

Protocol for Assessment of Speech Intelligibility and Speech Naturalness in Persons with Dysarthria in Kannada

Orin D'Silva & R Manjula*

Abstract

The purpose of this study was to develop a protocol in Kannada for assessing speech intelligibility and speech naturalness in individuals with dysarthria. The protocol thus developed was administered on 5 individuals with dysarthria within the age group of 25-70 years who presented with four different conditions of dysarthria: spastic, flaccid, hypokinetic and hyperkinetic. Speech samples collected from these clients was subjected for inter-judge and intra-judge reliability. Alpha reliability measure was employed and it was found that there was good inter-judge and intra-judge reliability indicating that the protocol developed is sensitive in assessing speech intelligibility and speech naturalness of individuals with dysarthria

Introduction

Communication is a two-way process by which information is passed from one person to another. Three criteria are met in communication: first, there is some output from the initiating person; second, there is some means of transmitting and conveying the information; and third, the recipient can receive and interpret the information. Effective communication depends on how intelligible the speaker is, how well the speaker's speech meets the cultural standards, how much the listener perceives or understands of what the speaker conveys. Hence nothing is more useful than to speak clearly i.e., speech should be intelligible. Speech intelligibility is defined as the degree to which the listener (Yorkston, Beukelman & Bell, 1986) understands the acoustic signal. It is simply and broadly defined as the understandability of speech. Implicit in the definition is a task in which a speaker produces a message and a listener who does not know the content of the message attempts to comprehend and/or reproduce it. Articulation, rate, fluency, vocal quality and intensity mainly influence this measure.

Speech is a very rapid and complex motor act which requires very finely tuned neurological regulation (Kent & Forner, 1980; Netsell, 1984). For speech motor act to take place, good co-ordination in muscle strength, speed of movement, appropriate range of excursion, accuracy of movement, motor steadiness and muscle tone is required in the speech organs. Damage to the nervous system impairs one or more of these neuromuscular functions and this may affect speech motor production. Dysarthria results from a disruption of muscular control due to lesions either in the central or peripheral nervous systems. Various clinical tools and tasks are employed by speech-language clinicians to assess the intelligibility of speech in individuals with dysarthria. These are often used in the assessment of speech impairment before and after the

* Professor of Speech Pathology, All India Institute of Speech and Hearing, Mysore, India
e-mail: r_manjula@yahoo.com

rehabilitation procedures to maximize functional communication competence in this population (Coombes, 1986).

Also, intelligibility of speech is defined as recognition of words or utterances in natural communication situations (Smith & Nelson, 1985). This has been used as an important index or measure to assess the overall speech adequacy of individuals with speech-language disabilities by various researchers and clinicians (Beukelman & Yorkston, 1980; Yorkston & Beukelman, 1980; Weiss, Gordon & Lillywhite, 1987; Bernthal & Bankson, 1998). The most common means of measuring speech intelligibility is to have a naïve listener orthographically transcribe the recorded speech sample of an individual with dysarthria. The number of words correctly identified by the listener divided by the total number of words in a sample provides a percentage intelligibility score. Several factors influence the measurement of intelligibility: the speaking task, the type of analysis carried out by the judges, the training of the judges and the medium used for judging the signal (Yorkston, Beukelman & Bell, 1988). Although there is no 'ideal' technique proposed for the assessment of speech intelligibility, several techniques are advocated and employed for quantifying speech intelligibility. In spite of the drawbacks each of these proposed scales/protocols/tests help as a handy tool in the treatment planning and assessment of progress in individuals with dysarthria. In a practical clinical situation it helps in deciding whether or not a program of remediation is required and in monitoring progress.

The common procedures adopted to assess the speech intelligibility in individuals with dysarthria include:

- Scaling procedure: wherein listener uses a rating scale and assigns ratings of overall speech intelligibility (Darley, 1969; Platt, Andrews, young & Nelson, 1978; Yorkston & Beukelman, 1978; Platt, Andrews, Young & Quinn, 1980)
- Identification task wherein the listener transcribes what the speaker says. Intelligibility of single words is measured by computing the percentage of correctly identified words (Tikofsky & Tikofsky, 1964; Platt, Andrews, young & Nelson, 1978; Yorkston & Beukelman, 1978)
- Acoustic analysis using instruments (analysis of acoustic waveforms - Kent & Netsell, 1975).

