

Assessment of speech in neurological disorders: Development of a Swahili screening test

N Miller, G Mshana, O Msuya, C Dotchin, R Walker, E Aris

Institute of Health and Society, Speech and Language Sciences, University of Newcastle, Newcastle-Tyne, UK
Nick Miller

Kilimanjaro Christian Medical Centre, Moshi, Kilimanjaro Region, Tanzania
Gerry Mshana
Oliva Msuya

Department of Medicine, North Tyneside General Hospital, North Shields, UK
Catherine Dotchin
Richard Walker

Department of Medicine, Muhimbili University College Hospital, Dar es Salaam, Tanzania
Eric Aris

Corresponding author: N Miller (nicholas.miller@ncl.ac.uk)

Assessments for acquired motor-speech disorders that look at movements of the articulators would appear at first glance to be universal. This may be true for the most basic non-speech aspects of movement. We argue that assessments for *speech* motor control must be attuned to language-specific variables to be fully valid. We describe the rationale for, and development of a motor-speech-disorder screening test for Swahili speakers which includes impairment measures as well as measures of intelligibility and speech-voice naturalness. We further describe its initial validation in terms of content validity, feasibility of administration and scoring without requirements for lengthy training and technical expertise and application to groups of people with and without Parkinson's disease in Tanzania.

Results indicate that the protocol is ready to use in so far as it is acceptable to users (clinicians, patients), is feasible to use, shows good inter-rater reliability, and is capable of differentiating performance in healthy speakers and those whose speech is disordered. We highlight needs for further development, including issues around training, development of local norms for healthy speakers and for speakers with a variety of neurological disturbances, and extension of the tool to cover culturally valid assessment of impact of communication disorders.

Keywords: *speech disorders, assessment, Swahili, neurological, Parkinson's*

S Afr J CD 2012;59(1):27-33. DOI:10.7196/SAJCD.86

Communication changes are a feature of many neurological conditions. Yet in many African languages there is a lack of specifically designed assessments of communication. Swahili is one such language. It is spoken as a first or second language by over 60 million people, predominantly in East Africa. In Tanzania, Kenya, Uganda, the Comoros and the Democratic Republic of Congo it serves as a national language. Nevertheless there are no standardised assessments for motor-speech disorders. This article reports the development and initial validation of a Swahili protocol to evaluate speech changes in neurological disorders. It is intended for use by health professionals involved in managing people with neurological problems. As such it reflects an example of the challenge to devise what Pascoe and Norman (2011) have termed contextually relevant resources.

Speech (taken here to include voice) disorders affect an individual's ability to pronounce the sounds they need to say words and make themselves understood. Language disorders affect the ability to understand or retrieve words and sentences, whether in spoken or written form. Language and speech impairments can appear together or arise independently of each other. Our focus here is on speech.

Speech changes may represent the first or only signs of changing neurological function (Ball, Willis, Beukelman & Pattee, 2001; Harel, Cannizzaro, Cohen, Reilly & Snyder, 2004). Typically though, they co-occur with a range of other motor and cognitive impairments. Speech assessment is therefore important from a (differential) diagnostic perspective. Clinically it can also provide sensitive, non-invasive outcome measures of decline or recovery and of the effects of drug, surgical and behavioural interventions. Speech evaluation plays an important role too, since communication changes exercise a profound social and psychological influence on the person with the disorder, and

their family, even when changes are not so severe as to make speech unintelligible (Miller, Noble, Jones, Allcock & Burn, 2008).

Speech impairment can be measured in different ways. A subsystems approach (Duffy, 2005; Kent, 2009) singles out breathing (capacity; control over inspiration-expiration; co-ordination with vocalisation), voice (laryngeal function; voice loudness, pitch and quality), velopharyngeal efficiency (degree of hyper- or hyponasality) and articulation (pronunciation of sounds and syllables) to assess underlying speech motor performance. It is also vital to assess the consequences that changes across these subsystems might have for overall intelligibility.

Many speech assessments exist (Duffy, 2005; Kent, 2009). In so far as breathing for speech, voice loudness, rate of tongue movements in producing sounds and similar measures would appear to reflect universal aspects of motor performance, it is tempting to assume that these batteries can be applied without further consideration to any language. This may be true of the most basic speed, sustainability, force and range of movement tasks divorced from meaningful words. However, as soon as one deals with variables specific to a given language and the impact of changes on intelligibility, it is imperative not to simply translate assessments from other languages but to adapt or construct them anew for the target language (Pascoe & Norman, 2011; Wild *et al.*, 2005). Some examples illustrate this point.

