, sneezing, or giggling; (2) make "a more discrete sound"; and (3) emit recognizable syllables such as ah, ba, uh, ta, etc. Isaacs, Thomas, and Goldiamond (1966) used the following four steps to reinstate verbal behaviour in mute psychotics: (1) eye movement; (2) lip movement; (3) vocalization; and (4) a vocalization approximating the word "gum". Shaping was also successfully used by Panyan and Hall (1978) to teach vocal imitation to two severely retarded females. Due to limitations of human raters, the shaping procedure was broken down into only gross steps (i.e. correct imitation of tongue placement and mouth position, production of any vocalization, and imitation of the complete sound). These large steps can lead to frustration on the part of the trainee as there would be no distinction made between a completely incorrect vocalization and one that begins to approximate the target sound. Although these studies demonstrated success with the shaping procedure, it would be almost impossible to adequately replicate them, and even if this was possible, it would add little to the scientific understanding of the shaping procedure.

One speech shaping study which used specific criteria for shaping on an interval dimension involved shaping the length of saying the phoneme /u/ in college students (Lane, Kopp, Sheppard, Anderson, & Carlson, 1967). These authors instituted phase changes following 10 consecutive reinforced responses. They found that if the initial probability of responses selected for reinforcement is too low, shaping fails. If the probability of reinforcement is high (small shaping steps), shaping occurs but large changes are accomplished slowly and inefficiently.

Shaping the imitation of speech sounds and words can be seen as a first step in acquisition of useful speech. Yoder and Layton (1988) found that verbal imitation ability positively predicted the size of child-initiated spoken vocabulary acquired during training in a sample of 60 autistic children. In this study, the children with the higher verbal imitation scores pre-treatment were the ones who used more spontaneous words regardless of speech training method ("Speech only", "Simultaneous speech and sign", "Alternating speech and sign", and "Sign only"). Gaines, Leaper, Monahan, and Weickgenant (1988) also found that children with good vocal imitation were more likely to learn words either alone or with signs. Further support for the importance of teaching vocal imitation occurs in a study by Remington and Clark (1983) who found that following expressive sign training (using simultaneous speech and sign training), improvement in receptive speech was only evident in the child who was capable of verbal imitation.

It is important to begin vocal imitation at an early age since it has been found that younger children were more likely than older children to retain language which has been trained when tested at six month follow-up (Gaines et al., 1988). Due to limited human resources, an apparatus which could shape vocal imitation would be a valuable asset in the field of language training. Speech development is a long and tedious process often requiring months and years of training (Garcia & DeHaven, 1974). A computer is infinitely patient, is more accurate, reliable, and unbiased, and can eliminate repetitive training which would otherwise tie up a speech therapist's valuable time. One of the most valuable advantages a computer has over human speech trainers is the fact that the operations are repeatable. Another good reason to introduce computerized speech shaping is that it has been found in normal educational settings that automated instruction decreased the amount of learning time required by 20 to 40% (Kearsley & Seidel, 1985). If this finding could be generalized to a developmentally handicapped population, this would represent significant time savings.

One possible objection to the use of a computer in the shaping of vocal sounds may be the lack of the visual stimuli produced by lip movements. However, in studies of simultaneous communication training (signing and vocal) by Remington and Clark (1983) and by Carr, Binkoff, Kologinsky, and Eddy (1978), lip reading was not a salient variable in language acquisition for autistic children. However, if visual stimuli are desired, they can be added by video presentation which could be controlled by the computer.

Pear, Kinsner, and Roy (1987) have developed an apparatus which is designed to automatically shape sounds and which provides for precise measurement of the shaping process. This apparatus provides for the fulfilment of Martin and Pear's (1983) recommended guidelines for the effective application of shaping. The terminal behaviour is clearly identified and remains consistent from session to session since it is stored on a computer disk. The starting point is also clearly identified and a vocalization in response to "Say (sound)" can be precisely defined on an interval scale in terms of approximation to the desired behaviour. During the process of shaping, successive approximation criteria can be controlled by the computer in an exact manner not possible with human speech shapers. Finally, the shaping program is set up to be flexible depending on the student's progress. If the student's responses are correct for any given step, the computer automatically proceeds to the next shaping step. If the student's responses are incorrect, the program returns to an earlier step reducing the likelihood of extinguishing responding. The program is set up so that the speech trainer can determine both step size and number of correct/incorrect responses before the shaping step is advanced/ regressed.

In previous research, this apparatus has produced a slight downward trend in distance between target and response in vocal shaping and also produced moderate correlations with professional raters. Previous research (Desrochers, 1989) also suggests that use of this apparatus to shape speech may be more effective than human shaping since it is able to rate proximity to the target sound on a continuum whereas human raters tend to rate speech in a dichotomous manner (either like or unlike the target sound). It was expected that improvements in the shaping program would produce even better results.

The main purpose of this research was to examine the vocal imitation shaping methodology using the above apparatus for speech-deficient developmentally handicapped children. The methodology was examined using this apparatus because of its precision and the repeatability of the procedure. Since the majority of previous work involving examination of ideal shaping parameters has involved laboratory animals, it was expected that this research would shed some light in this regard on shaping the acquisition of vocal imitation in a human population. Training and control sounds of /a/ and /e/ were used as these are the phonemes that have previously been found to have the highest correlations between the computer and human ratings (Pear, Kinsner, & Roy, 1987).

One of the potential criticisms of much of the operant research is that when stimuli are chosen as reinforcers for a particular subject, the stimuli are not systematically tested to ensure that they are indeed good reinforcers. In a review of the 1986 Journal of Applied Behaviour Analysis, it was found that only three of 44 studies reported a systematic method of reinforcer selection (Mason, McGee, Farmer-Dougan, & Risley, 1989). To address this issue, the current research incorporated systematic identification of reinforcers and conducted a test of them. This procedure was derived from the work of Pace, Ivancic, Edwards, Iwata, and Page (1985) and Green, Reid, White, Halford, Brittain, and Gardner (1988) who developed and tested methods of identifying reinforcers. In addition, following on the work of Mason et al. (1989), this research used a daily presession mini-assessment to allow the child to choose what will be used as a reinforcer for that session.

Method

Subjects

The three participants were developmentally handicapped children who lived in a residential treatment facility, the St. Amant Centre, in Winnipeg, Manitoba. All three were assessed as capable of making vocal sounds but were unable to imitate both of the sounds /a/ and /e/. The children had no physical deformities that would prevent them from being able to emit the target sounds. Since the children were not capable of volunteering to participate, their parents were informed of the nature of the study and parental consent was obtained.

Amy was 4 years, 8 months old at the start of the study. She was diagnosed with spastic cerebral palsy with severe delayed development. At age four, she was assessed as functioning at the 12-18 month age level and was capable of babbling but was unable to imitate sounds.

Brian was 3 years, 10 months of age at the beginning of the study. He was diagnosed with profound mental retardation, spastic quadriplegia, cortical blindness, and refractory seizures secondary to post natal apnea. He was capable of emitting the sound "da da" spontaneously but did not imitate any other speech sounds.

Carol, who was 9 years, 8 months old at the outset was diagnosed with a ring chromosome defect of chromosome 4, microcephaly, seizure disorder and severe retardation. At age 9, she was assessed as functioning at the one year level. Carol was capable of emitting vocal sounds but did not imitate any of the sounds she was capable of producing.

Two other children had been started in this experiment but were discontinued at different points in the study. One child was discontinued early in the project due to not making any vocal responses to the computer command "Say [sound]" during five consecutive baseline sessions. The reason for this is not known since he did respond to the computer during the preliminary assessment. Another subject was discontinued during the training phase due to an increase in screaming and non-responding during the sessions. As this child was capable of functioning at a higher level than the other children used in the study and usually tried to please anyone he was working with, it was hypothesized that his apparent growing dislike for sessions was related to frustration that his spasticity prevented him from vocalizing the requested sound (this child suffered from severe spastic cerebral palsy).

Apparatus

An Apple IIe microcomputer with attached microphone, speaker, and double disk drive was used to train and record speech. A speech recognition card based on a SP1000 signal processor was used to analyze the quality of the speech utterances. A stereo frequency equalizer (Sound Shaper Two) was used to modify tape recordings of the vocalizations to restrict the bandwidth from 300 to 3000 Hertz. Special software controlled assessment, baseline, and training sessions. A User's Guide (Cairns, 1989) providing detailed information on use of this software is contained in Appendix A. Data was stored on floppy disks. A VHS video cassette recorder, camera, audio tape recorder, microphone, and tapes were used for procedural reliability and social validity.

To keep the microphone a constant distance from the child's mouth for accurate computer assessment of the sound, the child wore a set of headphones (similar to a Walkman radio) with a small microphone attached for recording the child's responses.

The computer stores reference sounds as a trajectory of a set of reflection coefficients and a logarithmic measure of voice energy, changing during the utterance and obtained using real-time linear predictive coding (LPC) of speech (Kinsner, Pear, & Roy, 1986). The subject's response is analyzed by the computer and assessed by the same characteristics of speech. The distance between the target and trial trajectories is measured using the second metric norm (range is 0 to 256). The second metric norm is defined as:

$$D = \sqrt{\sum_{j=1}^{n} (r_j - t_j)^2}$$

where r = the reference trajectory, t = the trial trajectory, and j= the reflection coefficients. The second metric norm is used because its reflection of energy more closely resembles the operation of people. However, future research should examine which norm is best suited to the shaping of speech sounds.

The reference sounds consist of a cluster of 12 templates made by six individuals of varying age and sex (2 children, 2 adolescents, and 2 adults uttering each sound twice). There were two reference clusters, one for /a/ and one for /e/. A template is defined as a vector containing the unique aspects of a spoken utterance in the form of LPC-10 (with 10 utterances in 0.2 sec.). Each response by the child was assessed against each template in the cluster and the distance score given by the computer was the lowest score obtained. To ensure that quality sounds were used for the reference templates, a speech therapist (M.Sc. Speech and Language Pathology) assessed if the sound produced by the models was the desired one. This speech therapist worked at the St. Amant Centre and volunteered her time. This assessment was conducted as the templates were being made. Templates that were judged to be of poor quality were rerecorded until the speech therapist was satisfied with the quality.

The computer also provides verbal reinforcement consisting of the word "Good" if the child's response is within the reinforcement range and is capable of activating a sensory reinforcer such as a video or audio tape. The reinforcement range is defined as a range of distance scores between 0 and any given criterion. A distance score is the difference between the reference and trial trajectories.

Procedure

The study involved four phases including initial assessment, reinforcement testing, baseline, and training. An experimenter was present at all times controlling the operation of the apparatus.

Baseline and training sessions used the phoneme /a/ for Brian and Carol, and the sound /e/ for Amy. The sound /e/ was used as a control sound for the children being trained with /a/, and /a/ was used as a control for the child trained with /e/. Control probes were presented once a week during baseline and training.

Assessment The synthesized auditory instruction "Say [sound]" was presented to the child. The computer scored and recorded the child's response according to the distance of the child's sound from the target sound. Five trials were presented of each sound (/a/ and /e/) and every vocal response was reinforced with praise (Good boy [girl], that's right. Very Good.) and either an edible or sensory stimulus identified as reinforcing to the child by the caregivers. To be included in the study, the child's response to each trial had to be unlike the target sounds.

Reinforcement Testing Prior to commencement of the study, the children selected to participate were tested for reinforcement efficacy using a procedure derived from the work of Pace et al. (1985) and Green et al. (1988). The reinforcement assessment consisted of two phases:

systematic identification and reinforcer test. Staff working with each child were asked what they thought was reinforcing for the child. These items and others, to make a total of 6 to 17 potential reinforcers for each child, were tested. (More items were tested for the children who approached few of the potential stimuli.) Potential reinforcers tested included both edibles and sensory stimuli. Stimuli used and method of presentation for each child are listed in Table 1.

During systematic identification, the children had from three to nine sessions of 20 trials each. Sessions were conducted until a minimum of two stimuli were identified that the child approached consistently. At the beginning of each of these sessions, the child was encouraged to sample each of four potentially reinforcing stimuli. If the child made an approach response, the stimulus was presented for another 5 seconds. If the child made an avoidance response, the stimulus was removed. An approach response was defined as movement toward the stimulus, maintenance of contact with the stimulus for 3 seconds, positive facial expression, or positive vocalization within 5 seconds of presentation of the stimulus. For Brian, approach also included quiet attending behaviour. An avoidance response was defined as a negative vocalization, pushing the stimulus away, or movement away from the stimulus. Following the

Table 1

Presentation of Stimuli During Systematic Identification

Stimulus	Approach	Presentation
		Amy
Praise	90%	"Good girl Amy. Very good."
Teddy Ruxpin	100%	Battery operated bear was activated and placed on the child's tray.
Shoulder Rub	100%	Experimenter rubs top of child's shoulders with her hands.
Radio	80%	Transistor radio turned on to a music station. Experimenter moved head with the music and smiled.
Elephant	80%	Elephant shaped rattle shaken by experimenter in front of child.
Ice Cream	40%	Spoon with vanilla ice cream placed at child's lips.
		Brian
Radio	90%	Same as for Amy.
Chocolate	20%	A piece of chocolate chip was placed at the child's lips with a spoon.
Music Clown	80%	Battery operated clown was activated and moved back and forth in front of the child.
Elephant	10%	Same as Amy.
Hooray	100%	Experimenter clapped hands and said "Yea Brian, hooray, yea Brian".
Kangaroo	0%	Plastic kangaroo that squeaked when squeezed was demonstrated by experimenter and placed on tray in front of the child.
Orange Drink	10%	Orange drink in a glass was put to the child's lips.
Toucan	80%	Stuffed toucan was moved in front of the child while experimenter said "Ooo oo ooo, oo oo ooo."
Bell	50%	School type bell was rung by Experimenter and placed on the table in front of the child.
Truck	90%	Plastic truck placed on tray in front of child.
		Carol
Radio	80%	Same as for Amy.
Elephant	0%	Same as for Amy.
Shoulder Rub	0%	Same as for Amy.
Sweet water	20%	Glass with sweetened water brought to child's lips.
Play with hands	30%	Experimenter manipulated child's fingers.
Music Clown	30%	Same as for Brian.
Water	50%	Same as for sweet water but using plain tap water.
Doll	30%	Soft doll placed on child's tray.
Bell	0%	Same as for Brian.
Toucan	0%	Same as for Brian.
Clock Phone	40%	Fisher Price toy manipulated by experimenter and placed on child's tray.
Smurf	10%	Stuffed toy placed on child's tray.
Rolly Clown	30%	Plastic clown with bells in it placed on child's tray.
Spoon	0%	Metal spoon placed in child's hand.
Pot	10%	Small blue coloured aluminum pot placed on child's tray.
Patty Cake	40%	Experimenter clapped hands with child and sang "Patty Cake".
Plastic Ring	0%	Brightly coloured plastic ring placed on child's tray.

reinforcement sampling, the four items were presented in a random order, 5 times each and approach responses were recorded. Once again approach resulted in a further 5 seconds of exposure to the stimulus and avoidance resulted in the stimulus being removed. The children were thus exposed to each of the potential reinforcers a total of 10 times.

The reinforcer test involved a minimum of one session for each potential reinforcer to which the child exhibited an approach response on at least 50% of presentations during the systematic identification phase. Sessions were repeated with a given stimulus if reinforcer efficacy was not clear. Each of these sessions started with 10 baseline trials in which the child was asked to perform an operant behaviour which was known to be within the child's repertoire. The operant behaviours used during the reinforcer test were "Touch the red square", imitate knocking on the child's tray, and "Look at me" for Amy, Brian, and Carol respectively.

During baseline the experimenter said "Good" following each correct response as the only consequence. This was followed by 10 trials with reinforcement for performing the requested behaviour. Reinforcement consisted of the stimulus being tested and the experimenter saying "Good girl [boy]. That's right. Good for you [child's name]." If the child did not perform the response on the first request for each phase, a prompt was given to the level required for the behaviour to occur, and the response was reinforced. That is, if the child did not perform the behaviour with a verbal prompt, the verbal prompt was repeated with the experimenter modelling the response and then repeating the verbal prompt. If the child still did not perform the behaviour, a physical prompt was provided. Only until the child had received the first reinforcement were prompts given. The next phase was 10 to 20 trials of return to baseline. For Amy and Carol, this return to baseline was the last phase for each session. For Brian, a second reinforcement phase was included in order to obtain a clearer picture of his response patterns.

Reinforcers used in the study were ones to which the child exhibited a high degree of approach behaviour (approached a minimum of 50% of presentations) and increased the performance of an operant behaviour over baseline by 20% or more. Each speech session began with a reinforcement preference test in which the child was presented with a variety of reinforcers as identified above, and the one selected by the child was used for that session.

Inter-Observer Reliability A person other than the experimenter viewed and rated a randomly selected 25% of the video recordings of the reinforcement assessment sessions to rate the approach or avoidance to the stimuli and the performance of the operant. IOR's calculated on these sessions using the formula:

$$\frac{\text{Agreements}}{\text{Agreements + Disagreements}}$$

indicated a high degree of inter-rater agreement, ranging from 95.7% to 100% with a mean agreement of 97.8%. The only procedural error observed involved one less trial being administered during systematic identification than was prescribed in the procedure.

Baseline

At the beginning of all speech sessions, the children were presented with each of the stimuli identified as reinforcing to that child. The stimulus to which the child exhibited the most positive response (on a continuum of laughing, babbling, smiling, making contact with the stimulus, and quiet attending with the former considered more positive than the latter) was selected as the reinforcer to be used for that session. The reinforcers were counterbalanced in terms of which was presented first for each of these mini-assessments.

