An Internet-Based Telerehabilitation System for the Assessment of Motor Speech Disorders: A Pilot Study

Anne J. Hill
Deborah G. Theodoros
Trevor G. Russell
Louise M. Cahill
Elizabeth C. Ward
The University of Queensland, Brisbane, Queensland, Australia

Kathy M. Clark
Princess Alexandra Hospital, Brisbane, Queensland, Australia

Purpose: This pilot study explored the feasibility and effectiveness of an Internet-based telerehabilitation application for the assessment of motor speech disorders in adults with acquired neurological impairment.
Method: Using a counterbalanced, repeated measures research design, 2 speech-language pathologists assessed 19 speakers with dysarthria on a battery of perceptual assessments. The assessments included a 19-item version of the Frenchay Dysarthria Assessment (FDA; P. Enderby, 1983), the Assessment of Intelligibility of Dysarthric Speech (K. M. Yorkston & D. R. Beukelman, 1981), perceptual analysis of a speech sample, and an overall rating of severity of the dysarthria. One assessment was conducted in the traditional face-to-face manner, whereas the other assessment was conducted using an online, custom-built telerehabilitation application. This application enabled real-time videoconferencing at 128 kb/s and the transfer of store-and-forward audio and video data between the speaker and speech-language pathologist sites. The assessment methods were compared using the J. M. Bland and D. G. Altman (1986, 1999) limits-of-agreement method and percentage level of agreement between the 2 methods.
Results: Measurements of severity of dysarthria, percentage intelligibility in sentences, and most perceptual ratings made in the telerehabilitation environment were found to fall within the clinically acceptable criteria. However, several ratings on the FDA were not comparable between the environments, and explanations for these results were explored.
Conclusions: The online assessment of motor speech disorders using an Internet-based telerehabilitation system is feasible. This study suggests that with additional refinement of the technology and assessment protocols, reliable assessment of motor speech disorders over the Internet is possible. Future research methods are outlined.

Key Words: telepractice, rehabilitation, assessment, speech disorders

Dysarthria is a collective term for a cluster of motor speech disorders of neurogenic origin that are characterized by weakness, slowness, imprecision, and/or incoordination of movement of the speech musculature (Darley, Aronson, & Brown, 1975; Yorkston, Beukelman, & Bell, 1988). This speech disorder can reduce speech intelligibility, which may have a significant impact on an individual's ability to communicate in social and vocational contexts and, hence, impact quality of life. It has been estimated that dysarthria accounts for up to 54% of communication disorders associated with stroke, traumatic brain injury, and progressive neurological disorders such

as Parkinson's disease and multiple sclerosis (Duffy, 2005).

The capacity to provide effective evaluation and rehabilitation to people with dysarthria, however, may be limited by individual difficulties in accessing speech-language pathology services. Difficulty accessing services may arise because of distance from service providers, individual physical disabilities, financial limitations, and institutional prioritization of cases. Telerehabilitation, a service delivery model in which rehabilitation services are provided at a distance using information and communication technology, may offer an alternative or adjunct service delivery model to facilitate or enhance access to services and, in turn, assist recovery from neurological impairment (Torsney, 2003).

Research into telerehabilitation in the area of speech-language pathology has revealed some promising outcomes, and it is generally accepted that certain speech-language pathology services are well suited to the telerehabilitation service delivery model (Brennan, Georgeadis, Baron, & Barker, 2004; Rosen, 2004). However, evidence to support the accuracy and validity of telerehabilitation services is still in its infancy. A literature review by Hill and Theodoros (2002) revealed that although research into the use of telerehabilitation in speech-language pathology began in the 1970s, only a small number of published empirical studies have been reported. Of these, only five specifically focused on the assessment of speech and language disorders (Hill & Theodoros, 2002).

The first study to investigate the assessment of communication disorders using telecommunication technology was by Wertz et al. (1987). Thirty-six clients were assessed using a battery of measures in three different conditions: traditional face-to-face, closed-circuit television, and "computer controlled video laser disk over the telephone" (p. 117). Although the results yielded an average of 92% agreement in diagnosis between the three conditions, it is notable that the diagnoses were of a general nature (e.g., aphasia vs. dysarthria) and perhaps not refined to a level required for treatment planning. Because of the small sample size in the original study, Wertz et al. republished the results with an additional 36 clients in 1992. Again, diagnosis was of a general nature, with no detail as to the number of clients diagnosed into each category (Wertz et al., 1992).

The next large-scale study to focus on the assessment of speech and language disorders by telerehabilitation methods was reported by Duffy, Werven, and Aronson (1997). In their study, assessment of communication disorders via satellite videoconferencing was compared with face-to-face assessment for accuracy of diagnosis. In this study, 8 clients underwent live assessment via satellite videoconferencing. The resultant diagnoses were compared with the diagnoses made by the on-site speech-language pathologist (SLP). To further test the reliability of the satellite videoconferencing method, 24 video-recorded samples of neurogenic speech disorders were transmitted across the satellite connection for diagnosis. Results were encouraging, with 96% agreement of diagnosis between the methods. Duffy et al. also conducted a retrospective analysis of 150 telerehabilitation speech and language consultations that took place at various Mayo Clinic centers between 1987 and 1994. Analysis revealed that telerehabilitation consultations were reliable, with speech and language diagnosis (e.g., dysphonia, dysarthria, aphasia, apraxia of speech) consistent with lesion localization and medical diagnosis. The frequency of uncertain diagnosis (13% of sample) in the telerehabilitation assessment was only slightly higher than that recorded in traditional face-to-face assessment at Mayo Clinic centers. However, it is important to note that several of the clients were recommended for a subsequent face-to-face assessment session. This finding suggests that some disorders may not be suited to assessment by telerehabilitation methods or that the technology of the time was not advanced enough for assessment of these disorders. As technology becomes more refined and more readily available, it is essential that researchers continue to develop and evaluate valid and reliable assessment and treatment protocols for a range of communication disorders.

Most studies of telerehabilitation in speech-language pathology have used expensive, specialized, videoconferencing equipment. Where the cost of installing and maintaining the equipment is significant, the equipment may itself become a limiting factor in accessing these services. If specialized videoconferencing equipment is required, clients are obliged to travel to a specific center that houses such equipment to access the service. In addition, videoconferencing using such specialized equipment for telerehabilitation usually requires that any materials used in the assessment or treatment of the client be either mailed or faxed to the client site.

