Development of a Computer-based Program for Training Speech Rate
Mise au point d'un programme informatisé pour l'éducation du débit de la parole

Nancy Thomas-Stonell
Department of Speech-Language Pathology & Audiology
The Hugh MacMillan Rehabilitation Centre

Michael McClean
Graduate Department of Speech Pathology
University of Toronto

Leslie Dolman
Designing Aids for Disabled Adults, Toronto

Key Words: computer, dysarthria, speaking rate, speech, intelligibility

Abstract

This paper describes the development of a Computer-Based Speech Training (CBST) program for modifying speaking rate. The program, called the "Stepping Stones Game," provides both visual and auditory feedback on overall speaking rate, articulation, and pause time. Preliminary results suggest that it is an effective tool for modifying speech rate and improving intelligibility in neurologically-impaired speakers.

Résumé

L'article décrit la mise au point d'un programme informatisé d'éducation de la parole pour modifier le débit de la parole. Appelé "Stepping Stones Game," le programme fournit des rétroactions visuelles et auditives sur le débit de la parole, l'articulation et le temps de pause. Les premiers résultats laissent supposer qu'il s'agit d'un instrument efficace pour modifier le débit de la parole et améliorer l'intelligibilité chez les personnes souffrant de troubles neurologiques.

The consequences of reduced speech intelligibility affect many aspects of life. Children with speech impairments experience difficulties with social interaction, academic performance and the development of language skills (Simon, 1985). Improved methods of speech production training are needed to meet the needs of these children.

Computer-Based Speech Training (CBST) is now recognized as a distinct area of study within the communication sciences and speech-language pathology (Watson & Kewley-Port, 1989). The general goal of work in this area is to enhance speech assessment and remediation processes through use of computer technology. Most CBST systems provide feedback on aspects of speech, such as pitch, loudness, and voicing. There is a need for CBST systems that address additional aspects of speech production, such as speech rate, nasality, and consonant articulation (Thomas-Stonell, 1989). The addition of these areas would make CBST systems more appropriate for remediation of the majority of speech production deficits. Recent advances in technology permit the development of such programs.

Training of speech rate and rhythm is particularly important for individuals who have speech difficulties resulting from neurological impairment (dysarthria). Excessive nasalization, disordered speech prosody, imprecise articulation, and variable speech rate are often associated with damage to the neuromuscular systems regulating speech production (Darley, Aronson, & Brown, 1969). This results in reduced speech intelligibility. Improving intelligibility is the primary clinical objective with dysarthric speakers (Yorkston, Beukelman, & Traynor, 1984). One approach to increasing intelligibility is to improve a speaker's ability to control speaking rate. Abnormal speech rate is prevalent among dysarthric speakers. Darley, Aronson, and Brown (1975) found that 80% of ataxic dysarthric speakers deviated from normal speaking rates. The benefits of rate control therapy are documented by marked enhancement of speech intelligibility scores in some patients (Yorkston, Beukelman, & Bell, 1988). At the later stages of recovery, an optimal speaking rate, which allows a speaker to maximize speech naturalness while maintaining speech intelligibility above 95%, may also become an important speech goal (Yorkston & Beukelman, 1981).

Techniques for rate control vary from rigid systems, such as pacing and alphabet boards that encourage pausing between words, to rhythmic cuing and eventual self-monitoring (Yorkston & Beukelman, 1981). Overall speaking time consists of two components - speaking or *articulation time* and *pause time* (Yorkston et al., 1988). Visual feedback in the form of oscilloscope displays has been used successfully to modify both pause and articulatory time (Berry & Goshorn, 1979). The effects of feedback on rate control are only beginning to appear in the literature (Yorkston, Hammen, Beukelman, & Traynor, 1990). Hyland and Weismer (1988) investigated the effect of visual/verbal, verbal/quantitative, and verbal/ qualitative feedback on rate control. They found that precise quantitative feedback (e.g., as the number of units in centiseconds that productions were greater or less than desired utterance duration) whether verbal or visual, is superior to qualitative feedback (e.g., as gen-

eral information as to whether productions were "a little fast" or "a little slow").