The synthesized auditory instruction "Say [sound]" was presented to the child. The child's response was scored and recorded by the computer according to the distance from the reference templates. Vocal responses to the computer instruction were reinforced on a fixed ratio schedule with praise ("Good boy [girl] [child's name], that's right, very good.") and the reinforcer chosen by the child at the beginning of the session. To establish a high rate of responding to the computer instruction ["Say (sound)"], the reinforcement schedule was gradually leaned from continuous to FR2 to FR3. (That is, reinforcement was leaned from every response to every second response to every third response.) The requirement for changing the schedule was that the child give a vocal response to the stimulus for 10 consecutive trials. If the child's response reached the target, reinforcement was given regardless of the point in the FR schedule. (This occurred for Amy only.) The target was defined as a low (< 80 for /a/ and < 110 for /e/) distance score between the trial and model trajectories as measured by the computer. These targets were determined by the average score (rounded up to the nearest 10) obtained by the experimenter emitting each of the target vowels over 40 trials. Baseline sessions consisted of a maximum of 40 trials (15 to 25 minutes). A session was terminated if the child did not give a vocal response for five consecutive trials. Baseline was continued until stability was reached according to the criteria outlined by Sidman (1960, p. 260). That is, the difference between the means of the first three of the last 5 (or 6 in the case of more than 5 baseline sessions) sessions and last two (or three in the case of more than 5 baseline sessions) could differ from the grand mean of these sessions by no more than 7%. Sidman suggested 5% for laboratory experiments with animals. It was decided to relax this criterion given the exploratory nature of the current work. This criterion was not adhered to for Brian as his data points showed an increasing trend in the direction opposite to that anticipated during training.

<u>Training</u>

The average score from the last 10 baseline trials was used as the starting rejection level which was maximum distance score which would receive reinforcement for the first training session. The reinforcement range was defined as the range of distance scores from zero to the rejection level for which the child received reinforcement. After the first training session, the average of the preceding 10 trials was used as the starting rejection level for each subsequent session. If the child's response to the computer instruction "Say [sound]" was within the reinforcement range, reinforcement (same as baseline) was given and the range for reinforcement decreased by three (moving closer to the target). If the response was out of the reinforcement range, the computer emitted a small beep and the trial was re-presented with the reinforcement range increased by one (moving farther away from the target sound). Training sessions also consisted of 40 trials and were terminated if no vocal response occurred for five consecutive trials.

Two additional training procedures were used due to variable responding under the initial training conditions.

Amy was given a second training phase in which a correct response resulted in the reinforcement range decreasing by three and two incorrect responses had to be emitted before the reinforcement range would expand. This phase was introduced in an attempt to exert more pressure on child's responding to shape in the desired direction.

All children were also exposed to changing criterion training conditions (Kazdin, 1982) due to an apparent lack of effect from the trial by trial shaping adjustments. In this phase, the rejection level was set at a value which remained constant throughout an entire session. For Amy and Carol's first changing criterion phase, the rejection level was set at a value that would have resulted in the child receiving approximately 30% reinforcement in the previous session. This rejection level was maintained until the child had three consecutive session means within five points of the criterion. When this condition was met, the rejection level was decreased by five. For the girls' second changing criterion phase, and Brian's only one, the initial rejection level for the phase was determined by a rejection level that would have resulted in 50% reinforcement in the previous session. The criterion was made more stringent when a child had three consecutive sessions with means at or below the rejection level. When this condition was met, the criterion was tightened by two to a maximum of five points depending on the level that would have resulted in 50% reinforcement had that level been used in the previous session. The criterion was made less stringent if the child had a session which resulted in no reinforcement or if the session means showed an upward trend for three consecutive sessions. When a criterion was being relaxed, the new rejection level was set at a value that would have resulted in 50% reinforcement in the previous session.

Control Probes Approximately once every five sessions during baseline and training, probes were conducted for both /a/ and /e/ to determine whether the sound being trained was showing improvement relative to the sound not being trained. The sound which was presented first was alternated between probe sessions. During probe sessions, each sound was presented by the computer three times and the child's response recorded. No reinforcements were given during the session but at the end of the session the child was reinforced for his/her participation. On days that control probes were conducted, no other session took place.

Social Validity At the end of the study its social validity was evaluated by two speech therapists. One had an M.Sc. in Speech and Language Pathology, and the other had a B.A. in Speech and Hearing Sciences. Both were employed as speech therapists at the St. Amant Centre and, therefore, had experience working with a developmentally handicapped population. Both therapists were paid for their time. They were given six coded (to mask training sequence) audio recordings from each of the children in the study. Three of these were straight recordings of the sessions (each child's worst, average, and best sessions as determined by session means). To be chosen for this assessment, a session had to have a minimum of 20 responses. The other three recordings were of the same sessions as above but the tapes had been put through a bypass filter which restricted the bandwidth from 300 to 3000 Hertz. Restricting the bandwidth makes the quality of sound received by the human raters similar to the information received by the computer. These tapes were coded to give the speech raters the impression that they were rating six sessions from each child in an attempt to avoid any biasing of effects. The speech therapists were asked to rate the sounds produced in terms of their proximity to the target sound according to the following rating scale: 1 = Matches target sound; 2 = Close to target sound; 3 = Some similarity but still unlike target sound; 4 = More unlike than like target sound; and 5 = Totally unlike target sound.

When the data was returned to the experimenter, one speech therapist indicated that she interpreted the scale as: 1 = correct; 2 = close approximation; 3 = gross approximation; 4 = any vowel type sound; and 5 = any other sound (vegetative or consonant-type).

The purpose of the speech rating was two-fold: (1) to determine if human raters believed that vocalizations were being shaped in a desirable direction, and (2) to determine if the human raters' judgements more closely approximated the computer under restricted or unrestricted bandwidth conditions. Previous moderate correlations between the computer and human ratings may be related to the difference in quality of speech given to the computer (i.e. restricted bandwidth) compared with the range of sound available to the human ear. Bell, Dirks, and Carterette (1988) found that error patterns in understanding speech sounds were significantly affected by an interaction of presentation level, bandwidth filtering, and positioning of consonants.

Inter-Observer Reliability As the computer objectively scored all responses from the speech shaping procedure, formal IOR's were not required.

Results

Reinforcer Assessment

As shown on Table 1, the children exhibited varying levels of approach to the stimuli presented. Carol demonstrated low approach levels to a large number of stimuli even though the ward staff had indicated her liking of a number of these potential reinforcers.

Figure 1 shows the results of the Reinforcer Test with the stimuli to which the children showed the highest degree of approach. From these results, it was decided to use the radio, shoulder rub, and elephant as reinforcers for Amy; the radio, hooray, and truck for Brian; and the radio and water for Carol. During the mini-assessments at the start of each session, Amy showed the most positive response to the shoulder rub 42.1% of the time, to the elephant 34.2% of the time, and to the radio 23.7% of the time. Brian "selected" both hooray and the radio 43.3% of the time, and the truck 13.3% of the time. The lower rate of selection of the truck is consistent with this stimulus being the least effective reinforcer of the three as shown in Figure 1. Carol responded the most positively to the radio 68.5% of the time and to the water 31.5% of the time. Although this difference would not have been predicted by the results shown in Figure 1, it is consistent with the approach level during systematic identification indicated in Table 1. Carol approached the radio 80% of presentations compared to 50% of the presentations of water. The order of presentation of the stimuli during the mini-assessment did not appear to affect the "choice" of reinforcer.

Speech Training

The session mean scores, rejection levels, and ranges of scores for Amy are shown in Figure 2. Means and standard

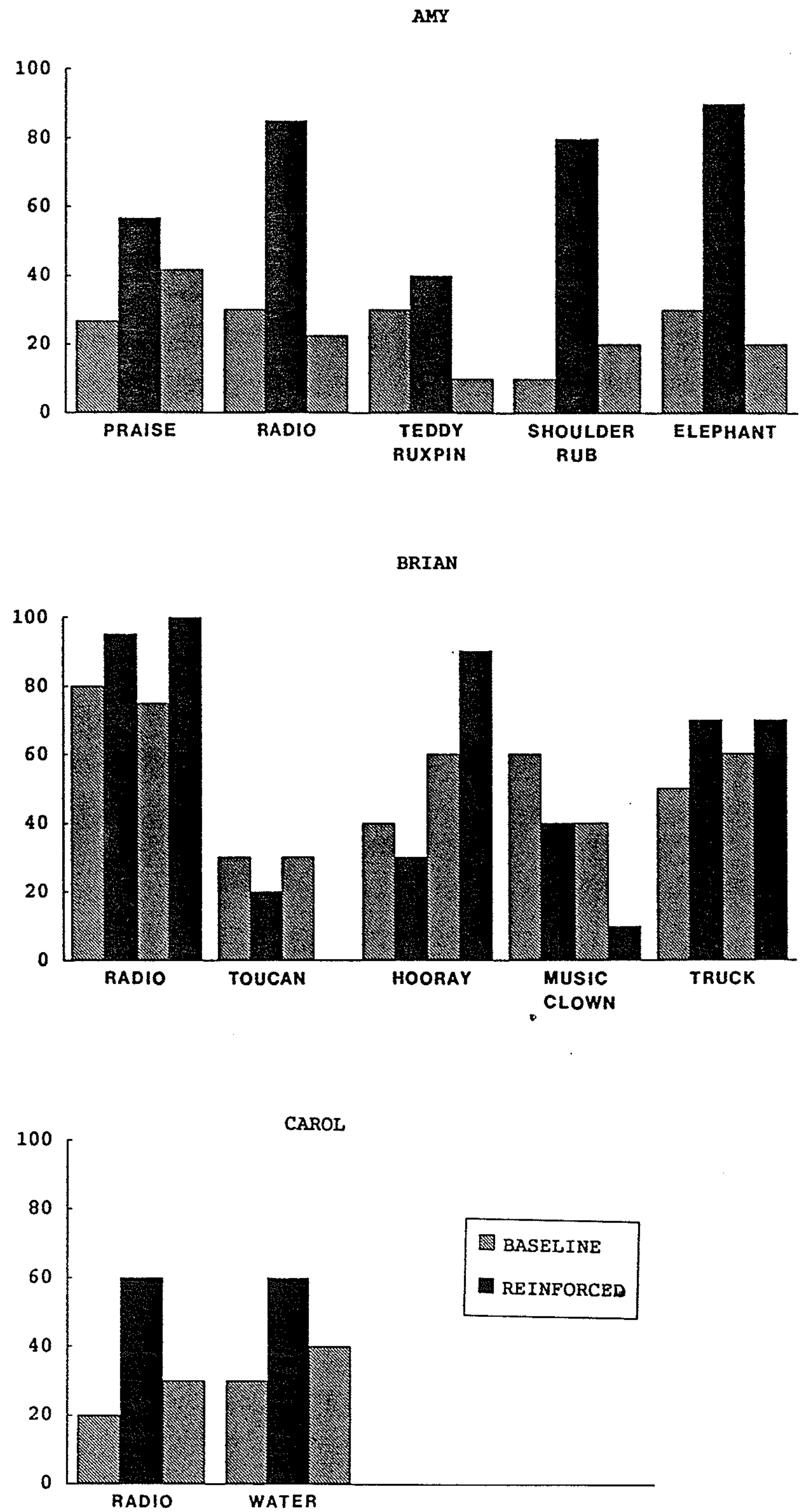

Figure 1. Percentage of trials on which each child performed the requested behaviour under baseline and reinforced conditions during the Reinforcer Test.

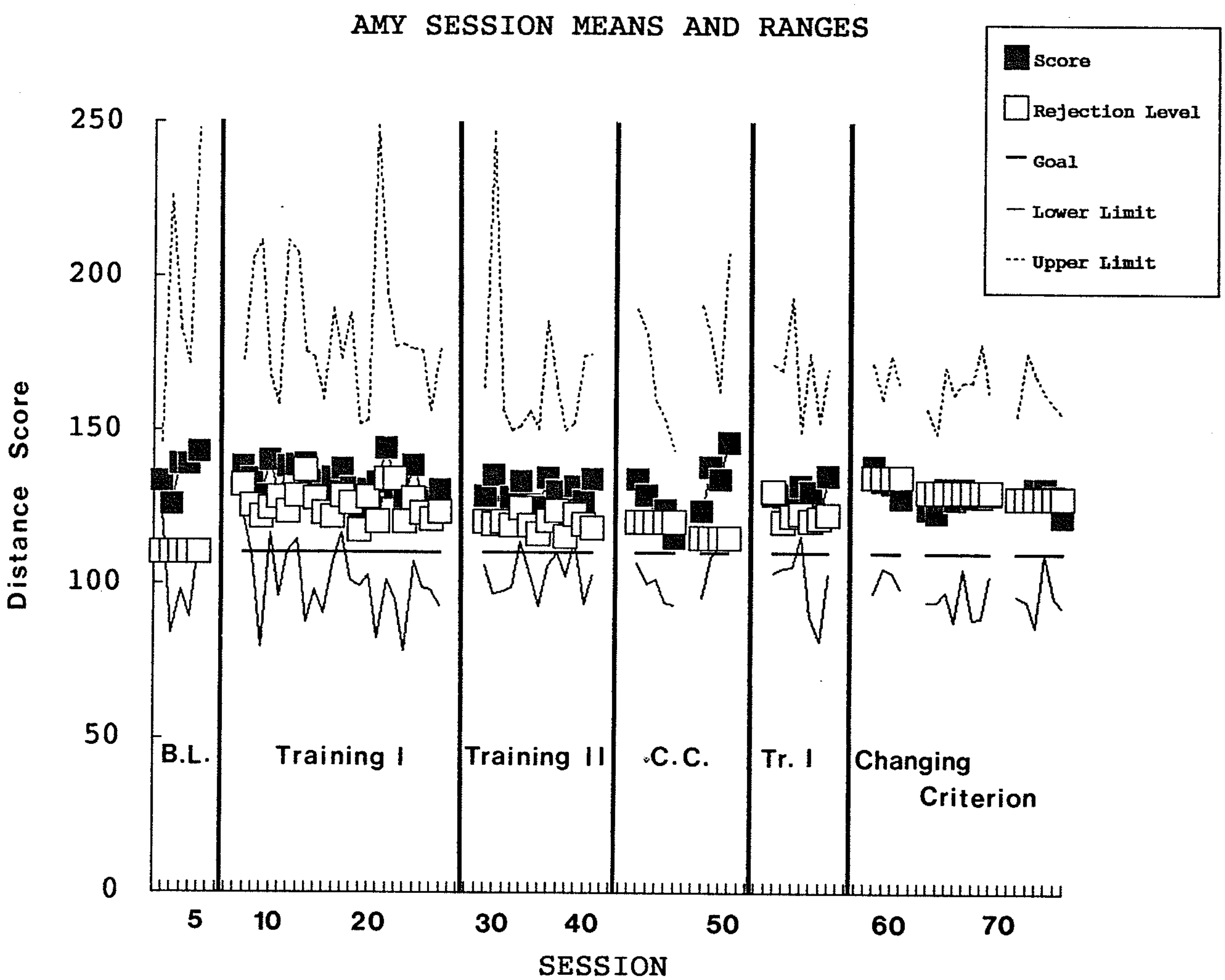

Figure 2. Session mean distance scores, rejection levels, and range of scores for Amy during Baseline (B.L.), Training I, Training II, and Changing Criterion (C.C.) phases.

deviations for these sessions are shown in Table 2. Satisfying the stability criteria outlined earlier, the difference between the mean of the first three baseline sessions from that of the mean of the last two baseline sessions was within 6.0% of the grand mean. There was also an increasing trend in session means during baseline. At the grossest level of analysis, the mean of the baseline phase was 136.1 compared to 129.4 on the last Changing Criterion phase and the mean of the last session conducted which was 122.3. Thus, there is some evidence of shaping having taken place.

During Training I and Training II, session means were variable. However, over these phases there was a downward trend with phase means of 136.1, 133.3, and 128.9 for Baseline, Training I, and Training II respectively indicating that shaping towards the goal was occurring. During subphase 1 of the first Changing Criterion phase, there was a consistent downward trend (indicating improvement) in sessions means. However, in the second subphase when the rejection level was made more stringent, there was an increasing trend indicating that this shaping step was too large for Amy and control over the behaviour was lost. The conditions of Training I were then reinstated to regain control of the behaviour before a second Changing Criterion Phase was introduced. As can be seen in Figure 2,

Table 2

Session Means and Standard Deviations for Amy

Session	Mean	S.D.
	Baseline	
1	133.6	7.9
2	125.7	34.1
3	139.2	21.0
4	139.1	21.1
5	142.9	30.6
Mean	136.1	23.0
	Training I	
6	138.3	13.2
7	134.4	19.6
8	128.4	29.6
9	140.4	12.6
10	127.8	15.6
11	138.2	22.4
12	138.5	27.7
13	138.9	25.2
14	133.2	18.2
15	128.5	14.3
16	134.7	18.5
17	137.3	14.5
18	130.9	17.8
19	124.2	16.0
20	130.8	25.7
21	133.1	40.4
22	144.3	22.8
23	131.1	19.8
24	128.2	22.0
25	138.5	18.3
26	127.9	16.6
27	126.9	13.4
28	130.6	18.8
Mean	133.3	20.1
	Training II	
29	128.5	15.7
30	135.5	30.7
31	119.4	17.6
32	128.1	12.8
33	133.3	10.5
34	125.7	13.4
35	124.9	15.8
36	134.5	16.0
37	130.0	12.2
38	122.5	13.0
39	131.7	9.3
40	127.0	19.0
41	134.0	17.9
Mean	128.9	15.7
	Changing Criterion I	
42	134.1	19.4
43	129.1	17.5
44	120.7	15.6
45	124.2	14.7
46	115.1	14.6
47	123.7	16.1
48	138.0	18.7
49	134.1	12.5
50	146.2	21.0
Mean	129.5	16.7
	Training III	
51	128.9	15.8
52	128.6	14.5
53	129.3	18.7
54	132.2	10.1
55	130.1	17.8
56	121.7	17.6
57	135.3	14.8
Mean	129.4	15.6
	Changing Criterion II	
58	138.4	14.9
59	133.7	14.9
60	132.9	15.5
61	128.1	15.3
62	124.4	16.1
63	123.5	16.1
64	131.5	16.6
65	127.7	14.1
66	131.3	13.9
67	128.8	18.0
68	128.8	16.1
69	129.6	13.8
70	128.1	13.9
71	127.8	22.2
72	131.6	20.4
73	131.6	16.0
74	129.0	14.5
75	122.3	15.5
Mean	129.4	16.0

during this phase session means followed the criterion quite closely with a gradual shaping toward the goal occurring. The range of scores across sessions was also less variable in this final phase than in previous phases.