In recent years, a number of studies into speech-language pathology services and telerehabilitation have examined the use of computer-based telerehabilitation systems (Brennan et al., 2004; Mashima et al., 2003). Such systems may help to alleviate difficulties in accessing rehabilitation services, because of the relatively low cost of installation and maintenance of computers as well as their general availability. Using a computer-based videoconferencing system that utilized Microsoft's Windows NetMeeting to share the Kay Elemetrics Multispeech Signal Analysis Workstation, Mashima et al. (2003) reported no significant difference between traditional and remotely delivered voice therapies for a number of outcomes. These outcomes included perceptual rating of voice quality, acoustic variance, speaker satisfaction, and laryngeal examination. Brennan et al. (2004) used the computer not only to enable real-time videoconferencing between locations but to concurrently play prerecorded audio files and accompanying images in a story-retelling task. The study extended telerehabilitation beyond the bounds of videoconferencing and into a dynamic and interactive mode only offered by computer-based telerehabilitation applications. This flexibility within computer-based telerehabilitation may allow for the use of established practices, such as the ability to deliver standardized assessment tools in an online capacity.

In Australia, the wide distribution of the population and increased use of technology in regional and remote areas

support the use of computer-based telerehabilitation systems. Australia's population is currently 20.2 million, of which 66.3% live in metropolitan areas and 33.7% live in regional and remote areas (Australian Bureau of Statistics, 2004). Census data from 2001 reported computer use in regional and remote areas to be 41%, only slightly less than the national average of 42% (Lloyd & Bill, 2004). Similarly, 2002 Australian Bureau of Statistics data reported that 46% of Australian households had home Internet access and that 58% of adults accessed the Internet (Lloyd & Bill, 2004). Within the older age group of adults, age 55 years and over, the rate of accessing the Internet has grown from 15% in 1998 to 55% in 2002 (Australian Bureau of Statistics, 2004).

The application of computer-based telerehabilitation is also supported within populations of people with disabilities. Within Australia, 48% of people with disabilities reported having used a computer in 2003, whereas 39% had accessed the Internet (Australian Bureau of Statistics, 2004). In the United States, a survey by Ricker et al. (2002) into the telerehabilitation needs of persons with acquired brain injury reported considerable interest within this population for accessing services from home computer systems. Although the sample was small (N = 71), 87% of respondents reported having used a personal computer in the past, and 59% currently used a computer. Furthermore, 73% of these computer users indicated that they would be "'generally' or 'very' comfortable with accessing the Internet for health-related reasons" (Ricker et al., 2002, p. 247).

For telerehabilitation to be truly effective, researchers must establish that clinically valid and reliable assessments can be performed from a distance. This requirement is important not only for initial assessment and diagnosis but for ongoing monitoring and maintenance of client progress. Furthermore, the current imperative of evidence-based practice compels researchers to validate telerehabilitation protocols, if this form of service delivery is to become widespread. In the current study, therefore, we explored the effectiveness of delivering a standard assessment for dysarthria over the Internet. In addition we explored the feasibility and usability of computer technology for the online assessment of dysarthria, as a means to inform design and technical questions for future studies.

Method

Participants

Speakers

A total of 19 speakers, ranging in age from 18 to 78 years (79% male) and diagnosed with a perceptible and stable dysarthria associated with an acquired neurological impairment, participated in this pilot study. The diagnosis of dysarthria had been established prior to participation in the study, by a hospital SLP with over 20 years experience with motor speech disorders. Speakers were excluded from the study if they had a history of a speech disturbance prior to the present speech disorder, a severe aphasic language disorder, or a coexisting apraxia of speech. Speakers were also excluded if they exhibited significant aided visual or hearing impairment, post-traumatic amnesia, respiratory dysfunction unrelated to the neurological disorder, a positive history of alcohol abuse, and/or dementia. Table 1 provides biographical and medical information on the speakers in the study. The severity of the dysarthria shown in Table 1 is the rating recorded in the speaker's medical chart. The speakers were not required to have any knowledge or skills in the use of computers, because they were not required to control the computer during the assessment. Speakers gave informed consent before inclusion in the study.

TABLE 1. Clinical details of speakers.

Speaker	Age (years)	Gender	Months postonset/diagnosis	Etiology	Severity of dysarthria per medical chart
1	23	M	5	TBI	Mild
2	21	F	7	TBI	Mild
3	22	F	7	TBI	Mild/moderate
4	36	F	7	HypoxicBI	Moderate
5	27	M	7	TBI	Mild
6	50	M	5	TBI	Moderate
7	69	M	3	CVA	Moderate
8	19	M	6	TBI	Moderate
9	52	M	5	HypoxicBI	Mild/moderate
10	22	M	3	TBI	Moderate/severe
11	71	M	36	PD	Minimal
12	77	M	96	PD	Mild
13	65	M	48	PD	Minimal
14	70	M	120	PD	Mild/moderate
15	59	M	84	PD	Mild
16	56	F	2	Neurosurgery for tumor	Mild/moderate
17	36	M	4	Hypoxic BI	Moderate/severe
18	18	M	5	TBI	Mild/moderate
19	78	M	14	CVA	Mild/moderate

Note. TBI = traumatic brain injury; BI = brain injury; CVA = cardiovascular accident; PD = Parkinson's disease.

Raters

Five SLPs participated in this study. These SLPs were blind to the severity of dysarthria ratings obtained from the speakers' medical charts.

Three SLPs participated in the assessment of the speakers: One was at a distant site to conduct the online assessment (Online SLP 1), the second was at the speaker's site during the online assessment to operate the telecommunication equipment (Online SLP 2), and the third was at the speaker's site to conduct the face-to-face assessments (Face-to-Face SLP). These SLPs were allocated to an assessment environment, such that a different SLP assessed the speaker in each environment. These three SLPs were responsible for providing ratings of the severity of dysarthria and scoring portions of a standardized dysarthria assessment protocol. Two additional SLPs, who were blind to the intent of the study, also participated. These two SLPs provided ratings of speech samples and scored an intelligibility test from recordings made during the assessment.

Procedure

Assessment Environments

A battery of perceptual assessments was administered to each speaker under two conditions: a clinical face-to-face environment and an online environment, with a 2- to 3-day interval between assessment sessions to minimize test–retest effects and fatigue. The assignment of assessment environment (face-to-face followed by online, or online followed by face-to-face) was randomized across speakers.

Face-to-face assessment. In the face-to-face environment, the assessments were conducted in a quiet room and under standard test-administration instructions. The relevant stimulus books were used to conduct the assessments. All audio recordings were made using a headset microphone (AKG Model C 420) and a minidisk recorder (Sony MRX 70). The Face-to-Face SLP used a stopwatch to time various tasks.