Our research has revealed only two rate-training software packages commercially available in North America. One is the Slow Speech-Rate Drill for Dysarthric Speakers (Goldojarb & Secor, 1988) from Sunset Software. The other is the Pacer/Tally software package (Beukelman, Yorkston, & Tice, 1988) from Communication Skill Builders. The Slow Speech-Rate Drill for Dysarthric Speakers displays text at controlled rates of presentation. The text varies in complexity from single words to sentences. The text format prevents the program's use with non-reading children. In addition, the software uses a drill and practice format, which may be of limited motivational value. There are no research studies evaluating the program's clinical effectiveness (Schwartz, 1988). The Pacer/Tally software package uses a text passage with a cursor which moves at preselected rates. The documentation includes supporting data for the program and clinical/research applications. A product review (Schwartz, 1989) rated the program as excellent, however, the text format limits its applicability to older school-aged children, adolescents, and adults.

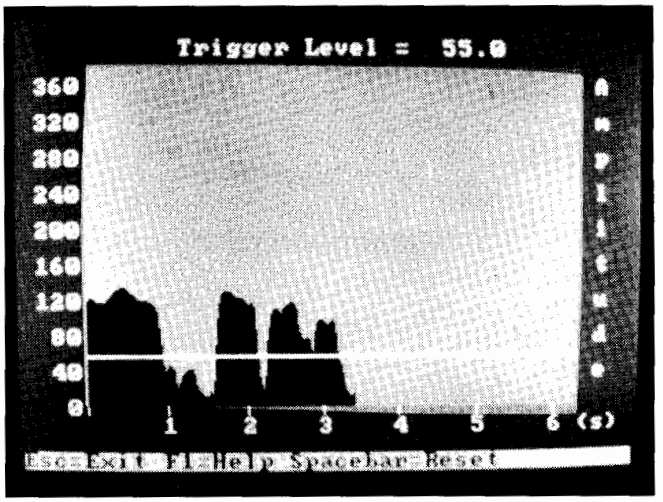
This paper describes the design and evaluation of a CBST software program for modifying speaking rate called "The Stepping Stones Game." This program is designed to be compatible with the IBM SpeechViewer system (Thomas-Stonell, 1989) and would enhance its utility with neurologically-impaired populations. The specific goals of the program are as follows:

1. To develop a program for training rate control that uses both visual and auditory feedback to modify speech rate and subtle timing changes within short utterances;
2. To create a program that is useful and appealing to young non-reading children as well as school-aged children, adolescents, and adults;
3. To provide maximum clinical flexibility so that speech rate can be systematically increased or decreased using a variety of rate modification techniques (i.e., pause insertion, speech phrasing, changes in articulation time).

The Stepping Stones Game

The Stepping Stones Game is designed to provide both visual and auditory feedback of speech using a goal-oriented game format. Malone (1981) developed a theory of intrinsically motivated instruction based on the study of several computer games. He identified three characteristics of intrinsically motivated learning environments: challenge, fantasy and curiosity. Challenge is hypothesized to depend on goals with uncertain outcomes. The game format was chosen to provide challenge. Intrinsic fantasies (i.e., fantasies which are intimately related to the skill being used) are said to be more interesting and more instructional than extrinsic fantasies. A stepping stones fantasy was chosen to reinforce the concept of pausing because a person naturally alternates movement and pausing when crossing water using stepping stones.