Sample individual session data showing Amy's worst, average, and best sessions are displayed in Figure 3. Changing Criterion 9 (session 50) was the last session of the first Changing Criterion phase when the criterion was too stringent. Training 18 (session 23) shows the trial by trial adjustments in the rejection level according to the trial score. In this session, the range of scores is lower than in Changing Criterion 9 and a few trials were at or below the target. Changing Criterion 18 (session 75) again shows a lower range of scores with more trials at or below the target.

Figure 4 shows the percentage of trials reinforced, along with the mean distance scores and rejection levels, for Amy across the various phases of the study. The average percentage of trials reinforced during training was 37.4 for Training 1, 30.3% for Training II, and 35.5% for the second Training I phase. At the beginning of the first C.C. phase, reinforcement levels increased, but then dropped off to 0. During the final C.C. phase, Amy's average reinforcement level was 55%. Given that this is the only phase where control over the behaviour is demonstrated, it

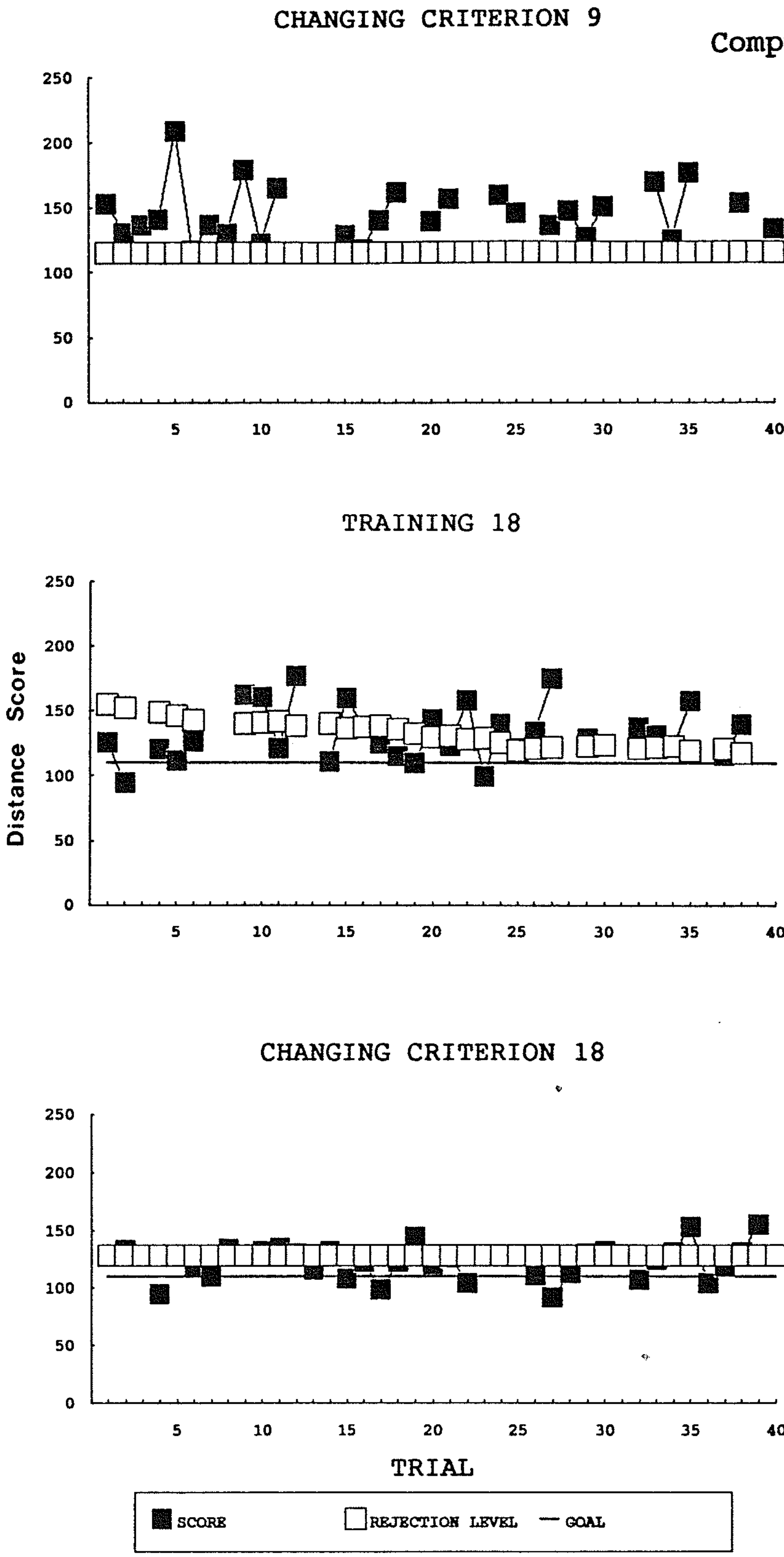

Figure 3. Sample sessions completed by Amy showing sessions with the highest, average, and lowest means (from top to bottom) and the training phase from which they were taken.

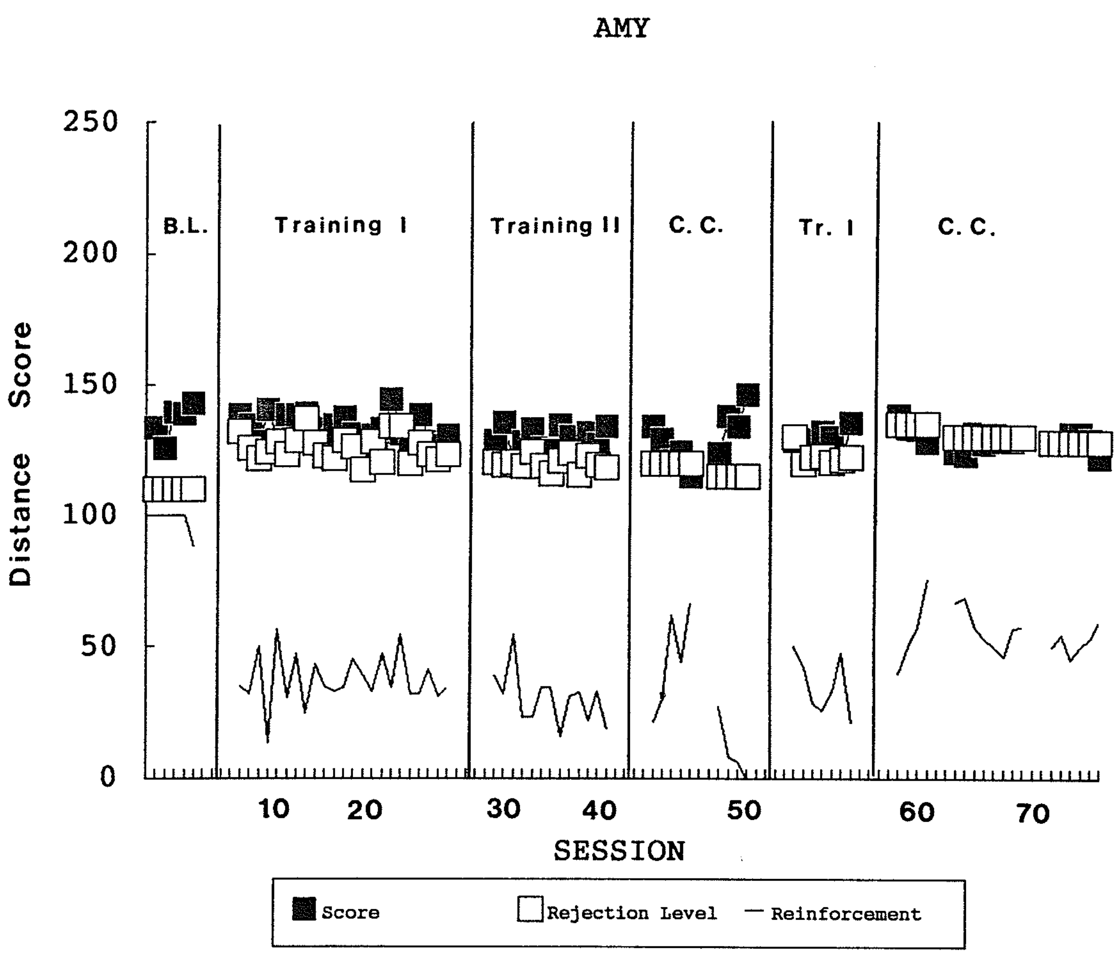

Figure 4. Percentage of trials reinforced compared to mean distance score and rejection level across sessions for Amy.

seems likely that this was a good rate of reinforcement for her to be shaped with.

Amy was the only child able to obtain scores at or below the target level distance scores. Throughout the various phases of the study, the average number of trials per session below this goal showed an increasing trend from Baseline (1.2) to the final Changing Criterion phase (3.0). The only exception to this trend was a decrease from 3.1 during the first Changing Criterion Phase to 2.3 in the second Training I phase. This is consistent with the Changing Criterion Phase being superior for shaping.

Session mean scores, rejection levels, and ranges of scores for Carol are shown in Figure 5. Means and standard deviations for these sessions are shown in Table 3. In accordance with the stability criteria, the difference between the mean of the first three of the last six baseline sessions and the mean of the last three was 3.3% of the mean of these six sessions. No trend is evident during Baseline. From a gross level of analysis, the mean of Carol's Baseline sessions was 153.3 compared to a mean of 146.1 during the final phase of the study. Thus, Carol showed overall improvement. During the first Training phase, there was an initial drop in mean session score but then mean session score showed an upward trend. When first changed to the Changing Criterion design, Carol's responding showed a

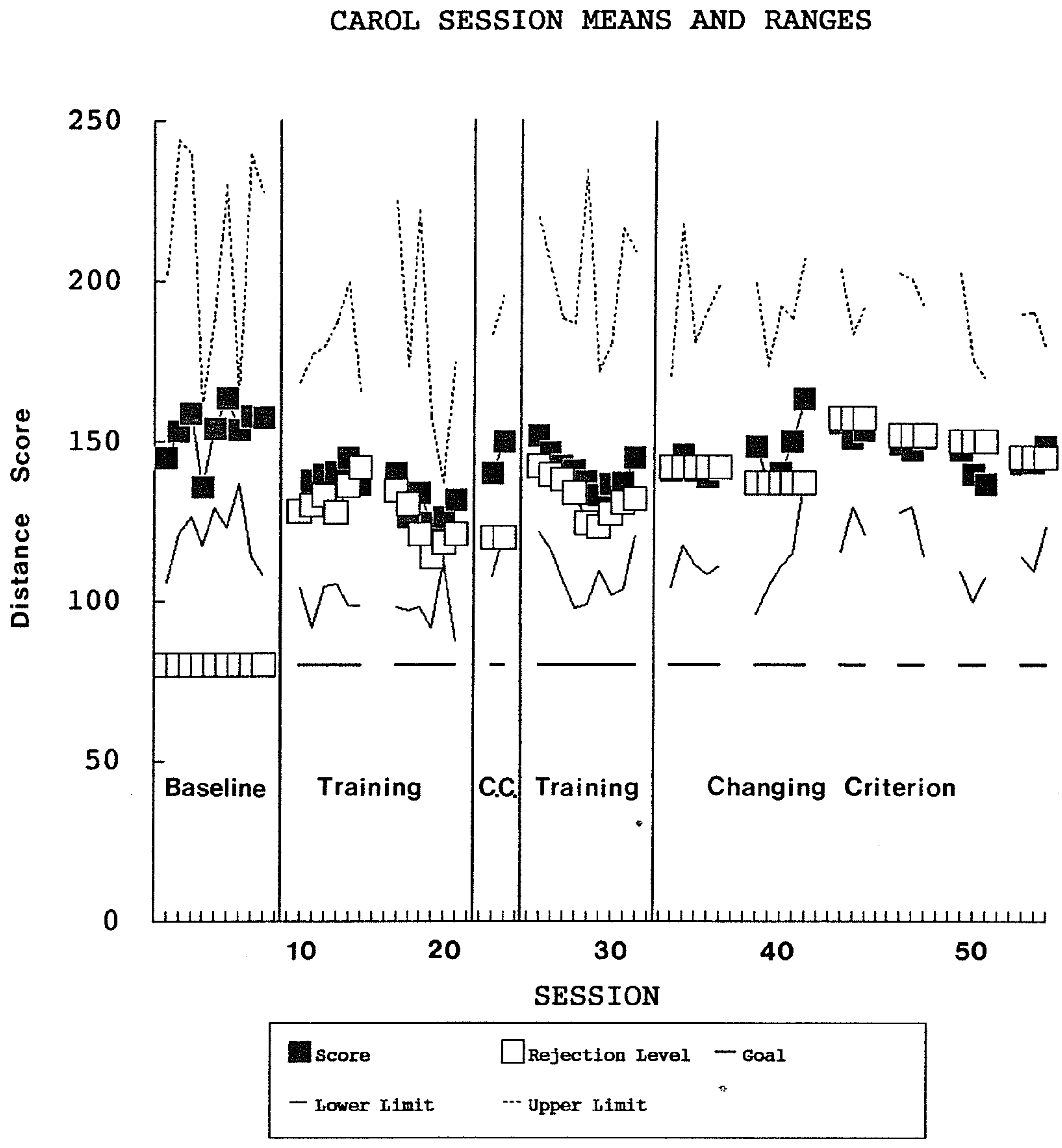

Figure 5. Session mean distance scores, rejection levels, and range of scores for Carol during Baseline (B.L.), Training I, Training II, and Changing Criterion (C.C.) phases.

Table 3

Session Means and Standard Deviations for Carol

Session	Mean	S.D.
	Baseline	
1	144.6	24.7
2	153.2	24.5
3	158.5	21.1
4	136.0	13.3
5	154.1	19.2
6	163.8	26.7
7	153.7	12.5
8	157.9	23.4
9	157.6	32.4
Mean	153.3	22.0
	Training I	
10	128.8	15.0
11	137.6	19.1
12	139.8	19.7
13	140.4	22.2
14	145.0	26.4
15	136.7	20.3
16	140.1	31.6
17	126.4	19.8
18	134.0	39.4
19	124.3	16.8
20	126.4	8.2
21	131.7	23.2
Mean	134.3	21.8
	Changing Criterion I	
22	140.2	19.4
23	150.0	18.8
Mean	145.1	19.1
	Training II	
24	151.9	22.2
25	146.6	21.1
26	142.2	19.1
27	141.0	19.0
28	137.5	26.7
29	133.4	15.1
30	136.5	16.7
31	137.2	19.4
32	145.1	17.3
Mean	141.2	19.6
	Changing Criterion II	
33	140.8	19.8
34	145.9	22.7
35	140.7	15.3
36	139.1	18.7
37	141.6	23.8
38	148.5	21.6
39	137.2	14.6
40	139.9	17.0
41	149.8	16.8
42	163.2	20.2
43	155.6	23.4
44	150.9	16.8
45	153.5	18.9
46	149.3	15.8
47	147.2	14.9
48	151.1	21.2
49	147.3	21.7
50	139.7	17.1
51	136.7	11.9
52	143.2	16.4
53	143.6	19.2
54	148.4	15.2
Mean	146.1	18.3

marked increase in distance from the criterion, perhaps caused by the criterion being set too stringently. With a change back to Training conditions, there was an initial downward trend but this was followed by an increasing trend. During the first subphase of the second Changing Criterion phase, Carol's mean scores closely matched the criterion. However, with the first tightening of the criterion, once again there was a rapid decline in accuracy. Relaxing the criterion then resulted in relatively close approximation to the criterion during the last four subphases (an exception being more accurate responding during the second from the last subphase). As for Amy, the final Changing Criterion Phase showed the most consistent range of scores with the high peaks of the previous phases eliminated.

Sample individual session data showing Carol's worst, average, and best sessions are shown in Figure 6. Baseline 3 (Session 3) shows that the majority of the scores are above 150 and several close to or above 200. Changing Criterion 20 (Session 20) has the majority of trials below 150 and none above 200. In this session, one can see that the majority of trials are relatively close to the criterion which was set for this session. Training 1 (Session 10), similar to Amy's training session shows the trial by trial adjustments in the rejection level. During this session, the range of scores was lower than that found in Carol's "average" session.