Online assessment: system architecture. The system consisted of two computers (see Figure 1). The first computer (Dell Pentium 4) was located in a hospital laboratory approximately 15 km from The University of Queensland. This computer was equipped with custom-built videoconferencing software that allowed the speaker to view Online SLP 1 throughout the assessment session. Two Web cameras (Logitech Pro 4000) were mounted on the computer monitor; one camera was used for the videoconference, whereas the other was used to capture high-quality video that could be delivered to Online SLP 1 by a store-and-forward mechanism. The Web cameras were set at a constant angle and focus. The speaker was positioned at a set distance (50 cm) from the monitor, which allowed Online SLP 1 to view the speaker's head and upper torso. A headset microphone (AKG Model C 420) was worn by the speaker at a distance of 10 cm from the mouth to record speech without obscuring the view of the face. The speaker's computer was responsible for the following:

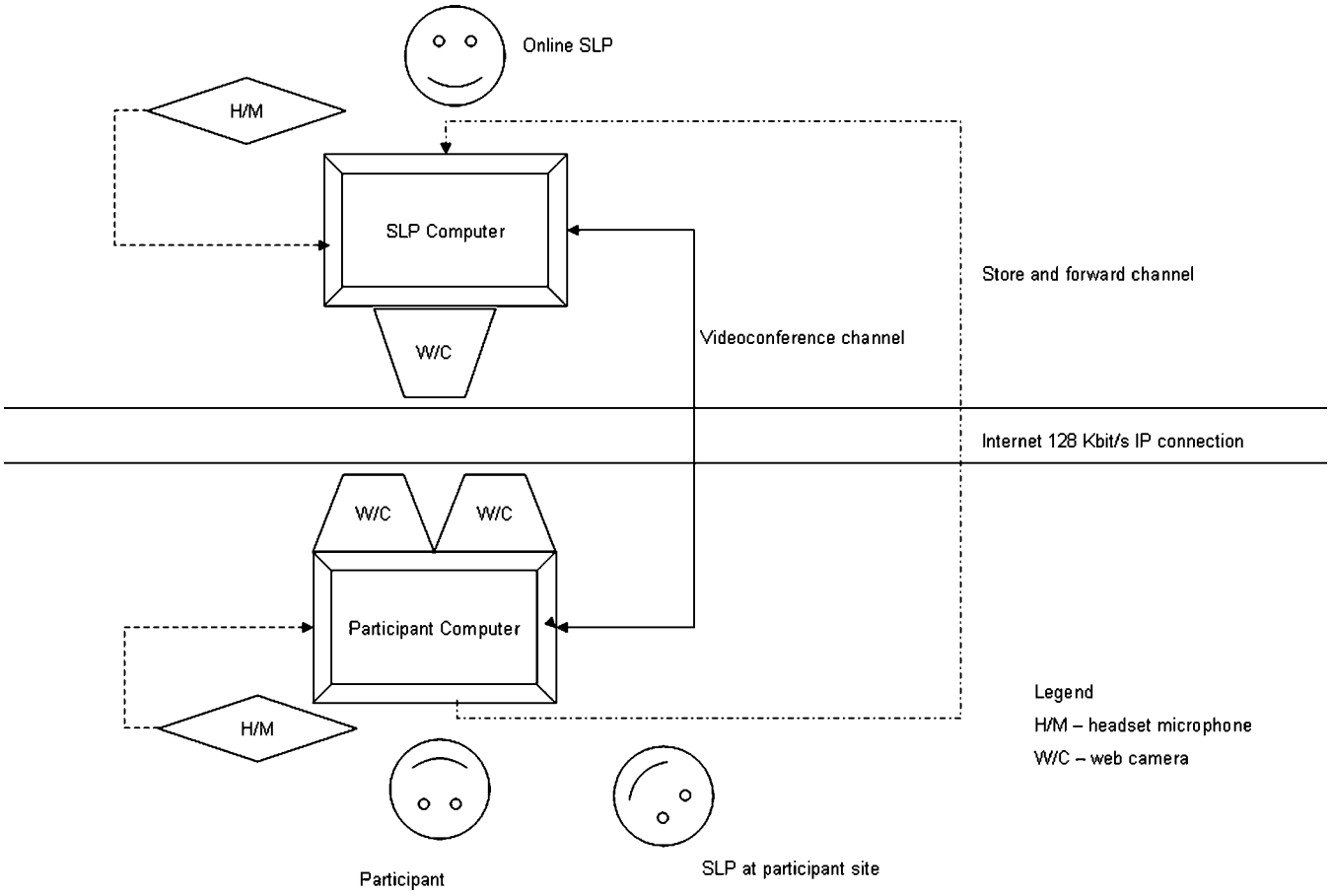
1. Real-time videoconferencing: The computer used custom-built software that enabled real-time videoconferencing between the sites. The software was specifically designed for operation on low bandwidths and Internet-protocol (IP) connections. The Internet bandwidths currently available in Australia in general and, more specifically, in those areas that may benefit most from this service delivery model (i.e., regional and remote areas) are not comparable with those available in most parts of America and Europe. Vast areas of Australia can only access bandwidth of 56 kb/s. Thus, the use of 128 kb/s was considered an acceptable starting point for our research, because it was consistent with the minimum bandwidth connection available in Queensland Health facilities, including those in regional and remote areas. Using this bandwidth would also allow the research to be rolled out to those regional and remote health facilities that may lack sufficient services.
2. Store-and-forward video banking: Using the second camera, the real-time video images were captured at 25 frames per second and then compressed at 235 kb/s using the Microsoft Windows Media Encoder Codec (Version 9). These video files were stored on the speaker's computer, until Online SLP 2 at the speaker's site forwarded them to Online SLP 1.
3. Store-and-forward audio banking: Real-time audio data (WMA, Microsoft) were captured and compressed at 70 kb/s. These audio files were stored and forwarded by Online SLP 2 at the speaker's site to Online SLP 1.

The quality of the compressed video was determined as adequate for the assessment of oromotor activities after a pilot study in which the investigators examined 10 randomized compression rates for the video recording of normal and disordered oromotor performances. One of the research SLPs independently judged the recorded videos at the various compression rates as either acceptable or unacceptable in quality for a reliable evaluation of performance and then rank ordered the compression rates. A compression rate of 235 kb/s was chosen as the highest quality possible without compromising the assessment procedures by extending the time taken to encode and transmit the video files.

The second computer (Dell Pentium 4) was located at the SLP's site at The University of Queensland and was equipped with custom-built videoconferencing software through which Online SLP 1 was able to conduct the online assessments. This computer was responsible for the following:

1. Real-time videoconferencing: The computer used the custom-built software to enable real-time videoconferencing between the sites. The software was specifically designed for operation on low bandwidths and IP connections.
2. Display of text and demonstration videos on the speaker's computer by Online SLP 1 through a chat facility within the software program. A chat facility allows text, audio, and/or video files to be sent and received between users.
3. Playback of the forwarded video and audio files, using the Microsoft WME Codec (Version 9).

FIGURE 1. Diagram of the online assessment environment. Web cameras are mounted above the computer monitor, and headset microphones are worn by both the speaker and the online speech-language pathologist (SLP). The videoconference channel permits communication between the computers, and the store-and-forward channel flows one way from the speaker's computer to the online SLP's computer.



Concerning participant confidentiality, the system had a number of built-in functions to improve data security over the Internet. The security of this software was developed with reference to NetMeeting security guidelines developed by the Systems and Network Attack Centre (SNAC), a division of the U.S. National Security Agency (Hayes & Pitsenbarger, 2003). The functions programmed into this system included the following:

1. The videoconferencing system was designed to allow a maximum of two concurrent users.
2. A secure data channel was incorporated for the transfer of sensitive information. This secure data channel was able to encrypt data, authenticate meeting participants using authentication certificates, and provide password protection.
3. Internet firewalls were installed on both computers to protect data stored from a videoconference, such as video and audio files.