There are a few standardized scales, tests and protocols developed to assess speech intelligibility in the Western countries. For example, "Assessment of intelligibility of dysarthric speech (AIDS)" given by Yorkston and Beukelman, (1981); Frenchay dysarthria assessment given by Enderby (1983). In the evaluation of communicative impairments in speech and language disabilities both subjective (perceptual), objective and acoustic analysis are often used. In the earlier days perceptual analysis was often used to test for phoneme intelligibility, word intelligibility and sentence intelligibility. The two most commonly used perceptual intelligibility measures are:

1. *Interval scaling (IS)*: Listener assigns to each stimulus a number that represents a linear partition of a scale (Darley, Aronson & Brown 1969).
2. *Direct magnitude estimation (DME)*: Listener assigns to each stimulus a number representing the ratio of the stimulus to a standard that is either specified by the examiner or selected by the listener.

Several techniques have been traditionally used for the quantification of speech intelligibility and comparisons have been made. They appear to be sensitive and reliable measures of dysarthric speech intelligibility. However no single technique appears to be applicable across different types of dysarthria. There is no tool or a protocol developed for the Indian population to rate the speech intelligibility in adult individuals with dysarthria. Study by Mary (1993) compared Tamil speaking children with spastic and athetoid cerebral palsy. No objective scores were used to delineate the intelligibility of speech.

Method

Participants: 5 clients in the age range of 25-70 years with confirmed diagnosis of dysarthria (by a neurologist & speech language pathologist) due to acquired condition whose native language was Kannada were included. The subjects selected did not have any hearing, visual or cognitive deficits as revealed by screening tests. All the subjects could express in sentences and could read Kannada text material.

Material: Protocol consisted of:

- i A word list consisting of 2 sets with 30 Kannada words in each set - Set I consisted of simple words without geminates and morpho-phonemic clusters. Set II consisted of complex words with geminates and morpho-phonemic clusters which added to the complexity of the word structure.
- ii A list consisting of 20 simple declarative sentences - The sentences were formed such that the level varied in terms of length and complexity. Using words with geminates and morphophonemic clusters and complex sentences were formed.
- iii A figure consisting of a 'market' scene which was depicted through line drawing was selected

Procedure

1. *Assessment of speech intelligibility:* Subjects were individually tested. The protocol consisted of 3 tasks, i.e., word intelligibility task, sentence task and narration task. Each target word and sentence for assessing word and sentence intelligibility task were written on flash cards and the subjects had to read them aloud. These responses were audio recorded. For narration the picture was shown to the subjects and asked to describe the activities in the picture. The speech sample obtained from the subjects was scored for percentage correct responses in terms of intelligibility of speech and speech naturalness by 3 literate judges who were speech-language pathologists having experience in handling individuals with dysarthria for a minimum of 1 year. They were kept blind to the purpose of the study and their mother tongue was Kannada.
2. *Word intelligibility task:* The words spoken by the 5 subjects were randomized within set 1 and 2 and a master tape was prepared for set 1 and 2. This was then subjected to perceptual judgment. The 3 judges carried out the task independently and without mutual consultation. The 3 judges were instructed as follows "You will hear a series of words which will be presented one by one with a time gap of 10 seconds in between. You have to listen to each word carefully and write the word as you hear them. You may rewind the recorder and listen

to the word again if you want to". Word intelligibility scores were calculated as the total number of intelligible words/correct words transcribed by the judge divided by the total number of words uttered by the subjects and multiplied by 100. This yielded the percentage score of speech intelligibility on word task.

$$\text{Word intelligibility score} = \frac{\text{No. of correct words transcribed by the listener}}{\text{Total no. of words in the sample}} \times 100$$