Figure 7 shows the percent of trials reinforced along with average distance score and rejection level, across sessions for Carol. This figure shows a pattern similar to Amy's with the mean number of trials reinforced during the

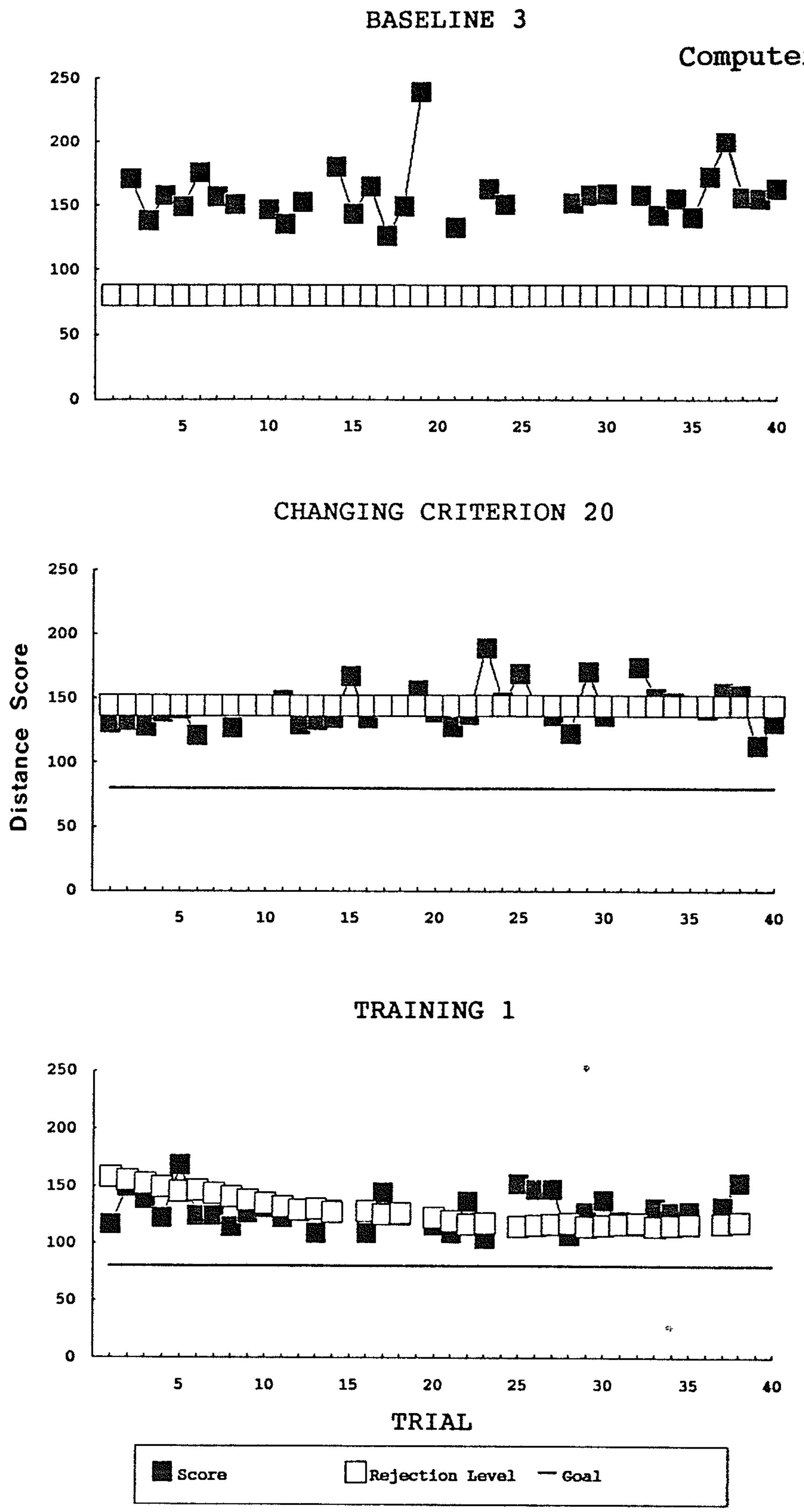

Figure 6. Sample sessions completed by Carol showing sessions with the highest, average, and lowest means (from top to bottom) and the training phase from which they were taken.

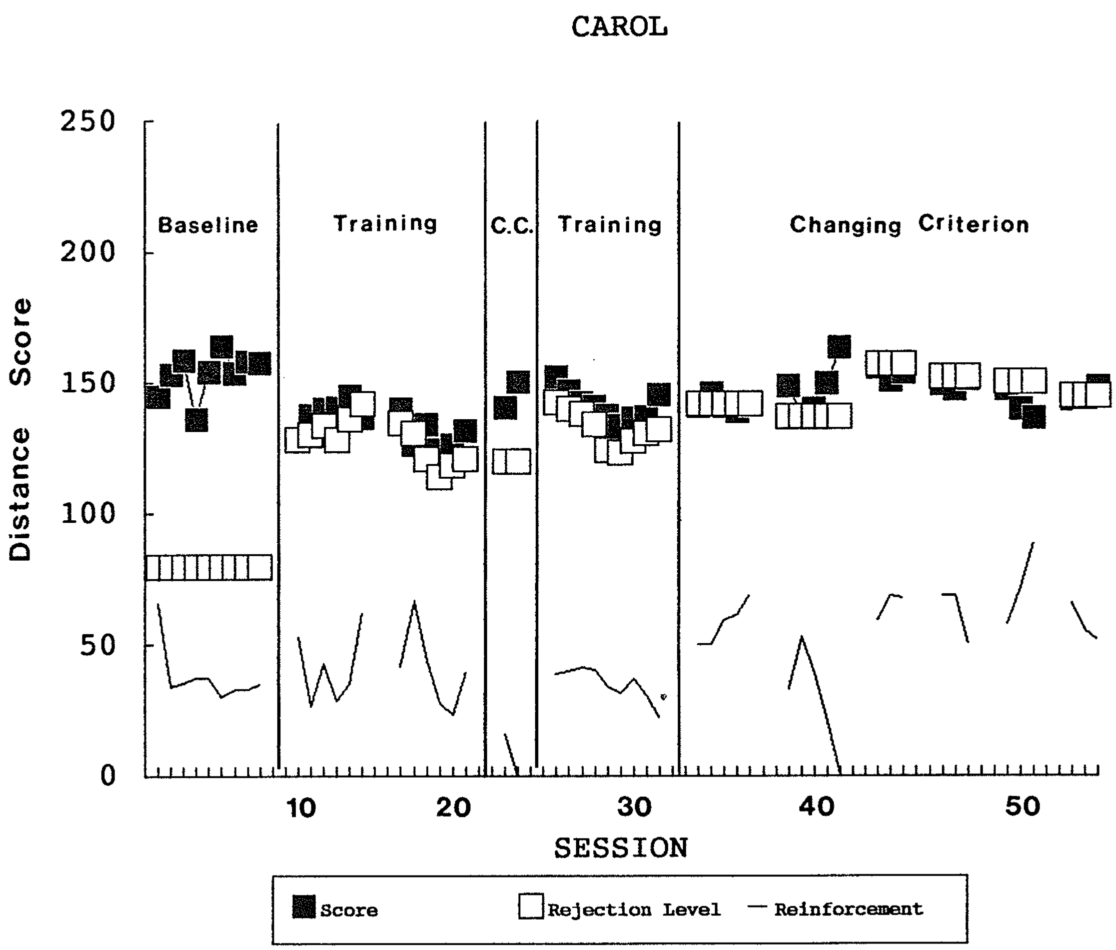

Figure 7. Percentage of trials reinforced compared with mean distance score and rejection level across sessions for Carol.

two training phases 41.0% and 35.1%, compared to 55.0% during the second Changing Criterion phase. During the first part of the second Training phase when Carol's distance scores were dropping, reinforcement was near constant at about 40%. This seems to indicate that for Carol, a reinforcement level between 40 and 55% is the best for maintaining a shaping procedure.

Figure 8 shows session mean scores, rejection levels, and ranges of scores for Brian. Means and standard deviations for these sessions are shown in Table 4. As previously mentioned, the stability criteria was not applied to Brian's Baseline due to the increasing trend in the

Table 4

Session Means and Standard Deviations for Brian

Session	Mean	S.D.	Session	Mean	S.D.
	Baseline			Changing Criterion	
1	123.7	24.8	14	150.3	29.7
2	127.6	20.3	15	144.0	24.4
3	132.9	19.8	16	146.7	36.2
4	137.3	24.7	17	152.8	20.4
5	143.0	28.7	18	139.5	22.3
Mean	132.9	23.7	19	148.1	26.8
			20	132.9	16.1
	Training		21	145.7	19.4
			22	146.2	21.1
6	156.5	29.3	23	143.2	23.2
7	140.0	23.7	24	137.2	20.0
8	151.7	24.8	25	146.1	20.5
9	141.5	25.3	26	133.3	21.0
10	146.0	32.0	27	143.6	24.7
11	144.7	24.7	28	150.0	23.5
12	160.5	20.9	29	150.9	21.0
13	157.1	27.4	30	141.7	22.3
Mean	149.7	26.0	Mean	144.2	23.1

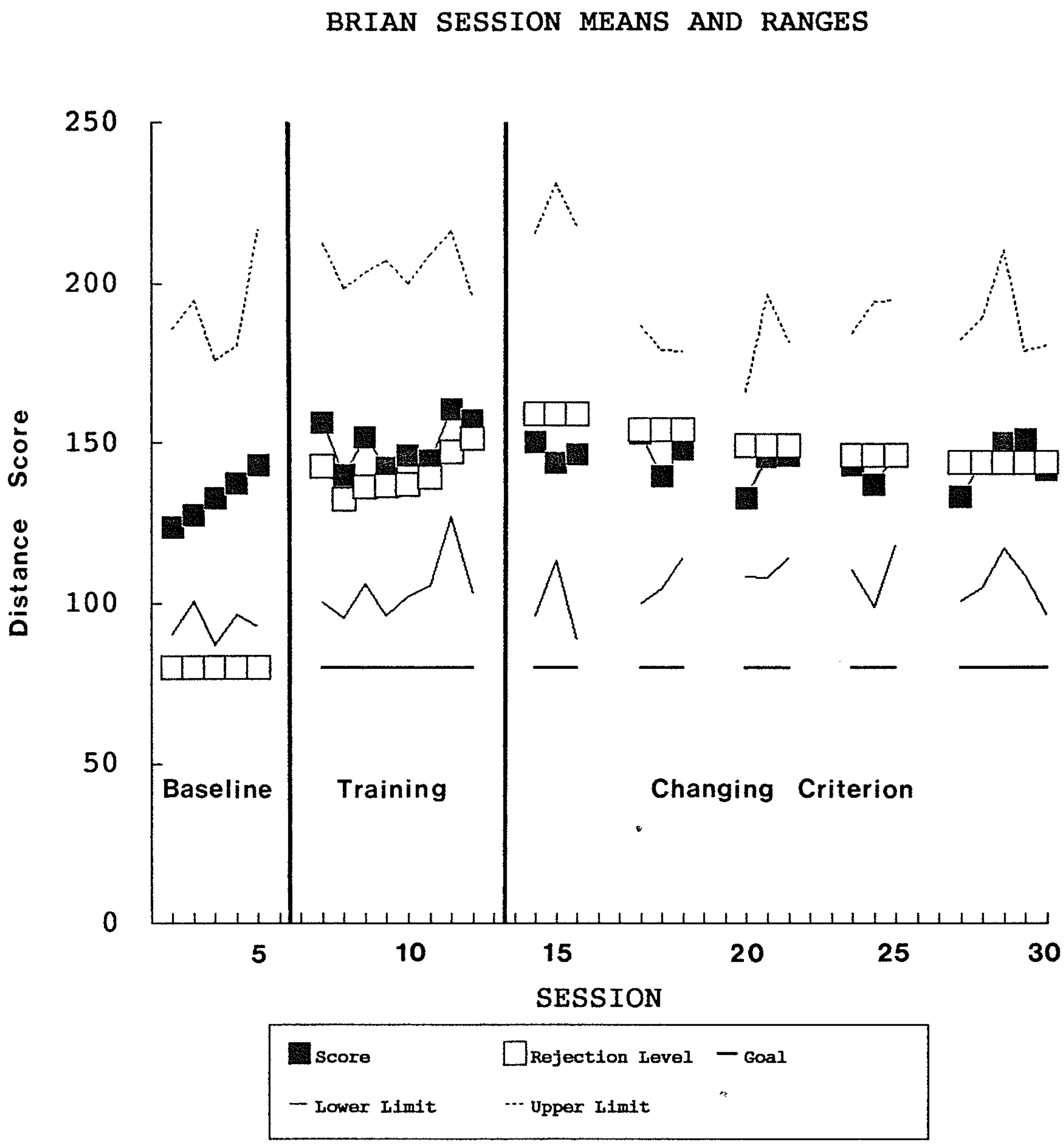

Figure 8. Session mean distance scores, rejection levels, and range of scores for Brian during Baseline (B.L.), Training I, Training II, and Changing Criterion (C.C.) phases.

scores. Although Brian never recovered the level of performance of his initial Baseline sessions, the Changing Criterion Phase was a distinct improvement over the training phase with 13 of 17 sessions during Changing Criterion below the mean of Training sessions. The average standard deviations for the phases (see Table 4) indicate that during the Training condition, Brian's within session variability increased from Baseline and recovered during the Changing Criterion phase. Responsiveness to the criterion can be seen in that in all but two sessions of this phase, Brian's mean score fell below the criterion.

Figure 9 shows Brian's worst, average, and best sessions. Training 7 (Session 12) shows an overall high range of scores and relatively few trials which received reinforcement. This can be contrasted to Brian's performance in Changing Criterion 14 (Session 27) where the rejection level is approximately the same and about half the trials are below the rejection level and none are above 200 as in Training 7. Baseline 1 shows a number of trials which came close to the goal and a majority of trials below 150.

Figure 10 shows the percent of trials reinforced, along with the mean distance score and rejection level, across sessions for Brian. Like Amy and Carol, Brian's average reinforcement rate during Training was about 30% (37.2%) compared to about 50% (61.4%) during the Changing Criterion

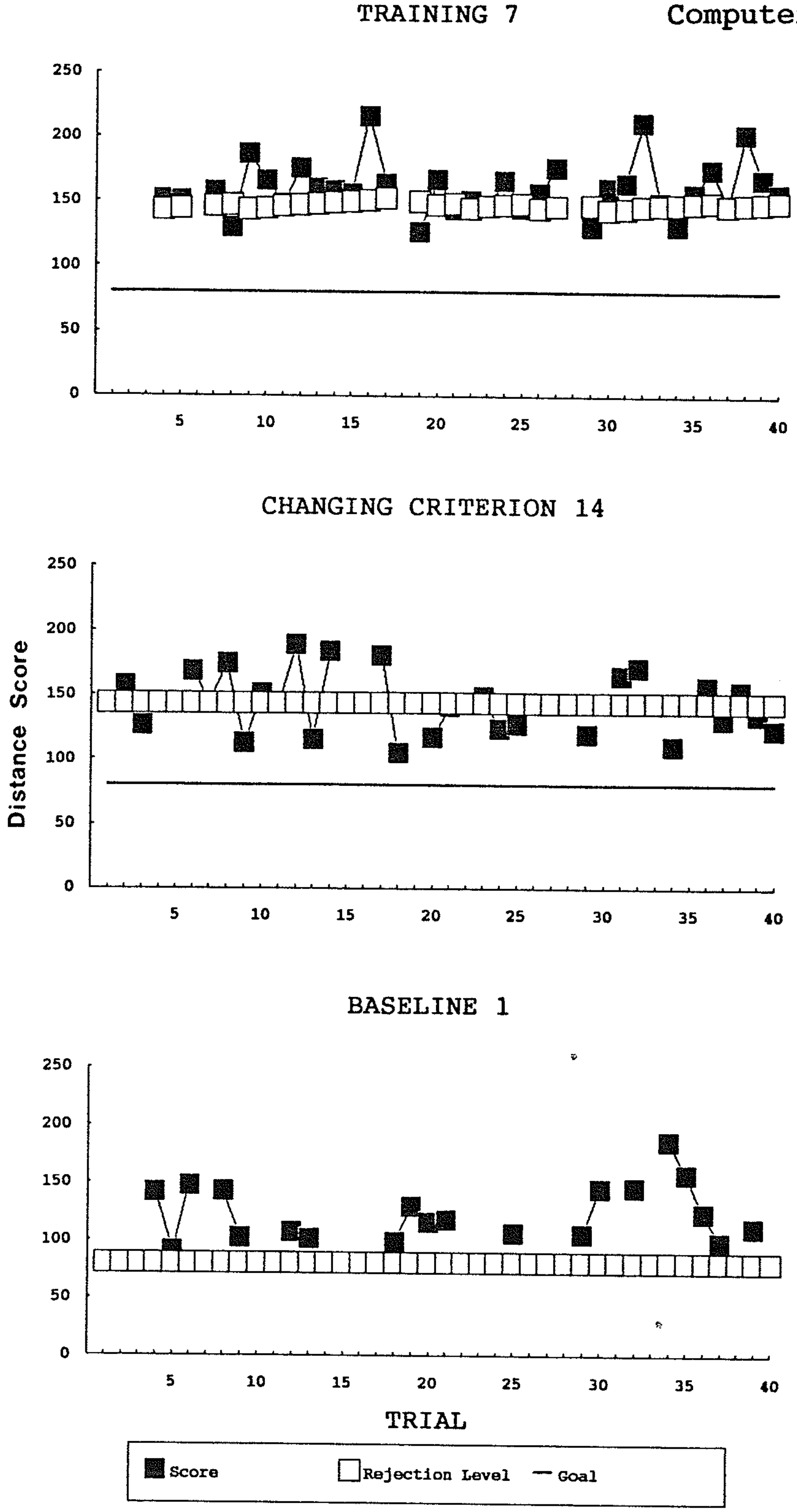

Figure 9. Sample sessions completed by Brian showing sessions with the highest, average, and lowest means (from top to bottom) and the training phase from which they were taken.