Online procedure. Online SLP 1 established the videoconference link between the two sites. The speaker was seated in front of the computer at a distance of 50 cm from the monitor, allowing for a view of the speaker's head and upper torso. The Web cameras had been prefocused for this distance. Online SLP 1 directed the entire session, giving instructions to the speaker as required for the assessments. This SLP also displayed text or demonstration videos as required on the speaker's computer via the chat facility. Online SLP 1 played back the forwarded audio and video files immediately after they had been sent to check the image quality and the adequacy of the performance by the speaker. Online SLP 1 usually rated the speaker's performance from the video and audio files after the assessment session was complete.

Online SLP 2 at the speaker's site was responsible for setting the encoders on the speaker's computer to either video or audio, when directed by Online SLP 1. Online SLP 2 also initiated and stopped the audio and video recordings, saved the files on the speaker's computer, and then forwarded these files to Online SLP 1.

Perceptual Assessment Battery

Conversational speech. In both environments, the speaker and the assessing SLP engaged in a 2-min dialogue

about a topic relevant to the speaker, such as the etiology of the speaker's communication difficulties, family, hobbies, or interests. This sample was later analyzed by the assessing SLP to determine severity of dysarthria and level of functional communication on a 7-point equal-appearing-interval scale (1 = *normal, no impairment* to 7 = *severe, speech is unintelligible*). At the end of the assessment phase, each of the three assessing SLPs rerated the speakers from the audio files recorded during the online assessment so that reliability statistics could be calculated for the online environment.

Reading aloud. Speakers read aloud a standard passage, "The Grandfather Passage" (Darley et al., 1975). In the online environment, this passage was displayed on the speaker's computer monitor in 36-point, boldface font. Similarly, in the face-to-face format, the passage was presented in 36-point, boldface font to the speaker on an A4-sized sheet of paper. The speaker was instructed to speak in his/her natural manner, using a normal speaking rate and loudness level. After all speakers had been assessed in both environments, all recordings were randomized and transferred to a CD for rating by the two additional SLP raters.

Twenty-six dimensions taken from the perceptual assessment outlined by FitzGerald, Murdoch, and Chenery (1987) were rated for each speaker. Five aspects of speech production were rated: prosody, respiration, phonation, resonance, and articulation. Defined, equal-appearing-interval scale ranges of 4, 5, and 7 points were used in this assessment. A 4-point equal-appearing-interval scale was used to describe the severity of dysfunction of the speech dimension, with 1 representing *normal* and 4 representing *severe impairment,* whereas a 5-point equal-appearing-interval scale was used to rate speech features according to their occurrence, with 1 representing *no occurrence* and 5 representing *frequent occurrence*. A 7-point equal-appearing-interval scale, with the middle rating of 4 representing *normal,* was used to rate those speech dimensions that could be rated on the same scale as being either too high or too low (e.g., pitch level). The two additional SLPs, blinded to the intent of the study, rated the speech samples independently; they were allowed unlimited time to perform the speech ratings. When the ratings between the SLPs differed by more than 1 point on the rating scales, an additional session was conducted to confer and produce a single consensus rating for each dimension. This consensus rating was used in the analysis of the results. In addition, 20% of the recordings were rerated by each of the two SLPs, to allow reliability statistics to be determined.

Assessment of Intelligibility of Dysarthric Speech (ASSIDS). The sentence level of the ASSIDS (Yorkston & Beukelman, 1981) was used to quantify speech intelligibility. Per the test manual, each speaker was required to read or repeat after the examiner 22 randomly selected sentences that varied in length from 5 to 15 words. In the face-to-face environment, the stimulus booklet was used to present the sentences to the speaker, whereas in the online environment, the sentences were displayed on the speaker's monitor in 36-point, boldface font. The speaker's speech was recorded onto a digital minidisk (face-to-face) or as an audio file (online). Samples were then randomly transferred onto CD and transcribed independently by the two additional SLPs. Transcription procedures followed those outlined in the test manual, with the listeners hearing the sentence twice only through headphones. The percentage of intelligible speech was determined at the sentence level. The communication efficiency ratio was determined by dividing the rate of intelligible speech (intelligible words per minute [IWPM]) by the mean rate of intelligible speech produced by a group of normal speakers who spoke at the average rate of 190 IWPM (Yorkston & Beukelman, 1981). For reliability statistics to be determined, 20% of the recordings were transcribed a second time by each of the two SLPs.

Frenchay Dysarthria Assessment (FDA). The FDA (Enderby, 1983) involves the standardized clinical evaluation of the functioning of the lips, tongue, jaw, velopharynx, larynx, respiration, reflexive activity of coughing and swallowing, and overall speech intelligibility. In this study, several of the subtests were excluded from the assessment battery. The Speech Intelligibility subtests were excluded because intelligibility was assessed independently using the ASSIDS. The Reflex subtests of cough, swallow, and dribble and the Palatal subtest of fluid management were also excluded. These subtests relied on the speakers' self-report of their abilities, and, therefore, we did not consider the subtests able to provide data that were suitable for determining the validity of using this assessment tool in the online environment. Two other subtests, Palatal Maintenance and Tongue at Rest, were also not assessed. These subtests were excluded because palatal movement and tongue features could not be rated in the online environment because of an inability to clearly see structures inside the oral cavity using the Web camera. Consequently, a total of 19 subtests on the FDA, including Respiratory, Laryngeal, Lip, Tongue, Jaw, and Palatal Function, were rated in both the face-to-face and online environments.

The speaker's performance was rated according to the FDA's 5-point equal-appearing-interval scale. To enable statistical analysis of the results obtained from the FDA, ratings were coded and a numerical value assigned to each variable assessed, such that "e" (*no function*) was represented by 1, and "a" (*normal function*) was represented by 5. The FDA was administered in accordance with the test manual instructions. Because of the potential for transmission delay and degradation of the visual image in the online environment, specific oromotor tasks requiring a clinical judgment of speed of movement were video recorded by the second Web camera. The FDA tasks that required this real-time video recording were respiration and the Lips and Tongue subtests. Likewise, those tasks that required high-quality audio recordings, including the Laryngeal and Palatal Function subtests, were recorded by the headset microphone, compressed, and then stored and forwarded to Online SLP 1 by Online SLP 2 at the speaker's site. At the conclusion of the data-gathering phase of the study, the online assessment sessions for each speaker were rerated by each of the three assessing SLPs so

that reliability statistics could be determined for the online ratings. The Jaw Function subtest of the FDA was rated live over the videoconference link, and, therefore, reliability analysis was not performed on this subtest.

Statistical Analysis

Two statistical methods were used in this study to compare the results between the assessment methods.