3. *Sentence intelligibility task*: The sentences spoken by the 5 subjects were also randomized across subjects and within the two sets respectively and a master tape for 2 sets was prepared which was subjected for perceptual judgement. The 3 judges carried out the task independently and without mutual consultation. The judges were instructed to listen to each sentence and do a word-by-word transcription of sentences. From the transcribed data, the percentage error was calculated i.e., number of words correctly identified in the sentences by the judges from the subjects' utterance, divided by the total number of words present in the sentences uttered by the subject and multiplied by 100.

$$\text{Sentence intelligibility score} = \frac{\text{No. of intelligible words identified in the sentences uttered by subjects}}{\text{Total number of words in the sentences}} \times 100$$

This calculation was done for each sentence in set 1 and set 2 respectively.

4. *Picture narration task*:

- a) The verbal description of the picture by each subject in the narration task was listened to by the judges and transcribed. The total number and percentage of intelligible words in the narration was calculated. The 3 judges performed the task independently without mutual consultation. The percentage correct identification was calculated for the picture narration as:

$$\text{Narration intelligibility score} = \frac{\text{No. of intelligible words identified in the narrated sample}}{\text{Total number of words in the narrated sample}} \times 100$$

- b) In order to check for correlation of the percentage scores offered by the judges in the word sentence and narration intelligibility task with the overall perceived intelligibility of the subjects' speech, three additional judges who did not participate in the assessment of speech intelligibility tasks were chosen. They were speech-language pathologists with a minimum of one year clinical experience with dysarthric clients. They were kept blind to the purpose of the study. These judges were asked to listen to a 30 second sample of the narrated speech sample of the five subjects and rate the narrated speech of each of the subject on a 4 -point rating scale with 0 indicating 'normal speech' and 3 indicating 'Severely unintelligible speech':

0 = Normal speech

1 = Mild unintelligible speech

2 = Moderate unintelligible speech

3 = Severe unintelligible speech

5. *Assessment of Speech naturalness*: The narration task selected in the protocol served as stimuli for assessment of speech naturalness also. The speech samples collected from subjects for assessment of speech intelligibility was judged on a 2 point rating scale by the same 3 qualified judges who assessed the speech intelligibility in words, sentences and narration sample.

- *Use of stress*: appropriate stress = 0; reduced stress/excess stress = 1
- *Use of intonation* : normal intonation = 0; excessive rise-fall/monotonous = 1
- *Use of pauses* : appropriate = 0; inappropriate = 1
- *Use of rhythm* : appropriate = 0; dysrhythmic = 1
- *Rate of speech*: normal = 0; abnormal (slow/fast/variable) = 1
- *Articulatory proficiency*: Good = 0; Poor (imprecise consonants/prolongation of phonemes /repetition of phonemes/distorted vowels/ irregular articulatory breakdown) = 1

It was hypothesized that a higher score in speech naturalness would be correlated with higher percentage scores obtained for speech unintelligibility and vice versa.

6. *Correlation between speech intelligibility and speech naturalness*: Percentage correct identification in speech intelligibility tasks (word, sentence and narration) and weighted scores under speech naturalness assessment task were observed to see if the percentage intelligibility scores for speech intelligibility task of words, sentences and narration correlated with the weighted scores offered by judges for the various factors under speech naturalness.

7. *Reliability testing*:

- a. *Inter-judge reliability*: The percentage intelligibility scores offered by the 3 judges for the words, sentences and narration were tested for inter-judge reliability.
- b. *Intra-judge reliability*: To test for intra-judge reliability 10 words and 5 sentences from narration of each subject were randomly selected and subjected to judgment by the same 3 judges after a time interval of 1 week. Word intelligibility score, sentence intelligibility score and narration intelligibility scores were calculated. These scores were compared across the subjects and tasks and with the intelligibility ratings obtained in the first instance.