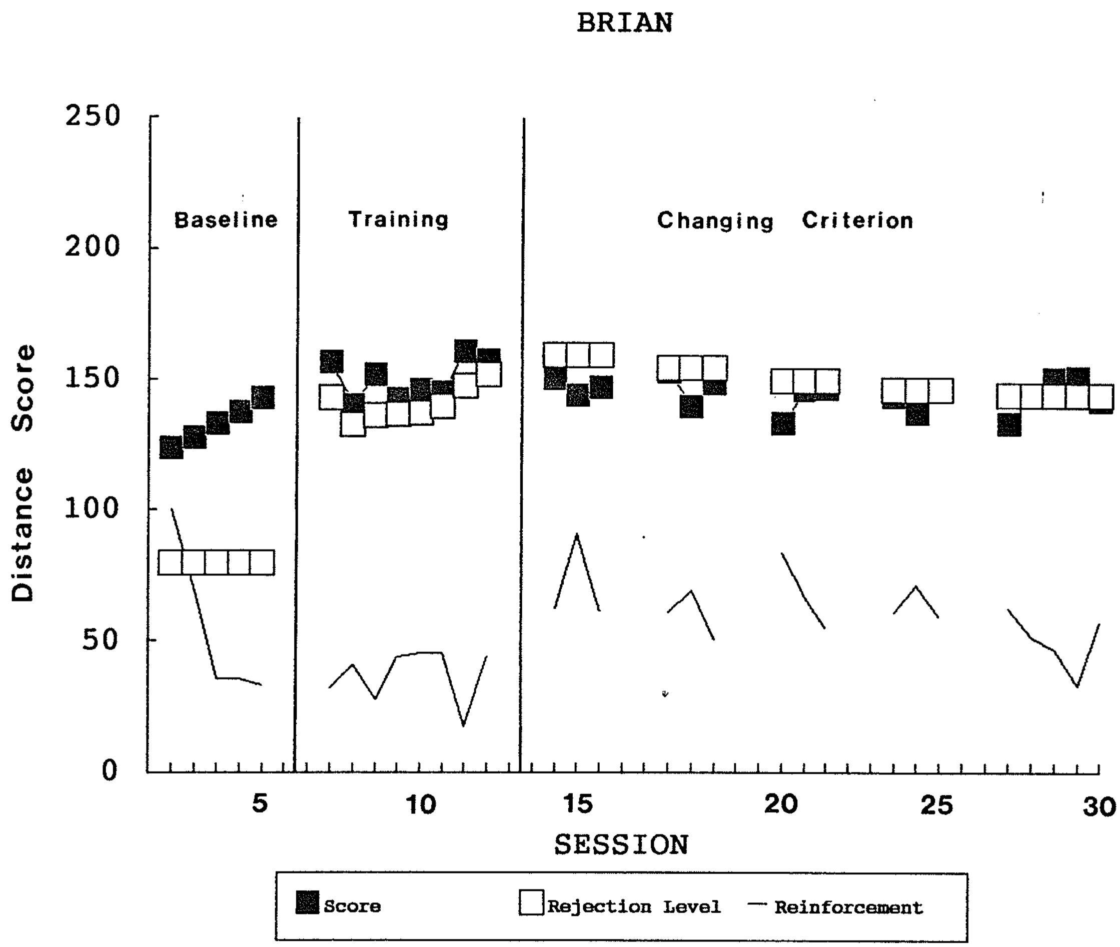

Figure 10. Percentage of trials reinforced compared with mean distance score and rejection level across sessions for Brian.

phase. With the close adherence to the criterion during the C.C. phase, one might conclude that this was a reasonable level of reinforcement for Brian to maintain responding that moves in the direction desired.

Control Probes

The results of the control probes are displayed in Figure 11. The sound which was trained for each of the children is shown as the solid blocks. This figure indicates that there was no generalization from training sessions to the probe sessions and no generalization to the control phoneme. These results could also indicate that the training had no effect. The only possible exception to this were Carol's responses during the Training phases when the sound /a/ was consistently lower than for /e/ indicating that during these phases Carol was better at imitating the sound that she was trained than on the control sound.

Voice Templates

Table 5 shows the percentage of trials that Amy's lowest distance score matched each of the different voice templates during her best (C.C. 18), average (Tr. 18), and worst (C.C. 9) sessions. Figure 12 is a scatter plot of these data with distance scores plotted against the templates. This shows that Amy's vocalizations were closest to one of the children's voices with the 13 year old female and the 6 year old male templates being matched the

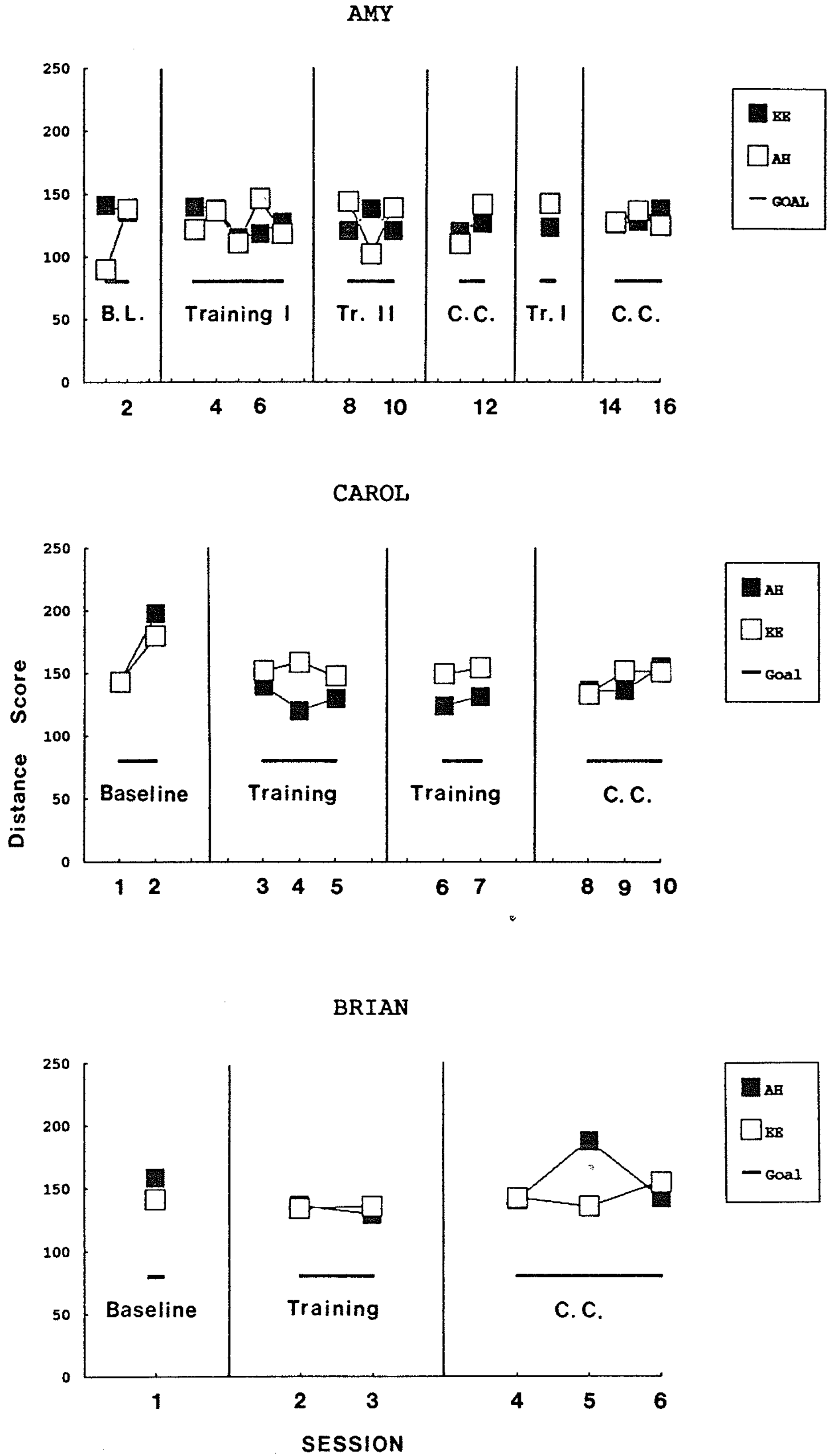

Figure 11. Mean distance scores for /a/ and /e/ during control probe sessions for Amy, Carol, and Brian during Baseline (B.L.), Training (Tr.), and Changing Criterion (C.C.) phases of the study. The phoneme which was trained is indicated by the filled squares.

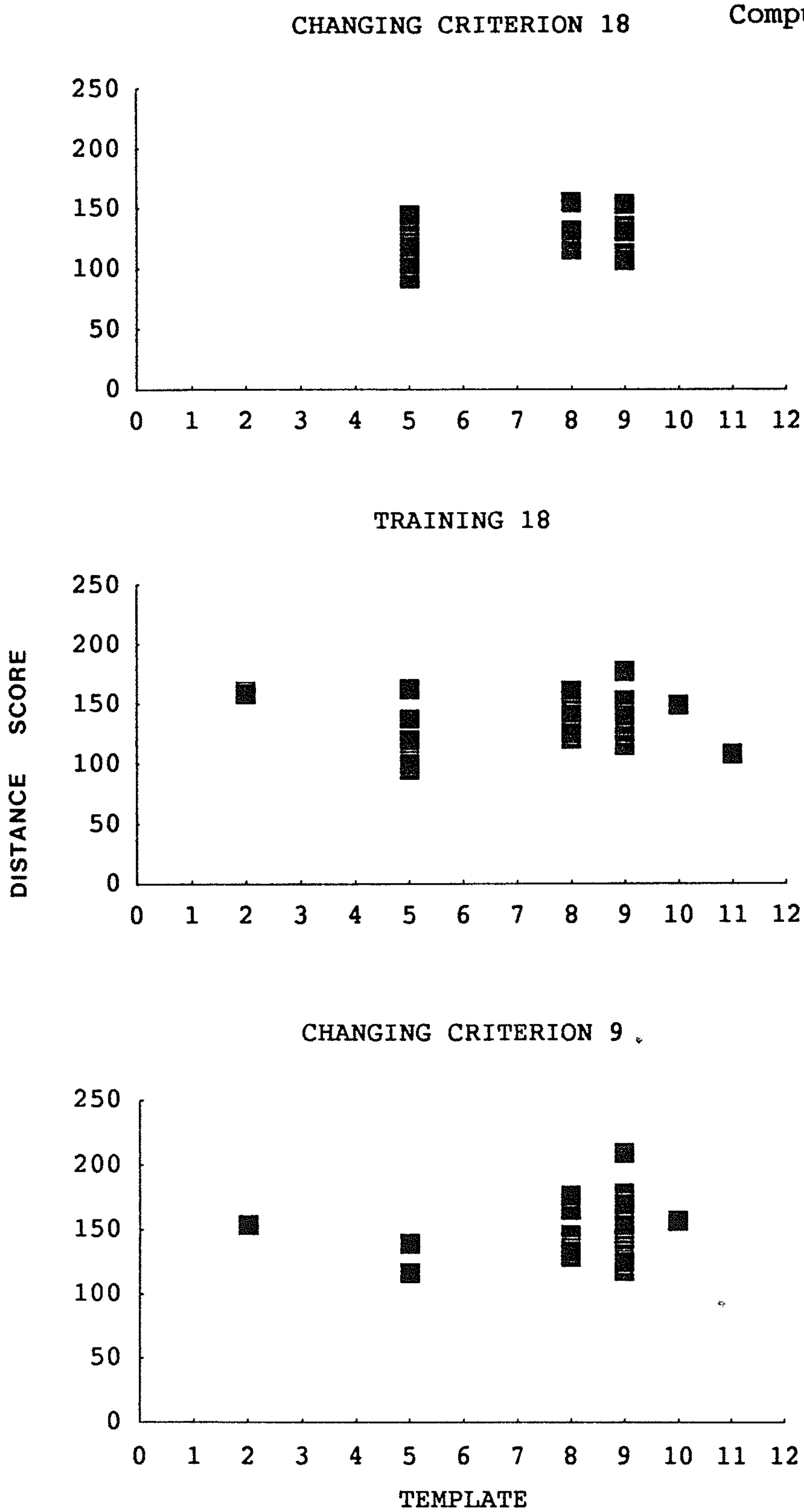

Figure 12. Scatter plots of distance score by voice template for Amy during her best, average, and worst sessions (top to bottom).

Table 5

Voice of Template and Percentage of Trials Each Template was the Lowest Score during Amy's Best, Average, and Worst Sessions

		Percent of Trials Lowest Score		
Template	Voice	Best	Average	Worst
1	Adult Female	-	-	-
2	Male 12	-	6.5	3.6
3	Male 12	-	-	-
4	Female 13	-	-	-
5	Female 13	37.5	45.2	7.1
6	Adult Male	-	-	-
7	Female 25	-	-	-
8	Male 6	18.8	19.4	28.6
9	Male 6	43.8	22.6	57.1
10	Male 3	-	3.2	3.6
11	Male 3	-	3.2	-
12	Adult Female	-	-	-

most often in the three sessions examined. As can be seen in Figure 12, there is no one template that was matched consistently for either high or low scores. However, for Amy it appears that with the passage of time and experience on the apparatus her vocal imitations became more concentrated in terms of the template matched (Training 18 was nearest to the beginning of the study and C.C. 18 was near the end).

Table 6 shows the percentage of trials that Carol's and Brian's lowest distance score matched each of the different voice templates during their best (Carol Training 1; Brian Baseline 1), average (Carol C.C. 20; Brian C.C. 14), and

Table 6

Voice of Template and Percentage of Trials Each Template was the Lowest Score during Carol's and Brian's Best, Average and Worst Sessions

		Percent of Trials Lowest Score					
		Carol			Brian		
Template	Voice	Best	Average	Worst	Best	Average	Worst
1	Male 3	2.9	-	-	-	-	-
2	Adult Female	5.9	-	-	5.0	3.7	2.9
3	Male 12	20.6	22.9	-	-	11.1	-
4	Male 12	5.9	-	3.2	-	3.7	8.8
5	Female 13	5.9	-	32.3	10.0	-	-
6	Female 12	-	-	-	-	-	2.9
7	Male 3	-	-	-	-	-	-
8	Adult Female	5.9	-	9.7	-	11.1	-
9	Male 24	-	-	-	-	-	2.9
10	Male 6	2.9	-	-	-	-	-
11	Adult Female	50.0	77.1	54.8	85.0	66.7	79.4
12	Male 6	-	-	-	-	3.7	2.9

worst (Carol Baseline 3; Brian Training 7) sessions. Figure 13 is a scatter plot of Carol's distance scores plotted against templates for each of these sessions. Unlike Amy, Carol's vocalizations were, during these sessions, most often closest to an adult female voice. As for Amy, there is no one template that was matched consistently for either high or low scores. Similar to Amy is the concentration of template matching towards the end of the study (C.C. 20) with her vocalizations most closely matching only two of the templates during this session. However, it appears that following baseline, there was an initial increase in experimentation with the vocalizations that Carol emitted in Training 1.

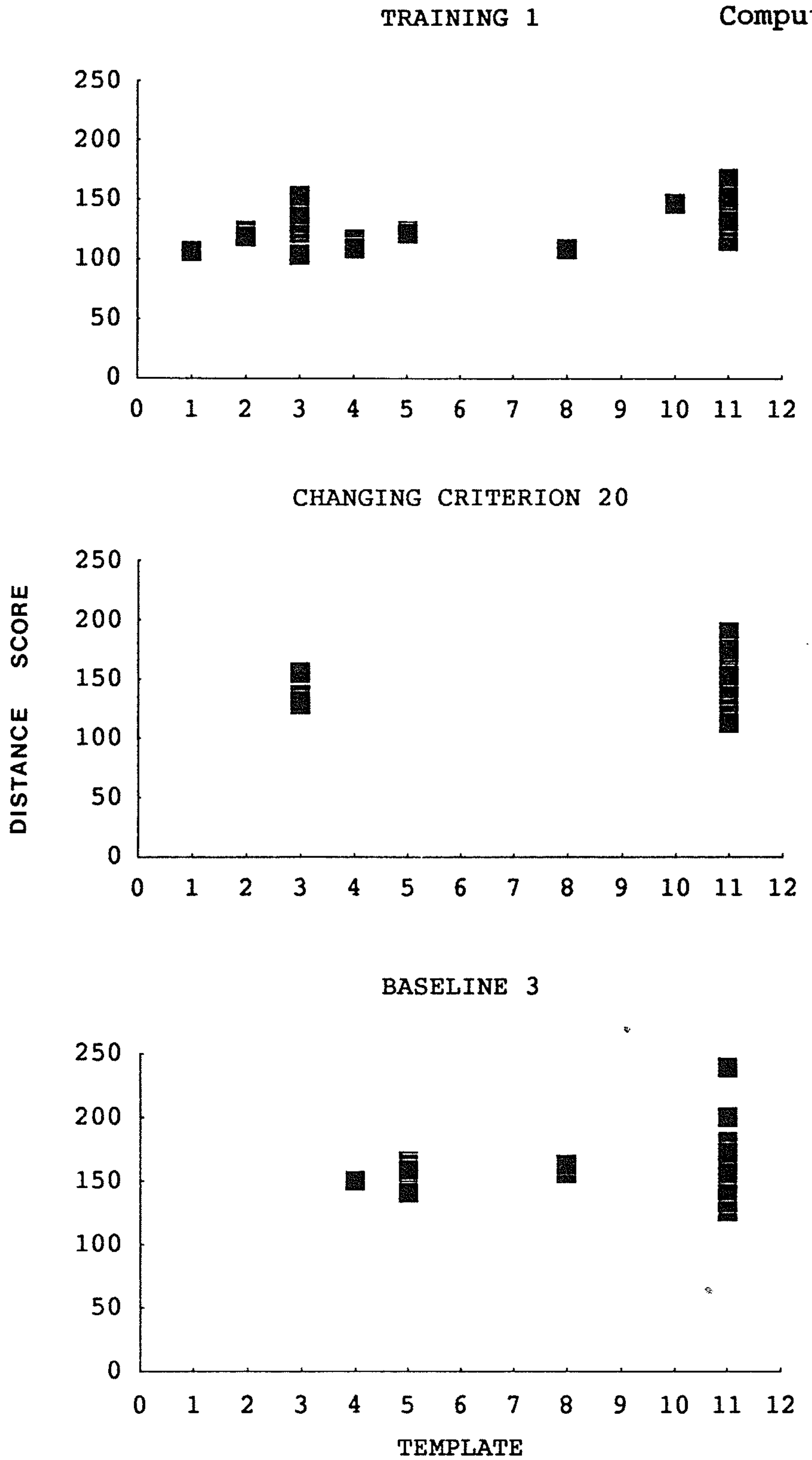

Figure 13. Scatter plots of distance score by voice template for Carol during her best, average, and worst sessions (top to bottom).