Bland and Altman (BA) technique. Consistent with other studies in telerehabilitation (Brennan et al., 2004; Loh et al., 2004; Russell, Jull, & Wootton, 2003; Russell, Wootton, & Jull, 2002), the first analysis we used was a method-comparison technique described by Bland and Altman (1986, 1999). The BA technique enabled the measurement of the magnitude of the differences between two assessment methods, in this case, face-to-face and online assessment. Bland and Altman suggested that by determining the 95% limits of agreement for the new method one may "provide an interval within which 95% of the differences between measurements by the two methods are expected to lie" (Bland & Altman, 1999, p. 135). This magnitude of the differences should be small enough that it will not affect clinical management. Therefore, clinical criteria that state the permissible level of difference need to be established prior to statistical analysis (see next section). If the limits of agreement between the two methods lie within the predetermined clinical criteria, the new method (online assessment) can be used with confidence interchangeably with the traditional method (face-to-face assessment).

Clinical criteria. The clinical criteria were set before data analysis began and were defined as the magnitude of difference permissible between the two environments such that it would not affect the clinical assessment. According to several authors (Kearns & Simmons, 1988; Kreiman, Gerratt, Kempster, Erman, & Berke, 1993; Sheard, Adams, & Davis, 1991) who have investigated the validity of perceptual analysis, the most common measure of interjudge and intrajudge consistency for rating scales is the percentage of scores that are within ±1 scale point across two rating occasions. Therefore, the clinical criteria for the 5-point and 7-point rating scales used in this study were set at ±1 scale point. The assessments that used these rating scales were the rating of the severity of dysarthria, several of the variables in the perceptual rating of the speech sample, and the FDA. Given the limited range of the 4-point scale as used in the majority of the perceptual ratings, we agreed that a clinical criterion of ±1 scale point was inappropriate and that exact agreement would be required. Therefore, because of the clinical criterion being zero, the BA analysis was inappropriate for the 4-point rating scale.

The clinical criterion for the percentage sentence intelligibility on the ASSIDS was set with respect to the test–retest speaker variability as reported in the test manual. The variability that was seen in people with dysarthria who were assessed face-to-face on the same day was found to be approximately 8.6% change in sentence intelligibility levels. Therefore, the clinical criterion was set within ±8.6% change in sentence intelligibility. If the differences between the two methods of assessment were comparable with the test–retest variability seen in speakers assessed face-to-face on the same day, the online method would be considered as accurate as the traditional face-to-face method. For the communication efficiency ratio, the clinical criterion was calculated using previous research data, in which the standard deviation of the mean difference between ratings for a group of neurologically impaired dysarthric speakers was 0.14 (Farrell, Theodoros, Ward, Hall, & Silburn, 2005). By applying a 95% confidence interval to this standard deviation, a clinical criterion of −0.27 for the communication efficiency ratio was established.

Percentage level of agreement. In addition to the BA method, a second level analysis was performed to compare the results with other studies. Percentage level of agreement has been used in a number of studies as an acceptable measure of consistency in the perceptual rating of speech dimensions (Kearns & Simmons, 1988; Kreiman et al., 1993; Sheard et al., 1991). These authors considered high agreement to be at a level equal to or greater than 80% agreement within ±1 scale point (percentage close agreement). Therefore, the clinical criteria for percentage level of agreement measures were set at equal to or greater than 80% close agreement for the 5-point and 7-point rating scales. Because of their limited scale, however, the 4-point rating scales required a clinical criterion of equal to or greater than 80% exact agreement. For the ASSIDS, the clinical criteria set for the second level of analysis were that at least 80% of the comparisons between the methods were within ±8.6% for the percentage sentence intelligibility and ±0.27 for the communication efficiency ratio.

Results

Reliability Within Methods

It was important to establish the degree of reliability in each method, because poor reliability within either method would have had a significant impact on the degree of difference between the assessment methods. Consequently, interjudge and intrajudge reliabilities for the nonstandardized assessment tools in each environment were determined using a Pearson's product–moment r correlation for interval data and a Spearman's ρ for ordinal data.

Results for the severity rating scale in the online environment revealed moderate to high interjudge reliability, whereas intrajudge reliability was determined as high for two of the SLPs and moderate to high for the other SLP (see Table 2). Reliability correlations were not calculated for the severity rating scale in the face-to-face environment because only a single live rating was recorded in this environment.

Interjudge reliability for the perceptual analysis of the speech sample was determined using Spearman's ρ correlations. Correlations indicated a moderate degree of interjudge reliability within the face-to-face and online environments for this assessment. Intrajudge reliability within each environment was found to be high (see Table 2).

Concerning the ASSIDS, moderate interjudge reliability was revealed for the face-to-face environment, whereas a

TABLE 2. Intrajudge and interjudge reliability statistics for assessments in both environments.

Assessment	Intrajudge reliability face-to-face	Intrajudge reliability online	Interjudge reliability face-to-face	Interjudge reliability online
Severity rating[a]	—	0.88**	—	0.57*–0.85**
	—	0.76**		
	—	0.97**		
Speech sample[b]	0.99**	0.88**	0.56**	0.62**
	0.90**	0.92**		
FDA[a]	—	0.85**	—	0.72**–0.77**
	—	0.90**		
	—	0.87**		
ASSIDS[b]	0.99**	0.97*	0.74**	0.90*
	0.99**	0.99*		

Note. Dashes indicate that data were not obtained; FDA = Frenchay Dysarthria Assessment (Enderby, 1983); ASSIDS = Assessment of Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1981). Statistic for severity rating, speech sample, and FDA was ρ; statistic for ASSIDS was r.

[a]Reliability calculated for three judges.
[b]Reliability calculated for two judges.
*$p < .05$, two-tailed.
**$p < .01$, two-tailed.

high interjudge reliability was found for the online environment. Pearson's r coefficient revealed high intrajudge reliability for each of the judges in both environments (see Table 2).

Interjudge reliability for the FDA as determined by Spearman's ρ correlation revealed moderate reliability among the three judges, with ρs = .72, .72, and .77, $p < .01$. The interjudge reliability as reported in the FDA (Enderby, 1983) test manual ranged from .79 to .92. Spearman's ρ correlation revealed moderate to high intrajudge reliability (ρs = .85, .87, .90). Reliability correlations were not calculated for the FDA in the face-to-face environment because only a single live rating was recorded in this environment. Table 2 outlines the results of reliability analyses for each assessment.

Differences Between Methods

Severity Rating of Conversational Speech Intelligibility

A normal distribution of difference scores was found for this rating scale. Using the BA method, the limits of agreement were calculated at ±0.9 of a scale point. Thus, the limits of agreement fell within the clinical criterion for this rating scale. The percentage level of agreement at percentage close agreement was 100% and, thus, within the clinical criterion.