Results and Discussion

The speech samples of subjects with dysarthria on the three tasks were subjected to perceptual judgment of speech intelligibility and speech naturalness. Three judges performed this and the percentage intelligibility scores for word, sentence and narration samples were obtained. Speech naturalness was assessed on a 5 point binary scale after listening to the narrated speech sample of the subjects. The scores obtained from the perceptual judgment were subjected to inter-judge reliability and intra-judge reliability checks. Alpha reliability co-efficient was calculated to test inter-judge reliability and intra-judge reliability co-efficient was calculated to check for the intra-judge reliability. To test if the perceptual judgment scores tallied with the overall perceived severity of the disorder 3 other judges rated the severity of speech of the subjects with dysarthria on a 4 -point rating scale after listening to a 30 second narrated speech sample of the subjects. All the judges were kept blind to the purpose of the study and the

demographic details of the subjects. Table 1 shows the inter-judge reliability scores for word, sentence and narration task respectively.

Table 1: Inter-judge reliability scores for word, sentence and narration task

Judges	Reliability scores for Word task		Reliability scores for Sentence task		Reliability scores for Narration task
	Level-1	Level-2	Level-1	Level-2	-
Overall judge reliability	0.95	0.96	0.96	0.97	0.85
Judge 1 Vs. Judge 2	0.95	0.91	0.98	0.94	0.98
Judge 1 Vs. Judge 3	0.94	0.98	0.92	0.95	0.96
Judge 2 Vs. Judge 3	0.91	0.94	0.91	0.97	0.96

From Table 1 it is seen that the overall inter-judge reliability across the judges for all 5 subjects at both the level 1 and level 2 was good for the word task. The overall percent intelligibility scores varied from 72%-95% between the judges across all 5 clients for level 1 and it from 57%-90% for level 2. Similar to the word task the percentage speech intelligibility scores for 5 subjects as scored by 3 judges showed good reliability for both the levels of sentence task. The overall percent score for level 1 varied from 82%-95% and that for the level 2 was from 55%-95%. For the narration task too the reliability scores obtained between judges was good and varied from 65%-95%.

Intra-judge reliability check: The scores obtained for the intra-judge reliability are shown in Table 2 for each task between the judgment of judge 1, 2 and 3 respectively in the 1st and 2nd instance (the judgment task was repeated after a gap of 1 week)

Table 2: Intra-judge reliability for word, sentence and narration task

Judges (1st instance Vs. 2nd instance)	Intra-judge reliability scores for word task		Intra-judge reliability scores for Sentence task		Intra-judge reliability scores for narration task
	Level - 1	Level - 2	Level - 1	Level - 2	-
Judge 1	0.71	0.89	0.95	0.86	0.99
Judge 2	0.86	0.88	0.92	0.88	0.95
Judge 3	0.77	0.90	0.84	0.92	0.98

From table 2 it is observed that the intra-judge reliability for judge 1 and judge 3 is much lower when compared to the reliability scores obtained for judge 2 for the 2nd instances of level 1 in the word task. For the level 2 of the word task the reliability scores are almost similar across the judges. The relatively poor intra-judge reliability of Judge 1 and 3 for simple words could probably be due to deployment of too many guesses for the simple target words rather than depending on the acoustic cues to perceive the right target word. The poor scores also reflect the fact that these judges had not retained the target words and did not follow any pattern in their judgment as the two judgments were made after a gap of 1 week. For the sentence task the scores indicated no significant difference across the judges' scores for both the levels of the sentence task. For the narration task also there is good intra-judge reliability.

The judges were practicing speech-language pathologists and had work experience with the dysarthria clients. The overall high inter and intra judge reliability scores suggests that the judgment responses for the speech intelligibility evoked from the protocol was reliable.

Table 3 shows the comparison of the speech naturalness rating of the subjects' speech by the first three judges and the overall speech severity rating offered by second three judges.