Brian's template matching data are shown in Table 6 and Figure 14. Like Carol, Brian's vocalizations during these three sessions, most often came closest to an adult female's voice. Like both Amy and Carol, there again is no one template that was consistently matched for either high or low scores. Unlike Amy and Carol, Brian's vocalizations did not become more concentrated in terms of the template matched in successive sessions. This may be related to the fact that Brian had the fewest number of sessions of all the subjects, only 30 compared to 54 for Carol and 75 for Amy. His pattern of results on this measure may be similar to the increase in experimentation in vocalizations seen in Carol's Training 1 session.

Social Validity

Correlations obtained between the computer and human raters and correlations between human raters were generally low. The correlations between computer distance score, Rater 1, Rater 1 Restricted Bandwidth (Rest. 1), Rater 2, and Rater 2 Restricted Bandwidth (Rest. 2) over all sessions rated are shown in Table 7. As can be seen, the highest correlations were obtained between a rater and her own ratings of the restricted bandwidth tapes. However, even these correlations were at the low end of those previously found between human raters. This indicates that restricting the bandwidth did have an impact on sound rating.

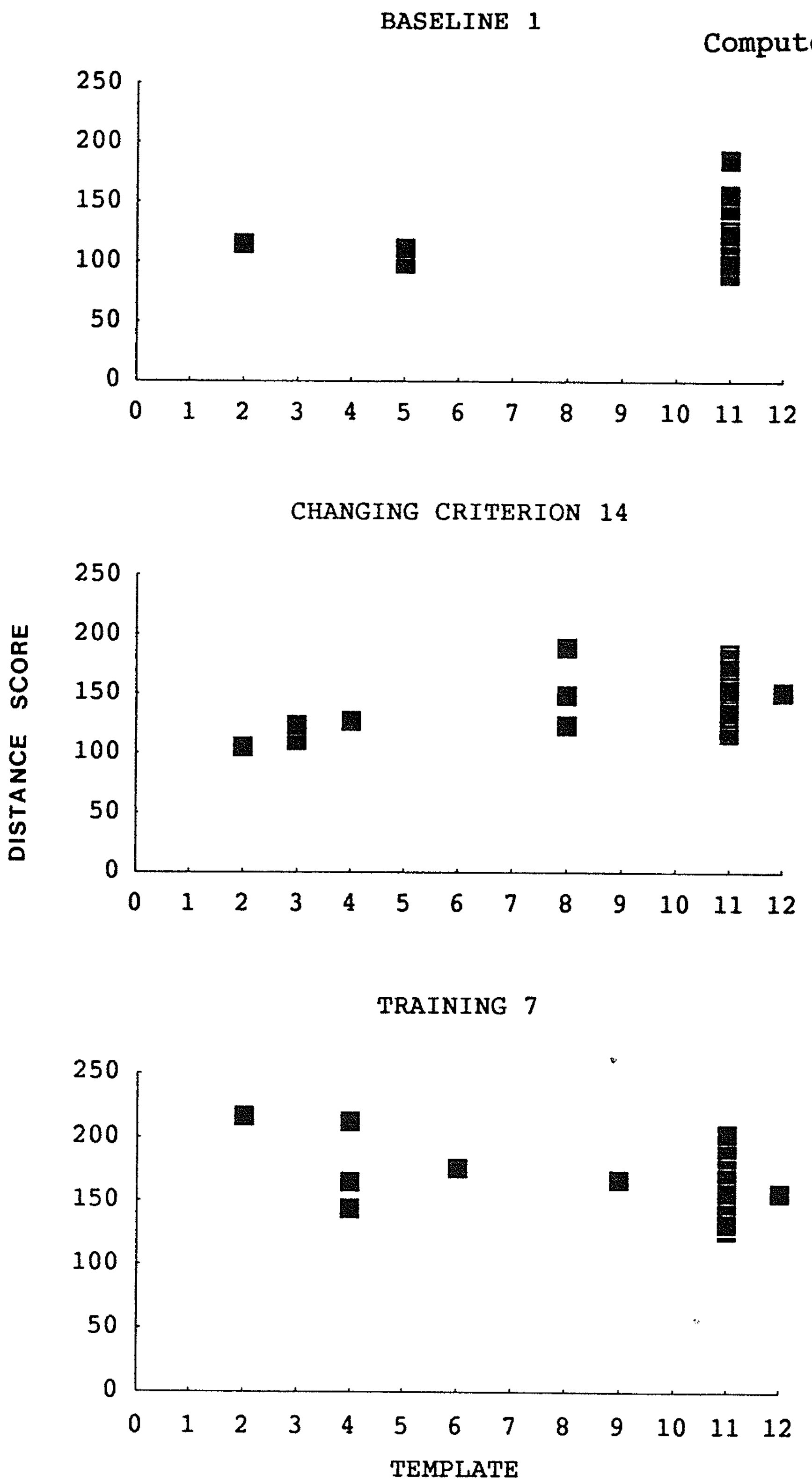

Figure 14. Scatter plots of distance score by voice template for Brian during his best, average, and worst sessions (top to bottom).

Table 7

Correlations Between Computer Distance Scores and Human Raters (Unrestricted and Restricted Band Widths) Over all Sessions Rated

	Distance Score	Rater 1	Rater 1 Restricted	Rater 2
Rater 1	-.05	-	-	-
Rater 1 Restricted	.03	.67**	-	-
Rater 2	.14*	.36**	.35**	-
Rater 2 Restricted	.08	.36**	.35**	.80**

* $p < .05$
** $p < .01$

A different pattern of correlations emerges if one examines the correlations for each child emitting the sounds. These data are presented in Table 8. In the overall analysis, Rater 2's judgements were more closely related to the computer distance scores. When correlations were calculated for each child, the only significant correlation between human raters and the computer was for Rater 1 under restricted bandwidth conditions for Carol. For Amy, all correlations between the computer and human ratings were in a negative direction. Rater 1's correlations with her own ratings under restricted bandwidth conditions were lower for Amy and Carol's vocalizations than that previously found between human raters (Pear, Kinsner, &

Roy, 1987).

Table 8

Correlations Between Computer Distance Scores and Human Raters (Unrestricted and Restricted Band Widths) Over all Sessions Rated for Each Child

Amy

	Distance Score	Rater 1	Rater 1 Restricted	Rater 2
Rater 1	-.12	-	-	-
Rater 1 Restricted	-.11	.44**	-	-
Rater 2	-.02	.57**	.45**	-
Rater 2 Restricted	-.16	.51**	.48**	.81**

Brian

	Distance Score	Rater 1	Rater 1 Restricted	Rater 2
Rater 1	-.12	-	-	-
Rater 1 Restricted	-.06	.84**	-	-
Rater 2	.15	.70**	.75**	-
Rater 2 Restricted	.05	.60**	.67**	.86**

Carol

	Distance Score	Rater 1	Rater 1 Restricted	Rater 2
Rater 1	.03	-	-	-
Rater 1 Restricted	.31**	.48**	-	-
Rater 2	.15	.27**	.41**	-
Rater 2 Restricted	.15	.31**	.37**	.63**

* $p < .05$
** $p < .01$

Figures 15, 16, and 17 show scatter plots of each of the human raters under restricted and unrestricted bandwidth conditions plotted against the computer distance score for each child. In Figure 15, for Amy who was trained with the phoneme /e/, restricting the bandwidth resulted in Rater 1 judging the quality of responses as poorer, whereas it had little effect on Rater 2. Figure 16, showing this data for Brian (phoneme /a/) indicates that restricting the bandwidth affected Rater 2's judgements in a negative direction more than it affected Rater 1's judgements. In Figure 17, for Carol (phoneme /a/), restricting the bandwidth resulted in Rater 2 judging vocalizations more positively and both raters corresponding more closely with the computer ratings.

Figure 18 shows the mean session rating for each child's best, average, and worst sessions (according to computer distance scores) for both raters under unrestricted (Rater 1 and Rater 2) and restricted (Rest. 1 and Rest. 2) bandwidth conditions. Rater 1, under both bandwidth conditions, rated Amy's and Brian's vocalizations as further from the target sound than did Rater 2. This was reversed for Carol where Rater 2 judged the vocalizations more poorly than did Rater 1. For Amy, both raters, under both bandwidth conditions, indicated a pattern of best, average, and worst sessions, completely different from that indicated by the computer ratings. Brian's "best" session was

AMY

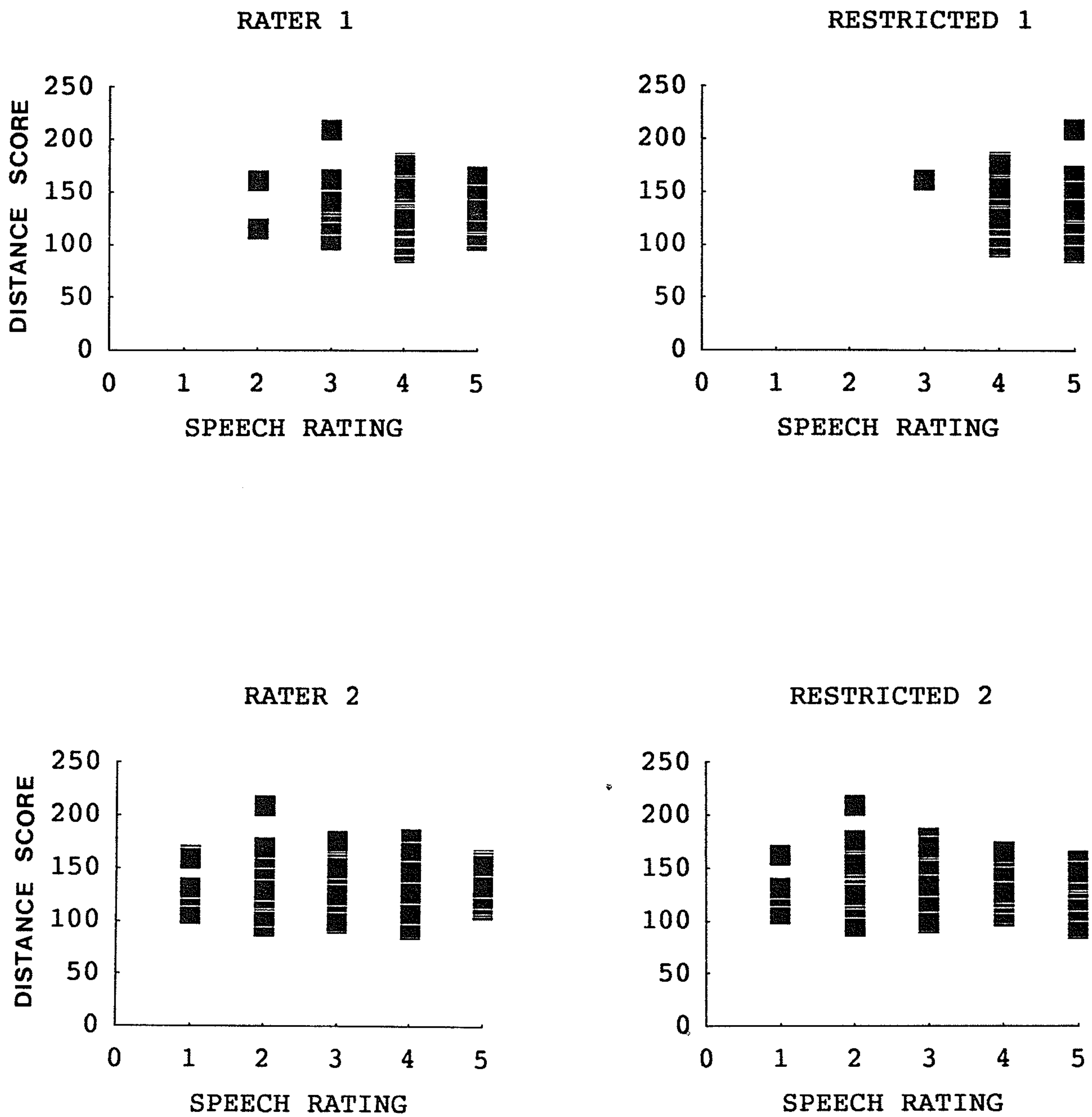

Figure 15. Scatter plots of each of the human raters under unrestricted and restricted band width conditions plotted against the computer distance scores for Amy (phoneme /e/).

BRIAN

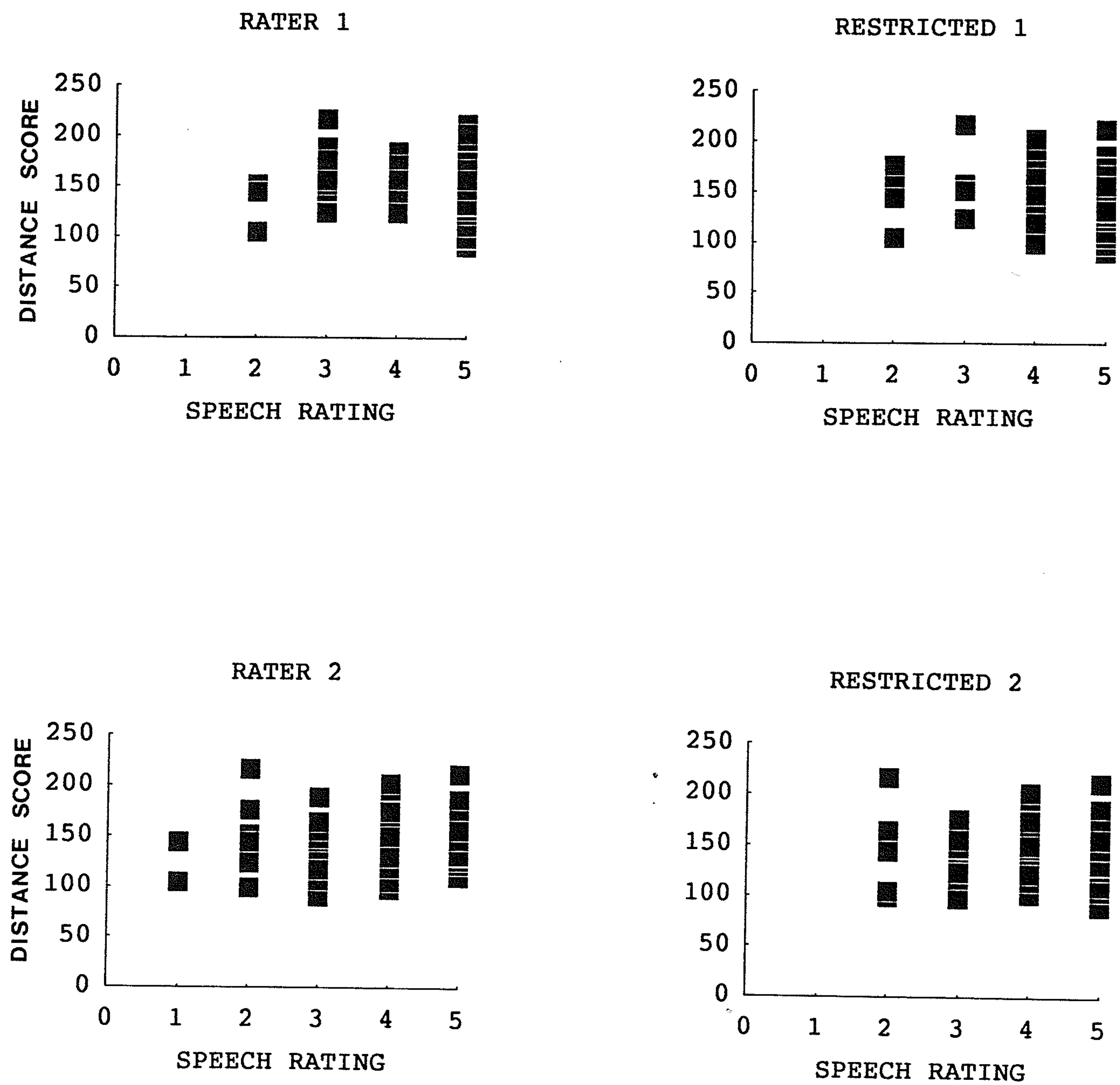

Figure 16. Scatter plots of each of the human raters under unrestricted and restricted band width conditions plotted against the computer distance scores for Brian (phoneme /a/).