Speech Sample Analysis

Tests of normality revealed normal distributions of difference scores for each of the 4-, 5- and 7-point rating scales. Four variables were rated on a 7-point equal-appearing-interval scale. The limits of agreement were within the clinical criterion for one of the variables (general rate), whereas the other three variables (pitch level, loudness level, and maintenance of rate) fell outside the criteria. However, all of the variables met the clinical criteria for percentage level of agreement analysis (see Table 3).

Five variables were rated on a 5-point equal-appearing-interval scale. For two of these variables (rate fluctuation and short rushes of speech), the limits of agreement using the BA method were determined to be within the clinical criteria, whereas the other three variables (pitch breaks, excessive loudness variation, and prolonged intervals) were outside the criteria. All of these variables met the clinical criteria for the second level of analysis, with all percentage level of agreement being above 80% close agreement (see Table 3).

Seventeen variables, representing all levels of the speech production process, were rated on a 4-point equal-appearing-interval scale. Six variables met the clinical criterion of equal to or above 80% exact agreement for the percentage level of agreement analysis, whereas 8 were above 70% exact agreement and the remaining 3 above 63% exact agreement (see Table 3).

ASSIDS

The difference scores were normally distributed for both the percentage sentence intelligibility and communication efficiency ratio. Using the BA method, the limits of agreement for the percentage sentence intelligibility were determined at ±8.84% and, thus, just outside the clinical criterion of ±8.6%. The limits of agreement for the communication efficiency ratio were found to be ±0.22 and, therefore, within the clinical criterion. The percentage levels of agreement for both the percentage sentence intelligibility and communication efficiency ratio were above 80%, at 83.33% and 100%, respectively.

FDA

We found normal distributions for tests of normality on the difference scores for each subtest within the FDA. The BA limits of agreement were outside the predetermined clinical criteria for all except three of the subtests. However, the percentage level of agreement results were within the criteria of equal to or above 80% close agreement, on all except four of the subtests. Table 4 outlines each subtest within the FDA, their limits of agreement, and the percentage level of agreement.

TABLE 3. Limits of agreement and percentage level of agreement for 26 perceptual speech dimensions measured from speech sample.

Perceptual speech dimension	Bland and Altman limits of agreement	% level of agreement
7-point scales		
Pitch level	±1.80	89.47[b]
Loudness level	±2.09	84.21[b]
General rate	±0.92[a]	100[b]
Maintenance of rate	±1.11	94.74[b]
5-point scales		
Pitch breaks	±1.22	100[b]
Excessive loudness variation	±1.10	94.74[b]
Rate fluctuations	±0.73[a]	100[b]
Prolonged intervals	±1.24	100[b]
Short rushes of speech	±0.62[a]	100[b]
4-point scales		
Variation of pitch	—	94.74[c]
Steadiness of pitch	—	89.47[c]
Variation of loudness	—	89.47[c]
Maintenance of loudness	—	78.95
Phrasing	—	73.68
Stress pattern	—	73.68
Breath support for speech	—	78.95
Hypernasality	—	78.95
Harshness	—	63.16
Strained–strangled	—	73.68
Intermittent breathiness	—	68.42
Hoarseness	—	63.16
Glottal fry	—	84.21[c]
Wetness	—	94.74[c]
Precision of consonants	—	89.47[c]
Length of phonemes	—	78.95
Overall intelligibility	—	78.95

Note. Dashes indicate that Bland and Altman's methodology (1986, 1999) was not applied to 4-point scales because of limited scale size.

[a]Within clinical criteria of ±1 scale point using Bland and Altman's methodology.
[b]Within clinical criteria of greater than or equal to 80% agreement at ±1 scale point.
[c]Within clinical criteria of greater than or equal to 80% exact agreement.

Discussion

The results of this study indicate that the majority of the perceptual assessment battery can be reliably administered over the Internet using the computer-based telerehabilitation system. Although the online assessment of motor speech disorders has previously been successful using specialized equipment and satellite connectivity (Duffy et al., 1997), the results from this study indicate that readily available equipment, such as computers, and easily accessible communication connections, such as the Internet, may be used to conduct reliable online assessments of motor speech disorders. The findings support the further development of computer-based telerehabilitation protocols for the assessment of communication disorders.

The magnitudes of the observed limits-of-agreement scores in the present study do not indicate a clinically significant difference between online and face-to-face environments for several of the assessments. In particular, the limits-of-agreement scores for the severity rating scale and the communication efficiency ratio from the ASSIDS fell within the clinical criteria for the BA analysis. This is not surprising because of the system's ability to record high-quality audio files at the speaker's computer and then store and forward such files to Online SLP 1. This feature of the system allowed the audio files recorded online to be comparable with those recorded in the face-to-face environment, thus reducing disparity between the assessment environments.

As reported in the results, the limits of agreement for the percentage sentence intelligibility (±8.84%) on the ASSIDS were marginally outside the clinical criteria of ±8.6%. Yorkston and Beukelman (1981) cautioned researchers to be conservative in their interpretation of intelligibility scores obtained from the ASSIDS, because of the considerable day-to-day variation in some speakers' performance. The 0.24% difference between the clinical criteria and the limits of agreement is, therefore, not considered to be clinically significant. Thus, the results of the current study support that a measure of sentence intelligibility comparable with face-to-face assessment can be reliably achieved online.

With regard to the perceptual rating of the speech sample, 6 of the 9 variables rated on 5-point and 7-point scales did not fall within the clinical criteria using the BA methodology. However, all 9 of these variables met the percentage level of agreement clinical criteria of equal to or greater than 80% close agreement. Of the 17 variables rated on the 4-point rating scale, 6 met the percentage level of agreement clinical criteria of equal to or greater than 80% exact agreement, 5 were just below 80% exact agreement, 3 were above 70% exact agreement, and the remaining 3 were above 63% exact agreement (see Table 3). Although chance-level agreement on a 4-point scale is 60%, only 3

TABLE 4. Limits of agreement and percentage level of agreement for 19 variables on the FDA.

FDA variable	Limits of agreement	% level of agreement
Respiration at rest	±1.30	100[b]
Respiration in speech	±0.89[a]	100[b]
Lips at rest	±1.94	84.2[b]
Lips spread	±1.18	100[b]
Lips seal	±1.68	89.6[b]
Lips alternate	±1.26	100[b]
Lips in speech	±1.19	89.5[b]
Jaw at rest	±0.79[a]	100[b]
Jaw in speech	0.00[a]	100[b]
Palate in speech	±1.99	79.0
Laryngeal time	±1.81	89.6[b]
Laryngeal pitch	±1.77	84.2[b]
Laryngeal volume	±1.89	73.7%
Laryngeal in speech	±1.74	89.5[b]
Tongue protrusion	±1.76	84.3[b]
Tongue elevation	±2.43	68.5
Tongue lateral	±2.37	63.2
Tongue alternate	±1.64	84.2[b]
Tongue in speech	±1.05	100[b]

[a]Within clinical criteria of ±1 scale point.
[b]Within clinical criteria of greater than or equal to 80% agreement at ±1 scale point.

variables approached this lower level of agreement. These findings indicate that perceptual ratings for the majority of speech characteristics defined in this assessment tool could be reliably achieved in the online environment. Given the similarity in the procedures used on these assessments, possible explanations for why 11 variables did not meet the clinical criteria for the percentage level of agreement analysis may include issues of speaker and judge variabilities. Speaker variability must be acknowledged as a possible factor in lower levels of agreement when ratings are completed on different occasions, as occurred in this study. According to Ludlow and Bassich (1983), the considerable day-to-day variation in the performance of some people with speech disorders impedes the attainment of acceptable agreement between two cases of assessment.