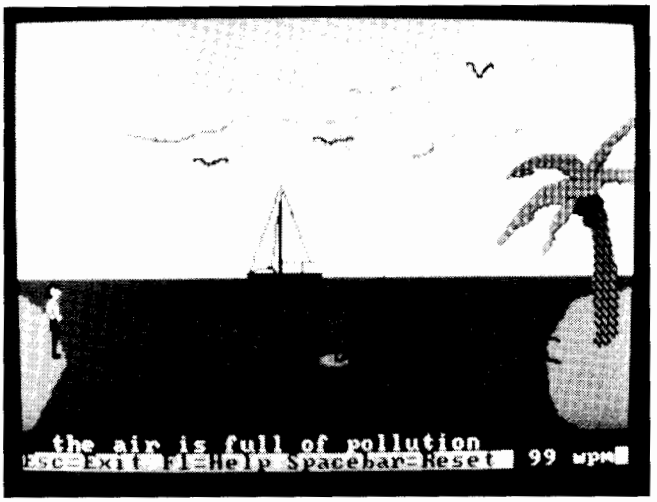
Figure 1. Screen 1 from the Stepping Stones Game. The sentence contour of the target utterance is displayed. The duration of speaking and pause time is determined by the points of intersection between the horizontal cursor (adjustable) and the sentence contour.



To play the game, children must match the temporal characteristics of a target utterance (see Figure 1). The clinician designs individualized target utterances for children to match. The child's own best production may also be used for target utterances. The target sentence is displayed as a series of stones which form a path to an island (see Figure 2). Acoustic energy associated with the speech signal causes a man to move across the screen to the island by stopping at each stepping stone. The stones correspond to pauses in the target sentence. Any speech sound above an adjustable threshold level causes the man to move across the screen. The duration of the acoustic signal below the threshold level defines pause length. To successfully cross to the island, children must match the pattern, both articulation (or speaking) time and pause time. If the pattern is matched correctly, the man reaches the island and jumps up and down, with accompanying music. If the target and child's pattern do not match, the man falls into the water (different music), disappears, and then reappears at the starting point for another attempt.

The game utilizes instructional visual and auditory feedback. The screen is constructed so that the distance between the stones corresponds to the length of speaking or articulation time. The size of the stones corresponds to the length of the pauses. If articulation time is too short, the man falls short of the stone and lands in the water. If articulation time is too long, the man overshoots the stone and

Figure 2. Screen 2 from the Stepping Stones Game. The stones correspond to pauses in the sentence. The distance between the stones corresponds to the length of speaking time.



falls into the water. Once the man does land on a stone, it turns green and begins to sink (visually displayed by shrinking), gradually becoming smaller until it has vanished completely. The man must pause on the stone until it turns green, but not so long that the stone sinks completely. A track feature (dotted line) displays the man's route after each trial. This allows both the clinician and the child to determine the cause of an error (i.e., speaking time too long or short, or pause time too long or short). Between trials the digitized target utterance can be replayed as many times as desired. This auditory feedback helps children learn to monitor their own speaking rate. After a trial, the contours of both the target sentence and the child's sentence can be displayed on a split screen. Auditory playback of both sentences is available so that both the clinician and child can see and hear the differences and similarities between the two sentences.

Hardware Requirements

The rate control game was designed to be compatible with the IBM SpeechViewer system (Crepy, Denoix, Destombes, Rouquie, & Tubach, 1983). The game requires the analog to digital (A-D) board used by the SpeechViewer system (SpeechViewer Hardware Option). This board processes speech at a rate of 9,600 samples per second with a 12 bit Analog-to-Digital converter. The speech is prefiltered with an effective roll-off of 28 dB per octave, attenuating the frequencies above 3.8 kHz (International Business Machines Corporation, 1988). The Audiotechnica microphone and speakers provided by the IBM Speechviewer system are used with the Stepping Stones Game. SpeechViewer requires an IBM PS/2 computer (either model 25 or 30) with 512K of memory and a colour monitor. The software is designed to execute on IBM DOS 4.0.