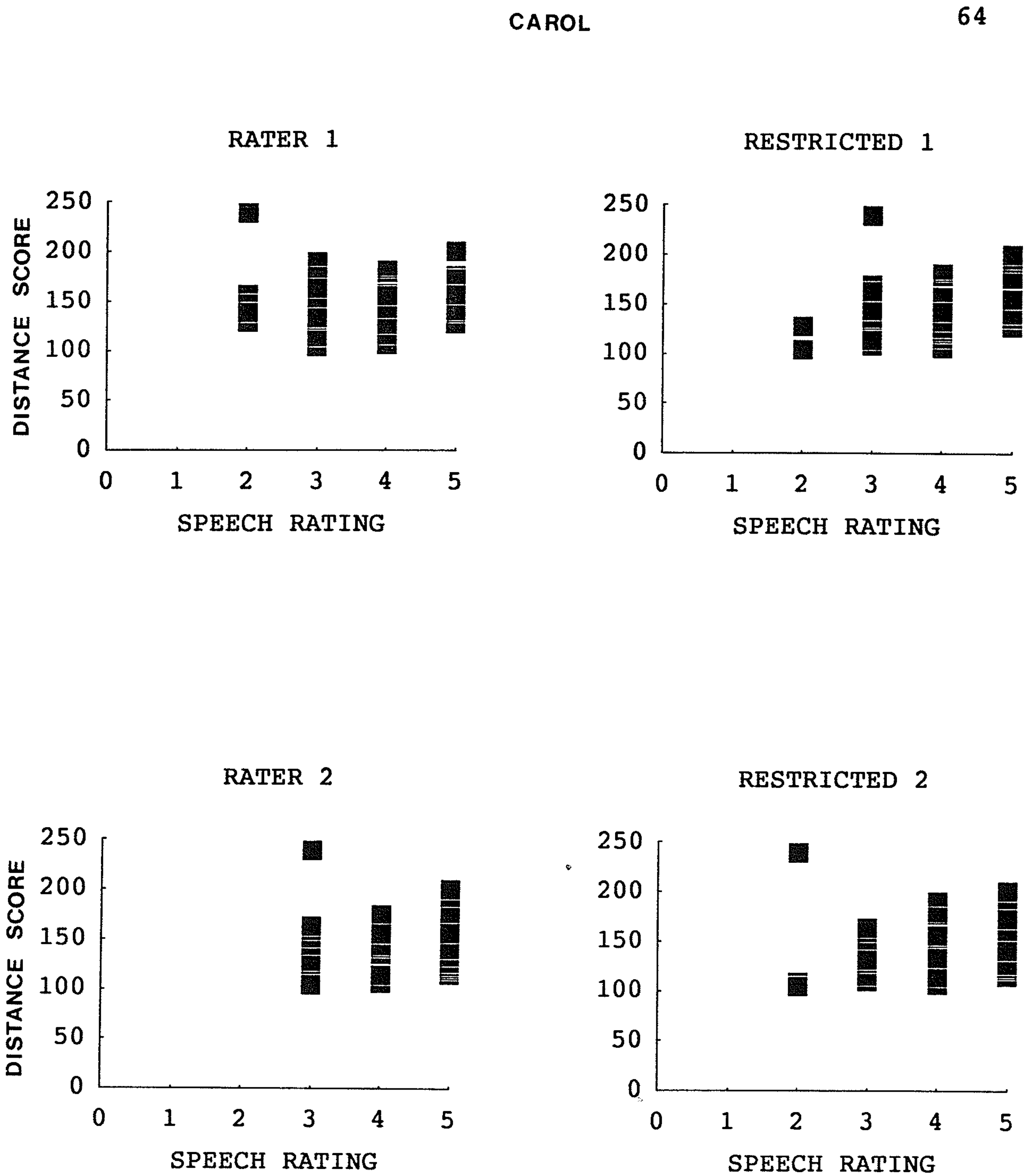

Figure 17. Scatter plots of each of the human raters under unrestricted and restricted band width conditions plotted against the computer distance scores for Carol (phoneme /a/).

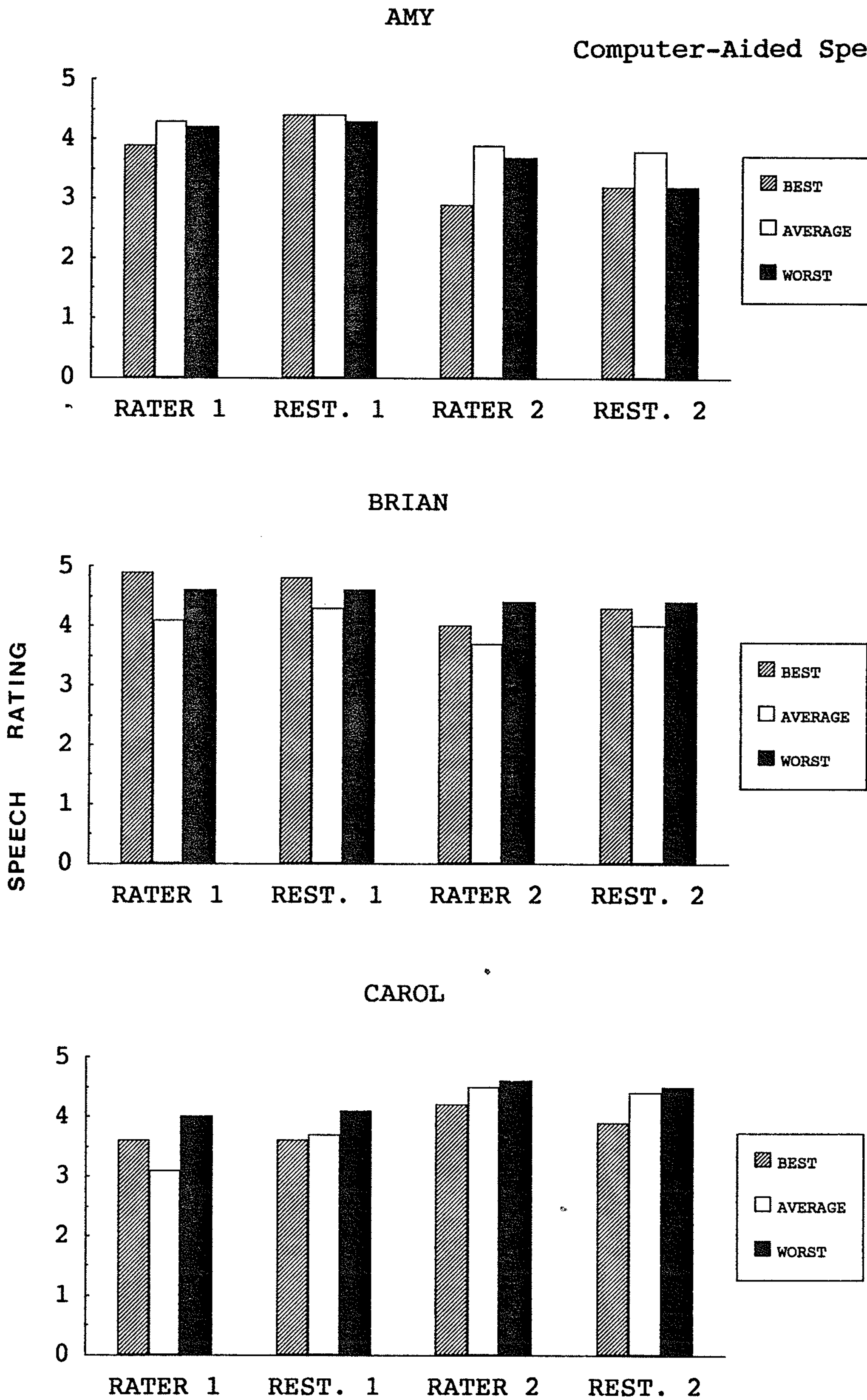

Figure 18. Mean session rating of each child's best, average, and worst sessions for both raters under unrestricted (Rater 1 & Rater 2) and restricted (Rest. 1 & Rest. 2) band width conditions.

consistently judged to be poor by both raters. However, his average and worst sessions show the same pattern as the computer ratings. With the exception of Rater 1's unrestricted bandwidth judgements of Carol's best and average sessions, there is agreement between the human raters' and the computer's assessments of best, average, and worst sessions.

Figure 19 shows the mean session rating (of the sessions rated by the speech therapists) nearest to the beginning (closed bars) and nearest to the end (open bars) of the study for each child for both human raters under unrestricted (Rater 1 and Rater 2) and restricted (Rest. 1 and Rest. 2) bandwidth conditions. With the exception of Rater 1 under restricted bandwidth conditions for Amy (which indicated no change), this figure shows that the session nearest to the end of the study was consistently rated as better than the session nearer the beginning of the study. Rater 1 indicated only slight improvement for Carol under both bandwidth conditions. Rater 2 indicated the most improvement for Amy, while Rater 1 indicated the most improvement for Brian and Carol. Thus, over the course of the study, both raters agreed that all three children improved at least somewhat in their ability to imitate the phoneme trained.

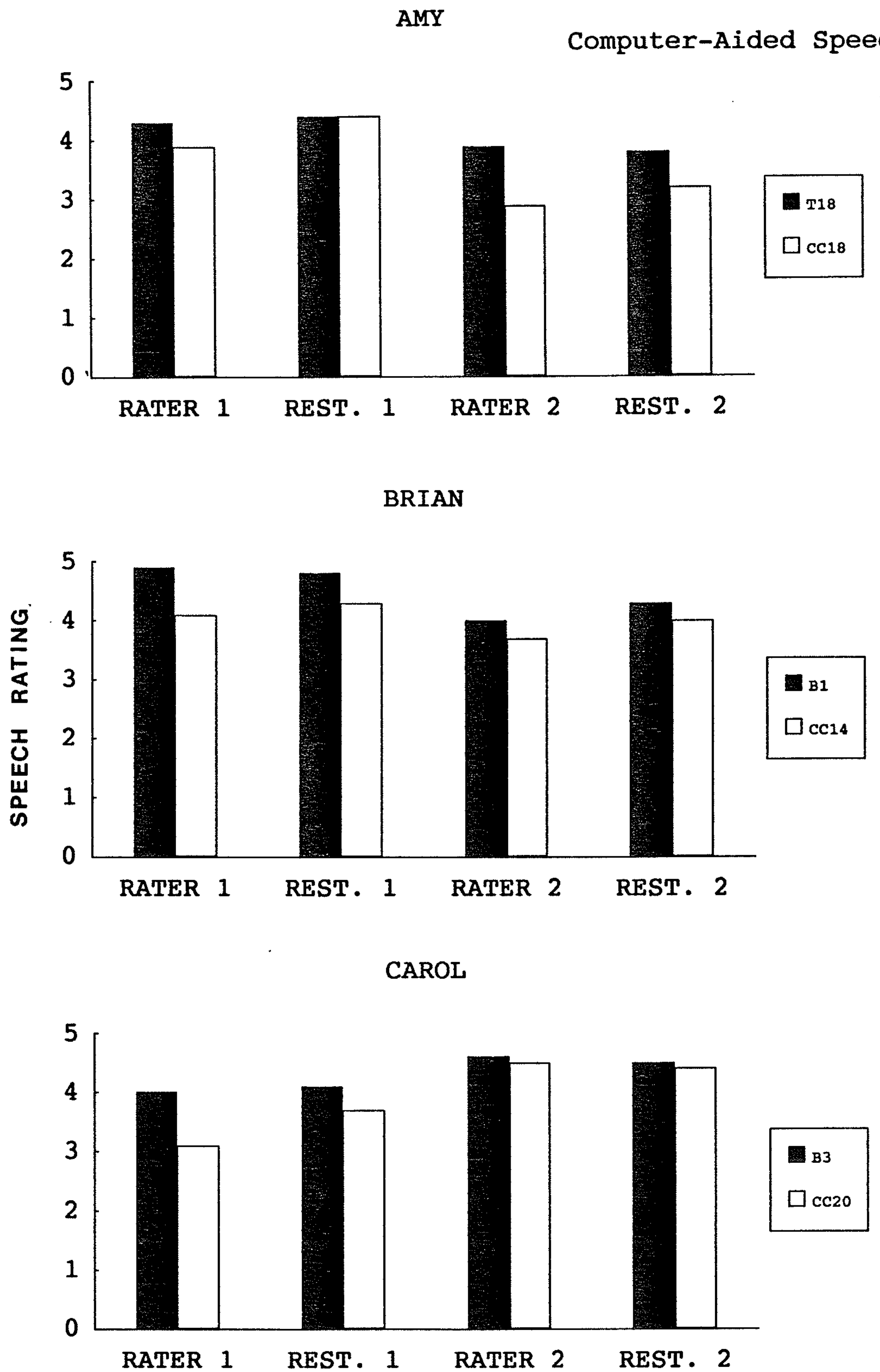

Figure 19. Mean session rating nearest the beginning (closed bars) and nearest the end (open bars) of the study for each child by both raters under unrestricted (Rater 1 and Rater 2) and restricted (Rest. 1 and Rest. 2) band width conditions.

Discussion

The abbreviated reinforcement assessment conducted at the beginning of this research was effective in identifying reinforcers which were effective enough to maintain responding over numerous sessions and a long period of time. Given the importance of using effective reinforcers, the time required to conduct this assessment can be considered reasonable.

The current results demonstrate the potential of a more precise methodology for the study of shaping and applied speech training. Due to the exploratory nature of the study and the fact that the project was terminated shortly after the Changing Criterion design started to demonstrate control over behaviour with a shaping trend, a more conclusive statement in this regard cannot be made at this time. Although both raters agreed that improvement in vocal imitation occurred for all three children, the current data is unable to differentiate if this improvement was related to shaping, practice effects, or maturation of the children. To obtain a clearer picture of this, human ratings of probe sessions and more intermediate sessions would be required. Although the achievement of shaping is only suggestive, the present research does demonstrate that trial by trial adjustments in the shaping parameters (Training conditions) are not effective. It also demonstrates that control over vocal behaviour can be achieved when the criterion is held constant over longer periods of time and when reinforcement is maintained at a high level. Under these latter conditions variability is less than under trial by trial adjustment conditions.

Unlike the laboratory animal research on shaping which indicated large rapid shaping steps maximized acquisition of the desired response, the present results indicate that vocal imitation in a developmentally handicapped population must proceed at a slower rate with small, slow step sizes, and that backtracking of the steps must occur rapidly to avoid serious loss of behaviour when performance declines. This finding is more in keeping with the results of Lane et al. (1967) who found that shaping the length of saying the phoneme /u/ fails if the probability of reinforcement is too low. It may be that the ideal speed of the shaping procedure will depend in large part on the normal rate of acquisition of a given behaviour in a given population. Vocal imitation acquisition is known to be a long and tedious process which, therefore, logically points to the use of small, slow shaping steps.

The consistency shown among all three children in their response to the different reinforcement rates during Training phases and Changing Criterion phases would seem to indicate that a reinforcement ratio of approximately 1:3 results in variable behaviour but a consistent shaping trend is obtained when the reinforcement rate is between 50 and 60%. Future shaping work in this area should therefore plan the shaping parameters to maintain reinforcement in this range. Although reinforcement rate was not specifically controlled, it is remarkable that the various training conditions used in this research resulted in very similar rates of reinforcement for all three children.

Correlations between the computer distance scores and human ratings were lower than those found by Pear, Kinsner, and Roy (1987). This may be related to the fact that one rater indicated the method by which she interpreted the rating scale and the method of interpreting the scale by the other rater is unknown. Also, the raters were not given any specific instructions to rate only the first portion of a vocalization. Both Brian and Amy, who had the lowest correlations, tended to babble with mixed use of vowels and consonants, whereas Carol more frequently emitted a singular sound. This may also account for the low correspondence between the computer and human assessments of Brian's "best" session (Baseline 1) since Brian's most frequent utterance early in the study resembled "eh da da". The computer would rate the "eh" relatively close to /a/ whereas the inclusion of the "da da" would spoil the response for the human raters.

In terms of social validity, agreement between the computer and human raters is necessary only at the two end points of the shaping process. That is, the computer and humans should agree that the subject is unable to correctly imitate the desired sound at the beginning of training, and that the subject improves in this ability at the end of training. The new method is not designed (nor should it be) to duplicate exactly the procedures that would be conducted by a human shaper.

The vocal rating by humans indicate a great deal of variability in the assessment of different sounds in different children, both between the computer and human raters, between human raters, and within raters according to whether bandwidth was restricted or not. The effect of restricting the bandwidth depended on both the individual rater and the child's voice being rated. In some cases it resulted in more favourable evaluations of the vocalization, and in other cases, less favourable. Unlike the finding that humans rated vocalizations in a dichotomous manner (Desrochers, 1989), these results indicated that these raters used the majority of the scale available to them. The variability of human raters emphasizes the advantage of the computer's precision and repeatable results.

The analysis of the template data reinforces the value of using a cluster of voice samples in that different subjects in this study more closely approximated different voice templates. Had only one template been used, the chances are high that a subject would be punished for approximating a target sound if his/her vocal characteristics differed from those of the target voice.

Ideally, future work with this type of apparatus will be more flexible in the forms of stimuli that it can present and analyze. That is, rather than presenting a single phoneme, it would be able to present whole and partial words since in learning speech children do not aim at single segments of speech pronounced in isolation. They aim at words and phrases, that is whole sequences of phonetic segments (Clark & Clark, 1977). The early words produced are those that have some meaning in the child's life. Therefore, more effective speech shaping might occur if the apparatus could train the name of the reinforcer selected by the child for any given session.

References

Alleman, H.D. & Platt, J.R. (1973). Differential reinforcement of interresponse times with controlled probability of reinforcement per response. Learning and Motivation, 4, 40-73.

Bell, T.S., Dirks, D.D., & Carterette, E.C. (1988). Interactive factors in consonant confusion patterns. Journal of the Acoustical Society, 85, 339-346.

Cairns, S. (1989). Speech shaping cluster program: Users guide. Unpublished manuscript, University of Manitoba, Winnipeg.

Carr, E.G.; Binkoff, J.A.; Kologinsky, E.; & Eddy, M. (1978). Acquisition of sign language by autistic children. 1: Expressive labelling. Journal of Applied Behaviour Analysis, 11, 489-501.

Clark, H.H. & Clark, E.V. (1977). Psychology and language: An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich.