Regarding judge variability, studies have confirmed that judges more frequently agree about normal or severe pathology, rather than mild-to-moderate levels of severity (Kreiman et al., 1993). As the majority of our speakers were classified in their medical charts as having mild-to-moderate dysarthria, this factor may have contributed to less than optimal perceptual rating agreement for the nine variables. Furthermore, some studies focusing on the perceptual rating of voice characteristics have revealed that vocal qualities such as harshness and hoarseness are particularly difficult for judges to rate accurately and are associated with low reliability (Ludlow & Bassich, 1983; Sheard et al., 1991). These were two of the variables that did not meet the percentage level of agreement clinical criteria. In addition, intercorrelation between speech dimensions may account for variability in the ratings. Intercorrelation occurs when judges report interference in their perception of the speech dimension in question, by the presence of other speech characteristics (Sheard et al., 1991).

Although most subtests of the FDA failed to meet the clinical criteria for the BA method comparison, the majority did meet the clinical criteria for the percentage level of agreement analysis. The four subtests that did not meet the second level of analysis were Palate in Speech, Laryngeal Volume, Tongue Elevation, and Tongue Lateral. The first subtest, Palate in Speech, perceptually assesses the degree of nasality present during speech, which is a notoriously problematic judgment even for experts (Chenery, 1998; Karnell, 1995; McHenry, 1999). Researchers have suggested that the perceptual judgment of nasality is often affected by other speech dimensions, whereas other researchers have gone farther to propose that nasality should only be judged in conditions that preclude intercorrelation between speech dimensions (Chenery, 1998; Theodoros, Murdoch, & Stokes, 1995; Theodoros, Murdoch, Stokes, & Chenery, 1993). This issue of intercorrelation between speech dimensions has been reported extensively in relation to motor speech disorders (Chenery, 1998; Kreiman et al., 1993; Sheard et al., 1991) and may have also influenced the judgments made in the other subtests that did not meet the clinical criteria. As previously discussed, intraspeaker variability in the current study was highly likely to have occurred and may have had a significant impact on the variability found in the FDA results and, thus, the degree of reliability attained.

Another factor that may have influenced the judgments made on these subtests is that of poor dimension definition. Tongue Elevation and Tongue Lateral subtests contained rating definitions that required two elements to be judged (e.g., time and range of movement or time and coordination of movements). The need to produce a single rating on two or more elements may have led to different judges emphasizing one element over the other and so producing different ratings even on the same performance. Feedback from the judges revealed that rating definitions also presented interpretation problems on the following subtests of the FDA: Lip Seal, Lip Alternate, Laryngeal Time, Tongue Protrusion, Tongue Elevation, Tongue Lateral, and Tongue Alternate. Other studies (Sheard et al., 1991) involving the perceptual rating of dysarthria have reported similar problems with the interpretation of rating scale descriptors. Of course, evaluation of the reliability within the face-to-face environment in the current study would have allowed for a more thorough examination of these issues.

Other possible factors limiting the FDA in an online environment relate to technical issues. The technical issues were particularly relevant to the ratings completed from video files. Camera position and focus were technical issues. The cameras used were fixed and could only be focused manually at the speaker's site. Although focus was constant for each assessment—having been preset for speaker positioning at a distance of 50 cm from the monitor—if the speaker moved from his or her initial position, the remote SLP was not able to adjust the camera position or focus. The speaker was required to return to his or her initial position in front of the camera. Speaker positioning was particularly significant when the speaker demonstrated head or body dyskinesis, as occurred in some speakers with Parkinson's disease. Another technical issue influencing the assessment was the lighting within the online environment. No additional lighting beyond standard overhead room lighting was used during the online assessment. We believe that lighting directed onto the speaker's face would have allowed for clearer vision of the face and oromotor movements. In addition, the use of a plain, contrasting backdrop would have enhanced the assessment of oromotor function. The issue of Internet speed is also relevant. Because the results from the current study are based on an Internet speed of 128 kb/s, we predict that increased Internet speeds or the use of broadband technology will lead to greater levels of agreement between assessment methods.

It is important to note that some subtests were excluded (e.g., Swallow) and others were not conducted because of inability to view structures inside the oral cavity in the online environment (e.g., Soft Palate). However, we hope that with improved camera versatility, such as improved focus and zoom capabilities, these subtests may be included in future protocols. Therefore, although the use of aspects of the FDA in an online environment can be both feasible and generally reliable, SLPs should conduct these assessments with caution and a clear understanding of the limitations in an online environment.

With the design and evaluation of protocols for the online assessment of dysarthria still in their infancy,

acknowledgment and discussion of limitations in this area are imperative. The scope and design of this pilot study precluded the investigation of some aspects of the use of telerehabilitation in speech-language pathology. These excluded items encompassed areas of clinical management such as the diagnosis of the type of dysarthria, evaluation of speaker satisfaction, and analysis of cost effectiveness of assessment by telerehabilitation methods. We acknowledge that diagnosis of the type of dysarthria is necessary to complete the clinical assessment process and allow treatment recommendations to be made. Furthermore, a diagnosis of type of dysarthria allows clinicians at different facilities to communicate more effectively on the subject of clinical management. Therefore, the diagnosis of type of dysarthria should be included in future studies so that the true clinical value of telerehabilitation methods may be fully appraised. The analysis of cost effectiveness in online assessment was considered outside the scope of this pilot study because the speakers were hospital inpatients; however, this topic should be addressed in larger experimental studies so that institutions have evidence to support advancement into the establishment of telerehabilitation speech-language pathology services. Likewise, the analysis of participant satisfaction—that is, of both speakers and raters—should be included in future research because satisfaction in a service delivery model is vital to successful clinical implementation.

This research project was the pilot phase for a larger study into the use of telerehabilitation applications for the assessment of motor speech disorders. The results from this pilot study and the limitations demonstrated will be used to improve the protocols for telerehabilitation assessment of motor speech disorders. Some of the design and technical issues designated for improvement as a result of this study include the automation of the store-and-forward mechanism between the sites and the use of tracker pods to allow the remote positioning of the Web cameras at the speaker's site by Online SLP 1. In addition, to more accurately compare reliability within and between the methods, two SLPs will rate motor speech function in the face-to-face environment.