Table 3: Comparison of speech severity and speech naturalness rating

Subjects	Speech severity rating by judges			Speech naturalness		
	Judge A	Judge B	Judge C	Judge A	Judge B	Judge C
S 1	Severe	Severe	Severe	83%	66%	83%
S 2	Mild	Mild	Mild	100%	100%	100%
S 3	Severe	Moderate	Moderate	66%	66%	66%
S 4	Mild	Moderate	Mild	83%	83%	100%
S 5	Severe	Severe	Severe	50%	50%	66%

Table 3 shows a comparison of ratings made by two groups of three judges, one for the severity of speech and another on the speech naturalness in the subjects. It is seen that the speech of subjects 1 and 5 have been rated as 'severe' degree of impairment and the percent scores for speech naturalness are also poor ranging from 66%-83% and 50%-66% respectively. The speech of subject 2 is rated as 'mild' degree of impairment and the percent scores for speech naturalness is high at 100%. The degree of speech impairment of subject 3 was rated as 'severe' by one judge and 'moderate' by two judges. The percent speech intelligibility was 66%. The degree of speech impairment of subject 4 was rated as 'moderate' by one judge and 'mild' by two judges. The percent speech intelligibility in this subject ranged from 83%-100%. Overall, the speech severity rating offered by three judges matched with the speech naturalness judgment given by three other judges. It is of importance to note that both the groups of judges were kept blind for the purpose of the study and the judgment was based on listening to the narrated speech sample of subjects and carried out independently by the judges in both the instances without mutual consultation.

Further, Spearman correlation test was run to verify the correlation of the 3 judges who rated the speech intelligibility scores for word, sentence and narration task and speech naturalness rating based on the narrated speech sample. This is shown in Table 4 and a significant correlation at a probability level of 0.05 was observed for the judgments made by all the three judges.

Table 4: Correlation values for speech intelligibility and naturalness ratings of 3 judges

Judges	Intelligibility Vs. Naturalness	Level of significance
Judge 1	0.913	P<0.05
Judge 2	0.913	P<0.05
Judge 3	0.889	P<0.05

Conclusion

From the reliability and correlation checks it is observed that the protocol developed is sensitive to assess speech intelligibility and speech naturalness in adult subjects with different types of dysarthria due to different neuropathologies. Thus protocol can be used as a clinical tool for assessing the level of speech intelligibility and speech naturalness in adult subjects with dysarthria, speaking Kannada language.

References

- Bernthal, J.E. & Bankson, N.W., (1998). Articulation and Phonological Disorders, *Fourth Edition, Boston: Allyn & Bacon.*
- Beukelman, D.R. & Yorkston, K.M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. *Journal of communication disorders, 13, 33-42.*
- Enderby, P. (1983). Frenchay dysarthric assessment. *British journal of disorders of communication, 51 (3), 165-174.*
- Kent, R. D., Weismer, G., Kent, J. E. & Rosenbek, J.C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of speech and hearing disorders, 54, 482-499.*
- Platt, L. J., Andrews, G., Young, M. & Quinn, P.T. (1980). Dysarthria of adult cerebral palsy: I. Intelligibility and articulatory impairment. *Journal of speech and hearing research, 23, 28-40.*
- Platt, L. J., Andrews, G., Young, M. & Neilson, P.O. (1978). The measurement of speech impairment of adults with cerebral palsy. *Folia phoniatica, 30, 50-58.*
- Smith & Nelson (1985). Cited in Perceptual judgment of speech intelligibility in cerebral palsy. *Unpublished Masters Dissertation, University of Mysore, Mysore.*
- Tikofsky, R.S. & Tikofsky, R.P. (1964). Intelligibility measures of dysarthric speech. *Journal of speech and hearing research, 7.*
- Tikofsky, R. (1970). A revised list for the estimation of dysarthric single word intelligibility. *Journal of speech and hearing research, 13.*
- Weiss, Gordon. & Lillywhite (1987). Clinical management of articulation disorders. *St Louis, The C.V. Mosby company, First Edition.*
- Yorkston, K. M., Strand, E. A. & Kennedy, M. R. T. (1996). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology, 5 (1), 55- 66.*
- Yorkston, K.M. & Beukelman, D.R. (1980). A clinician judged technique for quantifying dysarthric speech based on single word intelligibility. *Journal of communication disorders, 13, 15-32.*
- Yorkston, K.M. & Beukelman, D.R., Charles, D., Traynor. (1988). Articulatory adequacy in dysarthric speakers: A comparison of judging formats. *Journal of communication disorders, 21, 351.*