Software Description

The program was designed with two screens. A set-up screen (screen 1), shown in Figure 1, and a game screen (screen 2), shown in Figure 2. Screen 1 was designed for entering and displaying the target utterance. Screen 2, the game screen, displays the target utterance as stepping stones. The screen designs parallel the layouts, function keys, and format used by IBM in the development of SpeechViewer. When initiating a new game, the clinician is always placed into screen 1. This screen allows the input of a target sentence up to 10 words or 6 seconds in length. The intensity of the acoustic signal in decibels is displayed on the screen. This display is used to adjust a threshold level indicated by a horizontal line. The threshold level should be set by the clinician to a level slightly above the noise floor. The matching algorithm uses the points where the speech signal and the threshold cross to define the temporal structure of the target sentence. Specifically, these points define the onset, offset, and durations of speech and pause boundaries.

To successfully cross to the island, children must match the temporal pattern of threshold crossings contained in the target utterance. If a child were to produce a perfect match, the timing of these threshold crossings would be identical to the target utterance. Since this is very unlikely to occur, the tolerance level for the match can be varied in steps of ±13 ms to a maximum allowance of 299 ms. In order to obtain rate information, the clinician times each target sentence using the spacebar. The clinician then enters the number of words or syllables spoken and may type in the target sentence, if desired. The rate information, calculated either as number of words per minute or number of syllables per minute, is displayed in screen 2. This allows the clinician to modify speech rates in a controlled step by step manner. If the clinician typed in the target sentence in screen 1, it may be displayed across the bottom of the game screen by using a keyboard activated toggle switch. The Stepping Stones program allows target sentences to be saved (digitized) and recalled for use in other speech sessions. This feature allows the clinician to enter the same sentence at a variety of speaking rates. These sentences can be used across sessions to maintain consistency.

General Comments

Initial pilot testing of the program was completed on three subjects. The clinical work was intended to provide human factors information to the developers, as well as a preliminary assessment of clinical benefit. The rate training software

was evaluated during one phase of a single-subject multiple baseline across behaviors study which is described in a companion paper (Thomas-Stonell, McClean, & Hunt, 1991). All three subjects reported that the game format was fun, motivating, and easy to understand. They became skilled at interpreting the feedback displays and successfully modifying their own speech rate.

Target sentences were entered by the clinician and saved at various rates. Subjects were asked to match sentences at progressively faster speaking rates. The feedback provided by the game enabled the therapist to see whether mismatches were due to difficulties matching articulation time or pause time, and allowed the therapist to target the specific elements contributing to rate difficulties. The subjects occasionally used their own best productions as target utterances. Target sentences were evaluated by both the subject and the clinician to determine whether they were acceptable. Sentences were judged according to clinical goals of each subject. Acceptable target sentences were saved and used across speech sessions. Subjects were able to achieve a greater number of successful matches when using their own models.

Clinical goals for the three subjects varied. The goal for Subject 1 was to increase speech rate, while maintaining sentence intelligibility over 95% (optimum rate). For Subject 2, the goal was to reduce speech rate and maximize intelligibility. Subject 3 continued voicing between word segments which lowered his intelligibility. Therefore, the goal for Subject 3 was to insert short pauses between words to increase speech intelligibility, while increasing speech rate by shortening articulation time. The program was effective for modifying speech rate with all three subjects. The ability to maintain the newly acquired rate outside of speech sessions varied across subjects. The rehabilitation team observed changes in speech rate during other daily activities for two of the three subjects.

The Stepping Stones game allows the clinician to target such goals as: increasing or decreasing speech rate, and reducing or increasing articulation and/or pause time. Both natural phrasing and inserting pauses between words can be taught. Thus, it is possible to target overall rate modifications to enhance intelligibility, while preserving natural phrasing.

Acknowledgements

This research was supported by The Ontario Ministry of Colleges and Universities (Grant #T08-010) and IBM Canada Ltd. We wish to acknowledge the support of Mr. R. Mighton (IBM Canada Ltd.), Mr. John McTyre, Ms. Linda Boyer, Mr. John Roberts, Mr. Walt Nawrocki (IBM; Boca Raton) and the administrative support of Ms. Ruth Gannon, Dr. S. Naumann and Dr. M. Milner (HMRC).