Clinicians should be optimistic that as research continues to refine and evaluate assessment protocols for the use of computer-based telerehabilitation systems in speech-language pathology, solid evidence will emerge to support the reliable use of standard assessment procedures. It is only with this evidence that telerehabilitation will become an alternative service delivery model for speech-language pathology services.

Acknowledgment

Funding from The University of Queensland's Faculty of Health Sciences Small Grant Program (Grant B/502/SPA/01/UQSGS) supported this study.

References

Australian Bureau of Statistics. (2004). *Measures of a knowledge-based economy and society, Australia. Information and communications technology indicators.* Retrieved October 8, 2004, from http://www.abs.gov.au/Ausstats/abs@nsf/94713ad445ff1425ca25682000192af2/7599f94ffdbadccbca256d97002c8636!OpenDocument

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 8,* 307–310.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8,* 135–160.

Brennan, D. M., Georgeadis, A. C., Baron, C. R., & Barker, L. M. (2004). The effect of videoconference-based telerehabilitation on story retelling performance by brain-injured subjects and its implications for remote speech-language therapy. *Telemedicine Journal and e-Health, 10*(2), 147–154.

Chenery, H. (1998). Perceptual analysis of dysarthric speech. In B. E. Murdoch (Ed.), *Dysarthria: A physiological approach to assessment and treatment* (pp. 36–67). Cheltenham, United Kingdom: Stanley Thornes.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders.* Philadelphia: W. B. Saunders.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis and management* (2nd ed.). St. Louis, MO: Elsevier Mosby.

Duffy, J. R., Werven, G. W., & Aronson, A. E. (1997). Telemedicine and diagnosis of speech and language disorders. *Mayo Clinic Proceedings, 72,* 1116–1122.

Enderby, P. (1983). *Frenchay Dysarthria Assessment*. San Diego, CA: College-Hill Press.

Farrell, A., Theodoros, D., Ward, E., Hall, B., & Silburn, P. (2005). The effects of neurosurgical management of Parkinson's disease on speech characteristics and oromotor function. *Journal of Speech, Language, and Hearing Research, 48,* 5–20.

FitzGerald, F., Murdoch, B., & Chenery, H. (1987). Multiple sclerosis: Associated speech and language disorders. *Australian Journal of Human Communication Disorders, 15*(2), 15–33.

Hayes, J. M., & Pitsenbarger, T. H. (2003). *Microsoft NetMeeting 3.0 security assessment and configuration guide*. Retrieved July 11, 2003, from http://nsa2.www.conxion.com/support/guides/sd-6.pdf.

Hill, A., & Theodoros, D. (2002). Research into telehealth applications in speech-language pathology. *Journal of Telemedicine and Telecare, 8,* 187–196.

Karnell, M. P. (1995). Nasometric discrimination of hypernasality and turbulent nasal airflow. *The Cleft Palate-Craniofacial Journal, 32,* 145–148.

Kearns, K. P., & Simmons, N. N. (1988). Interobserver reliability and perceptual ratings: More than meets the ear. *Journal of Speech and Hearing Research, 31,* 131–136.

Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research, 36,* 21–40.

Lloyd, R., & Bill, A. (2004). *Australia online: How Australians are using computers and the Internet*. Canberra: Australian Bureau of Statistics, Australian Census Analytic Program.

Loh, P. K., Ramesh, P., Maher, S., Saligari, J., Flicker, L., & Goldswain, P. (2004). Can patients with dementia be assessed at a distance? The use of Telehealth and standardised assessments. *Internal Medicine Journal, 34,* 239–242.

Ludlow, C. L., & Bassich, C. J. (1983). Relationships between perceptual ratings and acoustic measures of hypokinetic speech. In M. R. McNeil, J. C. Rosenbek, & A. E. Aronson (Eds.), *The dysarthrias: Physiology, acoustics, perception, management* (pp. 163–195). San Diego, CA: College-Hill Press.

Mashima, P. A., Birkmire Peters, D. P., Syms, M. J., Holtel, M. R., Burgess, L. P., & Peters, L. J. (2003). Telehealth: Voice therapy using telecommunications technology. *American Journal of Speech-Language Pathology, 12,* 432–439.

McHenry, M. A. (1999). Aerodynamic, acoustic, and perceptual measures of nasality following traumatic brain injury. *Brain Injury, 13,* 281–290.

Ricker, J. H., Rosenthal, M., Garay, E., DeLuca, J., Germain, A., Abrahams-Fuchs, K., et al. (2002). Telerehabilitation needs: A survey of persons with acquired brain injury. *Journal of Head Trauma Rehabilitation, 17,* 242–250.

Rosen, M. J. (2004). Telerehabilitation. *Telemedicine Journal and e-Health, 10,* 115–117.

Russell, T., Jull, G., & Wootton, R. (2003). Can the Internet be used as a medium to evaluate knee angle? *Manual Therapy, 8,* 242–246.

Russell, T., Wootton, R., & Jull, G. (2002). Physical outcome measurements via the Internet: Reliability at two Internet speeds. *Journal of Telemedicine and Telecare, 8*(Suppl. 3), 50–52.

Sheard, C., Adams, R. D., & Davis, P. J. (1991). Reliability and agreement of ratings of ataxic dysarthric speech samples with varying intelligibility. *Journal of Speech and Hearing Research, 34,* 285–293.

Theodoros, D. G., Murdoch, B. E., & Stokes, P. D. (1995). Variability in the perceptual and physiological features of dysarthria following severe closed head injury: An examination of five cases. *Brain Injury, 9,* 671–696.

Theodoros, D. G., Murdoch, B. E., Stokes, P. D., & Chenery, H. J. (1993). Hypernasality in dysarthric speakers following severe closed head injury: A perceptual and instrumental analysis. *Brain Injury, 7,* 59–69.

Torsney, K. (2003). Advantages and disadvantages of telerehabilitation for persons with neurological disabilities. *Neuro-Rehabilitation, 18,* 183–185.

Wertz, R. T., Dronkers, N. F., Berstein-Ellis, E., Shubitowski, R. E., Shenaut, G. K., Knight, R. T. (1987). Appraisal and diagnosis of neurogenic communication disorders in remote settings. *Clinical Aphasiology, 17,* 117–123.

Wertz, R. T., Dronkers, N. F., Berstein-Ellis, E., Sterling, L. K., Shubitowski, R. E., & Shenaut, G. K., et al. (1992). Potential of telephonic and television technology for appraising and diagnosing neurogenic communication disorders in remote settings. *Aphasiology, 6,* 195–202.

Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of Intelligibility of Dysarthric Speech*. Austin, TX: Pro-Ed.

Yorkston, K. M., Beukelman, D. R., & Bell, K. R. (1988). *Clinical management of dysarthric speakers*. Austin, TX: Pro-Ed.

Received April 4, 2005
Revision received August 18, 2005
Accepted October 20, 2005
DOI: 10.1044/1058-0360(2006/006)

Contact author: Anne J. Hill, Division of Speech Pathology, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia.
E-mail: a.hill@shrs.uq.edu.au