Davis, E.R. & Platt, J.R. (1983). Contiguity and contingency in the acquisition and maintenance of an operant. Learning and Motivation, 14, 487-512.

Desrochers, M.N. (1989). Computer-based versus human assessment of vocal responses with developmentally handicapped individuals. Doctoral dissertation, University of Manitoba, Winnipeg.

Eckerman, D.A., Hienz, R.D., Stern, S., Kowlowitz, V. (1980). Shaping the location of a pigeon's peck: Effect of rate and size of shaping steps. Journal of the Experimental Analysis of Behaviour, 33, 299-310.

Eisenberger, R. (1989). Can response force be shaped by reinforcement? Perceptual and Motor Skills, 68, 725-726.

Gaines, R.; Leaper, C.; Monahan, C; & Weickgenant, A. (1988). Language learning and retention in young language-disordered children. Journal of Autism and Developmental Disorders, 18, 281-296.

Garcia, E.E. & DeHaven, E.D. (1974). Use of operant techniques in the establishment and generalization of language: A review and analysis. American Journal of Mental Deficiency, 79, 169-178.

Green, C.W.; Reid, D.H.; White, L.K.; Halford, R.C.; Brittain, D.P.; & Gardner, S.M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. Journal of Applied Behaviour Analysis, 21, 31-43.

Hingtgen, J.N. & Trost, F.C. (1966). Shaping cooperative responses in early childhood schizophrenics: Reinforcement of mutual physical contact and vocal responses. In Ulrich, R., Stachnik, T., & Mabry, J. (Eds.) Control of human behaviour. Glenview, Il.:

Scott, Foresman & Co.

Isaacs, W., Thomas, J. & Goldiamond, I. (1966). Application of operant conditioning to reinstate verbal behaviour in psychotics. In Ulrich, R., Stachnik, T., & Mabry, J. (Eds.) Control of human behaviour. Glenview, Il.: Scott, Foresman & Co.

Kazdin, A.E. (1982). Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press.

Kearsley, G. & Seidel, R.J. (1985). Automation in training and education. Human Factors, 27, 61-74.

Kinsner, W., Pear, J.J., & Roy, D.K. (1986). A speech recognition system and resynthesis subsystem. Technical Reports. Winnipeg: University of Manitoba, Microelectronics Centre, Dept. Electrical Engineering and Dept. of Psychology.

Kirkpatrick, M.A. & Fowler, S.C. (1989). Operant force-band differentiation by rats using two different response topographies. Bulletin of the Psychonomic Society, 27, 52-54.

Lane, H., Kopp, J., Sheppard, W., Anderson, T. & Carlson, D. (1967). Acquisition, maintenance, and retention in the differential reinforcement of vocal duration. Journal of Experimental Psychology Monograph Supplement, 74, (2), 1-16.

Martin, G. & Pear, J. (1988). Behaviour modification: What it is and how to do it (3rd Ed.). New Jersey: Prentice-Hall. (Ch. 5, pp. 67-82).

Mason, S.A., McGee, G.G., Farmer-Dougan, V., & Risley, T.R. (1989). A practical strategy for ongoing reinforcer assessment. Journal of Applied Behaviour Analysis, 22, 171-179.

Midgley, M.; Lea, S.E.G.; & Kirby, R.M. (1989). Algorithmic shaping and misbehavior in the acquisition of token deposit by rats. Journal of the Experimental Analysis of Behaviour, 52, 27-40.

Pace, G.M.; Ivancic, M.T.; Edwards, G.L.; Iwata, B.A.; & Page, T.J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. Journal of Applied Behaviour Analysis, 18, 249-255.

Panyan, M.C. & Hall, V. (1978). Effects of serial versus concurrent task sequencing on acquisition, maintenance, and generalization. Journal of Applied Behaviour Analysis, 11, 67-74.

Pear, J.; Kinsner, W.; and Roy, D. (1987). Vocal shaping of retarded and autistic individuals using speech synthesis and recognition. Proceedings of the Ninth Annual Conference of the IEEE Engineering in Medicine and Biology Society: Vol 3 (p. 1787-1788).

Piscataway, N.J.: IEEE.

Pear, J.J. & Legris, J.A. (1985). Shaping by automated tracking of an arbitrary operant response. Journal of the Experimental Analysis of Behaviour, 47, 241-247.

Powell, R.W. & Kelly, W. (1975). A method for the objective study of tool-using behaviour. Journal of the Experimental Analysis of Behaviour, 24, 249-253.

Remington, B. & Clarke, S. (1983). Acquisition of expressive signing by autistic children: An evaluation of the relative effects of simultaneous communication and sign-alone training. Journal of Applied Behaviour Analysis, 16, 315- 328.

Sidman, M. (1960). Tactics of scientific research: Evaluating experimental data in psychology. New York: Basic Books.

Skinner, B.F. (1953). Science and human behaviour. New York: Macmillan. (Ch. 6, pp. 91-107).

Skinner, B.F. (1954). The science of learning and the art of teaching. Harvard Educational Review, 24, 86-97.

Skinner, B.F. (1962). Two "synthetic social relations". Journal of the Experimental Analysis of Behaviour, 5, 531-533.

Yoder, P.J. & Layton, T.L. (1988). Speech following sign language training in autistic children with minimal verbal language. Journal of Autism and Developmental Disorders, 18, 217-229.

Appendix A

Speech Shaping Cluster Program

User's Guide

Sharon Cairns

University of Manitoba

1989

TABLE OF CONTENTS

SPEECH SHAPING CLUSTER PROGRAM
USER'S GUIDE

Introduction

The speech shaping cluster program, designed by Deb Roy, University of Manitoba, is extremely user friendly. In fact, it is so user friendly that a user's guide is not really necessary.

The cluster program is designed to allow an experimenter to store on disk a variety of speech utterances in "clusters". Later, when training speech, the subject's vocalization is compared to the reference values of all the utterances in the cluster and distance scores (second norm) will be computed. The screen display will show you all of the distance measures for each reference but when the data is stored on disk, only the lowest distance measure (closest approximation of the sound) will be recorded.

Ready for more detail? Let's have a look at the program and how to get started.

GETTING STARTED

Turn the computer on. Insert "Cluster" disk into drive A. The program will boot by itself, and the screen will show the following display:

```
Do you wish to:
       (T)rain Speech
       (R)ecord Speech
  Press 'T' or 'R' to select...
```

RECORDING SPEECH

As prompted, to record speech, simply press 'R'. Very briefly on the screen you will see:

```
Run Template Trainer II
```

Then the Record Main Menu will be displayed from which you may make a selection.

```
   Speech Recognition System
1)  Edit Templates
2)  View Templates
3)  Disk I/O
4)  List Templates
5)  Test Recognition
6)  Group Templates
7)  Quit System
      Enter Number to Select
```

1) EDIT TEMPLATES

This selection accesses a menu from which you create new template clusters, retrain or edit a template in a cluster, clear templates or return to the main menu.

```
Edit Templates
 1)  Train New Template(s)
 2)  Retrain a Template
 3)  Clear Templates
 4)  Quit to Main Menu
 Enter Number to Select
```

1) <u>Train New Template(s)</u> This is the selection to make if you wish to record a new cluster of sounds. Select by entering '1'. The screen display will then go to:

```
          Make Templates
How many templates do you wish to
   train < Default = 1 >
```

Enter the number of sounds you wish your cluster to contain, e.g. '4' and press return. The next screen you see is:

```
          Make Templates
How many templates do you wish to
   train < default = 1 >
   How many passes will be made for
   each template < default = 2 >
```

In response to this screen, enter that you wish to make one pass, otherwise the computer will use averaged sounds for the reference template. Enter '1', press Return to reveal the next screen.

```
Name of Template 1?
```

Enter the name you wish to call your first sound. Use a descriptive name for easy identification of your results. e.g. 'ahfemchild' (for a female child saying ah). Press Return to see the next screen.

Say 'ahfemchild'

This screen prompts you to have the speaker say the target sound into the microphone. The prompt will be the name you previously assigned the sound. Once the computer has recorded this sound, you will then be prompted for the names and vocalizations for the remaining number of templates as in the above two steps. Following the recording of the chosen number of templates you will be returned to the Edit Templates Menu.

2) Retrain a Template - Choosing option 2 gives you the opportunity to re-record one or more of the templates recorded above. You may require this option if for some reason the uttered sound was not of the desired quality, or if there was a loud background noise which might distort the template's value. To use this option enter '2' and the following screen appears:

How many templates do you wish to retrain?

Enter the number of sounds you wish to revise. You may choose any number so long as it does not exceed the number of templates initially recorded. Press Return for the next prompt.

Number of Template to be retrained

Enter the number of the template you wish to revise (i.e. if you were dissatisfied with the third sound recorded, enter '3'). Press Return for the next screen.

How many passes will be made for
each template <default = 2>

Once again you want to make only one pass to avoid averaged sounds, so enter '1' and press Return.

Name of Template # (shows no. chosen above)
?
<Default = (name previously given that template)

Pressing Return keeps the name previously chosen for that template or you can choose a new name simply by typing a new name in before pressing return. When prompted, have the speaker say the sound and then the computer returns you to the Edit Templates Menu.

3) <u>Clear Templates</u> - This option can be chosen if you were not satisfied with any of your recording and you wish to start over rather than editing all of the recordings. To choose enter '3'.

Are you sure you want to clear?
All current data will be lost

Here you have a chance to change your mind or proceed with the erasure of the recordings. Enter 'Y' if you want to erase, 'N' if you do not want to erase. Then once again you are returned to the Edit Templates Menu.

4) <u>Quit to Main Menu</u> - Choosing '4' returns you to the Record main menu. The first thing you should do here is choose 3) Disk I/O by pressing '3' otherwise you will lose your recordings.

2) VIEW TEMPLATES

This option simply allows you a view of the templates.

Note: never choose this option prior to saving the templates, otherwise they will be lost.

3) DISK I/O

Although I will go through these in sequence for purposes of uniformity, recall that the very first thing you want to do in this sub-menu is go to 3) Save Present Templates. Disk I/O gives you the following menu:

```
    Disk I/O
1) Catalogue Disk
2) Load New Templates
3) Save Present Templates
4) Quit to Main Menu
```

1) Catalogue Disk - Entering '1' displays a list of all files contained on the disk. From this list you will be able to identify the names of clusters that have been recorded and then make your choice of one to load. Once the list of files has been displayed, pressing any key will return you to the Disk I/O Menu.

2) Load New Template

```
              Load Template
Are you sure you want to write over the
  present templates? (Y/N)
```

'N' returns you to the Disk I/O Menu. Before you activate 'Y' ensure that you replace the write protect tab on the disk (just to avoid any possibility of erasing anything you don't mean to erase. Now enter 'Y'.

```
Enter base name of template set to be
  loaded?
```

Now you can enter the name of a cluster that has previously been saved on the disk. If the name you enter is on the disk, that cluster will be loaded for training purposes. If you did not enter the name of the cluster correctly you will see the following message:

```
Disk I/O Error ... Any Key to Return
```

Pressing any key returns you to the Disk I/O Menu from which you can access a catalogue of the disk files.

3) Save Present Templates - This is the option for saving the templates you have just recorded. Entering '3' prompts you as follows:

```
   Save Templates
Name of Templates?
```

Enter the name of your cluster of sounds. Once again try to be descriptive to aid in future identification.

```
Storing Template in Memory
```

```
Saving Templates on Disk
```

The above screens are displayed automatically once you have entered the name of your cluster. In order to save the cluster to the disk, the write protect tab must be removed from the training disk otherwise you will be given an error message. Once the cluster is storied on disk, you are returned to the Disk I/O menu.

4) Quit to Main Menu - This option does exactly as it states, It returns you to the Record Speech Main Menu.

4) LIST TEMPLATES

This option simply gives you a list of the templates that have been recorded. Return will take you back to the main menu.

5) TEST RECOGNITION

This option is most useful if each of your templates in a cluster is a different sound and you want to test that trial sounds come out closest to that sound's reference template.

6) GROUP TEMPLATES

This option is not needed in simple speech training, therefore, you may ignore it.

7) QUIT SYSTEM

Turns off the program. You may return to the choice of training or recording speech by pressing control open apple reset or by turning the machine off and back on again. It is recommended that you use the option of turning the equipment off and back on again since a very large portion of the available memory is used by the program.

TRAINING SPEECH

From the first menu when you turn the computer on, press 'T'. The following screens will then appear:

Run Cluster

Loading the Speech Training System

```
Loading Default Template Cluster
```

```
          Speech Training System

    1)  Disk I/O (Template:  AH1.Cluster)
    2)  Change Recognition Parameters
    3)  Train Vowel
    4)  Monitor Energy Level
    5)  Change Goal
    6)  Set Auto-Reinforcement
    7)  Quit System
           Enter Number to Select
```

1) DISK I/O

Choosing 1 will allow you to load the comparison cluster of your choice. The following menu will appear when you enter '1'.

```
            Disk I/O

      1)  Catalogue Disk
      2)  Load New Templates
      3)  Save Data to Disk
      4)  Load New Speech Recording
      5)  Quit to Main Menu
    Present Template:  AH1.Cluster
    Present Speech Recording AH2.D
    Enter Number to Select
```

1) <u>Catalogue Disk</u> - This choice will provide you with a catalogue of the files on the disk. It is useful if you forget the name of the cluster you wish to use.

2) <u>Load New Templates</u> - This choice allows you to enter the

name of the cluster you wish to use for training. When prompted, simply enter the name of the cluster you want to use.

3) Save Data to Disk - After your training session is complete, return to this choice. Then place a blank, formatted disk in the drive to save your data.

4) Load New Speech Recording - This is the verbal instruction given by the computer (e.g. "Say Ah"). You may load any file that has been previously digitally recorded. These can be recognized in the files by ones that end with (.D) e.g. AH2.D. Once again simply enter the recording name you wish to use when prompted.

5) Quit to Main Menu - As suggested by the name, this choice will return you to the Training Main Menu.

2) CHANGE RECOGNITION PARAMETERS

Entering '2' will give you the following screen:

```
        Change Recognition Parameters
Rejection Value = 190
Reject Par Inc Step = 2 (steps closer to target)
Reject Par Dec Step = 1 (steps away from target)
Number of 'Hits' to Progress = 1
Number of 'Misses' to Regress = 1
Space to Select, Arrows to Change
Press Q to Quit to Menu
```

As indicated, use the space bar to highlight the option you wish to change. The <-- arrow decreases and the --> arrow increases the value of the highlighted parameter.

The rejection value is the starting point for training. You will need to change this at the start of each session depending where the child left off at the end of the last session.

The Reject Par Inc Step is set at the default value of 2. This is the number of steps you wish the rejection value to shrink (Increase closeness to target). For the current study, this will have to be set at 3 as we wish the target range for reinforcement to contract by 3 each time the child hits the reinforcement range.

The Reject Par Dec Step is just the opposite of the Inc Step. It is the number of steps the reinforcement range will expand (decreased precision) each time the child fails to reach the reinforcement range. The current study is using a value of 1 for this parameter.

The number of Hits and Misses are the number of times you would like the child to hit or miss the range for reinforcement before the parameters Step.

3) TRAIN VOWEL

Choosing this option from the main menu gives you the following choice:

```
Perform Recognition
Will there be Experimenter Input?
```

'N' (No) will result in trials being advanced automatically. 'Y' (Yes) means the experimenter will have to push Y or N on the blue box following each trial. The advantage of this method is that it gives you the opportunity to agree (Y) or disagree (N) with the computer's assessment. If you disagree, the trials on which you disagreed will be earmarked on your data output.

Before you respond to this question, turn the speaker on. After you enter the 'Y' or 'N' to the question, the computer will proceed with the first trial by saying "Say AH". Once the child responds you will be shown the following screen:

```
                 Rejection = ###
Sound Name          :  Score (e.g. 71.1969101)
Sound Name          :  Score
Sound Name          :  Score
Sound Name          :  Score

     Trial # 1 Correct (or incorrect)
    Closest is 71.1969101; (Sound Name)
      Correct Evaluation?   (Y/N)
```

If the child's answer was correct (at or below the rejection value) the computer will say "Good". If the response was incorrect the computer will emit a small beep.

If you have chosen to have experimenter input, the computer will wait until you press the 'Y' or 'N' on the blue box and then will proceed with the next trial. If you chose to have no experimenter input, the computer will flash the above screen and then proceed to the next trial.

From here, simply proceed to carry out the number of trials you wish to conduct. Then you must return to Disk I/O to save your data.

4) MONITOR ENERGY LEVEL

You should make this selection before beginning any training session. It is used to test if the mic batteries are good and to adjust for ambient noise. As soon as you hit '4' the following screen will appear:

```
              Press 'Esc' to Quit to Menu


_____________________________________________________
0      5      10      15      20      25      30      35

                    Energy Level
```

Put the headphones on and make some test sounds and adjust the level by turning the knob indicated on the diagram. When you are satisfied with the level, press 'Esc' to return to the main training menu.

5) CHANGE GOAL

This option is used to adjust the minimum value the region for reinforcement will contract to during training. It is important to set this at a realistic level otherwise in the course of training the child will be emitting good quality sounds and not receive reinforcement. One way of estimating this value would be for you to do a practice run with yourself emitting the training sounds. Go through a number of trials and use the lowest value you are able to reach (this should be somewhere in the range of 50 to 90). Round the number off to a whole number for purposes of entering it here. When you select '5' you will see the following screen:

```
Currently, Goal = 50
  New Goal?
```

Enter the number.

6) SET AUTO-REINFORCEMENT

This option is used if you have hooked up some form of reinforcement device to the computer and wish this to be automatically activated when the child emits a correct response.

7) QUIT SYSTEM

Once again, this option does exactly as the name implies - it takes you out of the speech training system.