Address all correspondence to: N. Thomas-Stonell, The Hugh MacMillan Rehabilitation Centre, 350 Rumsey Rd., Toronto, ONT, M4G 1R8.

References

Berry, W. R., & Goshorn, E. L. (1979, November). *Oscillo-scopic feedback in the treatment of ataxic dysarthria.* A paper presented at ASHA Convention, Atlanta, Georgia.

Beukelman, D. R., Yorkston, K. M., & Tice, R. L. (1988). *Pacer/tally for pacing and tallying responses* [computer program]. Tucson Az: Communication Skill Builders.

Crepy, H., Denoix, B., Destombes, F., Rouquie, G., & Tubach, J.P. (1983). Speech processing on a personal computer to help deaf children. In R. Mason (Ed.), *Information Processing 83: Proceedings of the IFIP Ninth World Computer Congress* (pp. 19-23). Paris, France: North-Holland.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Clusters of deviant dimensions in the dysarthrias. *Journal of Speech and Hearing Research, 12,* 462-496.

Darley, F. L., Aronson, A. E. & Brown, J. R. (1975). *Motor Speech Disorders.* Toronto: W. B. Saunders Company.

Goldojarb, M., & Secor, J. (1988). *Slow Speech-Rate Drill for Dysarthric Speakers* [Computer program]. Los Angeles, CA: Sunset Software.

Hyland, J. D., & Weismer, G. (1988). The effects of three feedback modes on the ability of normal geriatric individuals to match speaking rate. *Journal of Speech and Hearing Disorders, 53*(3), 271-279.

International Business Machines Corporation (1988). *IBM Personal System/2 SpeechViewer: A Guide to Clinical and Educational Applications.* Atlanta, GA: Author.

Malone, T. W. (1981). *Toward a theory of intrinsically motivating instruction. Cognitive Science, 4,* 333-369.

Schwartz, A. (1988) [Review of Slow speech-rate drill for dysarthric speakers]. *Computer Users in Speech and Hearing, 4*(1), 62-63.

Schwartz, A. (1989). [Review of *Pacer/Tally for Pacing and Tallying Responses]. Computer Users in Speech and Hearing, 5* (1), 40-43.

Simon, C. (Ed). (1985). *Communication skills and classroom success: Therapy methodologies for language-learning disabled students.* San Diego: College Hill Press, (pp. xi).

Thomas-Stonell, N. (1989). Resource Reviews [Review of Speech-Viewer] *Journal of Speech-Language Pathology and Audiology, 13*(2), 59-60.

Thomas-Stonell, N., McClean, M., & Hunt, E. (1991). Evaluation of the SpeechViewer computer-based speech training system with neurologically-impaired individuals. *Journal of Speech-Language Pathology and Audiology,15,* (add final pages). Manuscript submitted for publication.

Watson, C. S., & Kewley-Port, D. (1989). Advances in computer-based speech training (CBST): Aids for the profoundly hearing-impaired. *Volta Review, 91*(5), 29-45.

Yorkston, K. & Beukelman, D. (1981). Ataxic Dysarthria: Treatment sequences based on intelligibility and prosodic considerations. *Journal of Speech and Hearing Disorders, 46,* 398-404.

Yorkston, K., Beukelman, D., & Bell, K. (1988). *Clinical Management of Dysarthric Speakers.* Boston: College-Hill Press, Inc.

Yorkston, K., Beukelman, D., & Traynor, C. (1984). *Computerized assessment of intelligibility of dysarthric speech.* Publications: Tigard Oregon.

Yorkston, K., Hammen V. L., Beukelman, D., & Traynor, C. D. (1990). The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders, 55*(3), 550-559.