Human spoken languages use a vast array of sounds. However, they differ in which sounds precisely are used, which sounds contrast with others to signal differences in meaning, which sound combinations are permissible, and where in a word sounds may appear – whether it can occur initially, medially and finally or only one of two of those options;

in which positions in words they differ in the nature and functions of stress and intonation patterns used. So, for instance, English does not have the implosives of Igbo (Nigeria) and Hausa (Nigeria, Niger and regions of West Africa) or the clicks of Nama (Namibia, Botswana). Luganda (Uganda) operates with 5 vowels, Twi (Ghana, Ivory Coast) with 15, English, including diphthongs, with nearer 20. Hindi (India) and English both use aspirated and non-aspirated /p/ sounds. While in Hindi they signal differences in meaning (/kapi/ (copy) - /kap^hi/ (meaningful)), in English they do not (their distribution is determined by surrounding sounds). Swahili and English both have a /ŋ/ sound, but in English this cannot occur in initial position in words. Both languages have /m/, /t/ and /t/ sounds, but while the combinations /mt/ (mchuzi, curry/sauce) and /mt/ (mtori, banana soup/porridge) at the start of a word are permissible in Swahili, they are not in English.

Contrasting intonation patterns form part of the grammatical systems of languages (e.g. signalling a statement versus a question), but how these are applied across languages is not uniform. Further, in most European languages changes in tone on a vowel signal predominantly affective nuances. In many other languages a system of tones operates that signifies contrasts in word and grammatical meaning, depending on whether the word is spoken with a rising, falling, high or low pitch (Wong, Perrachione, Gunasekera & Chandrasekaran, 2009).

A further important observation regarding cross-language differences in speech disorders centres on cultural variables in perceptions of change. While one may be able to ascertain that a person can generate a given force in the lips when pronouncing /p/ or sustain /a:/ at 60 decibels for 10 seconds, such measures bear no direct relationship to the speaker's and listeners' perception of if, how and why that represents a problem or not. The acceptability of different rates of speech or loudness levels, the significance of different alterations to voice quality (breathy, creaky; harsh, soft), the tolerance for amount and degree of dysfluencies are all strongly rooted in language- and culture-specific variables (Altenberg & Ferrand, 2006; Bebout & Arthur, 1992; Kita, 2009; Mackey, Finn & Ingham, 1997; Yiu, Murdoch, Hird, Lau & Ho, 2008).

Given that languages differ in their sound structure and use and the perception of change in neurological disorders is strongly influenced by sociolinguistic and cultural variables, it follows, as argued above, that simply translating an assessment devised around the structure and rules of one language will be invalid if applied to a structurally different one. The solution is either to adapt an extant test to the structure of the new language or create a test specifically tailored to the new language.

We aimed to create a screening assessment for speech-motor dysfunction in Swahili speakers covering quantification of underlying speech-motor impairment and activity limitation levels (intelligibility and speech naturalness) (Kent & Kent, 2000; Yorkston, Strand & Kennedy, 1996). Given the context in which the assessment is designed to be applied (by clinicians from all backgrounds, since it may well be the case that there is no trained speech-language therapist (SLT) available; health services with minimum resources), the aim was to select measures that could be accomplished with minimum training in application, scoring and interpretation, without technical equipment beyond paper, pencil and (stop)watch and if possible a simple audio-recording device. In this initial development phase we also aimed to ascertain whether the assessment was suitable for purpose, i.e. to detect differences between speakers with and without a neurological illness and be sensitive to possible changes in speakers over time.

Methods

Assessment rationale

Impairment measures

The underlying impairment to speech in neuromuscular disorders stems from alteration in the range, strength, sustainability, stability and co-ordination of movements of the muscles/movements involved in breathing, phonation, velopharyngeal function and pronunciation (Duffy, 2005). Assessment of these variables is typically achieved through maximum performance speech tasks (Duffy, 2005; Kent,

Kent & Rosenbek, 1987) that challenge the patient to produce a sound or word as fast, loud, long, high or low as possible. We followed a subsystems approach to assessment, adapting standard recommended clinical tasks with demonstrated validity (Kent, 2009) that assess breath capacity and control for speech, voice loudness, pitch and stability, and tongue and lip movement, and fulfilled the conditions required for minimal equipment and training (Appendix A). The following paragraphs elucidate.

Prolongation of /a:/ for as long as possible gives an estimate of air reserve for speech (Kent, 2009). The task can also serve as a basis for voice assessment through attention to perceived control and appropriateness of loudness and pitch, stability (tremor; inappropriate swings in pitch or loudness), and voice quality (e.g. harsh/strained (spastic) v. breathy/weak (flaccid), diplophonic (cord palsies)). These can be measured instrumentally, but for present purposes of minimal technical outlay they can be scored perceptually (naked ear) using rating scales. Control over loudness, pitch level and range can be further evaluated by asking the speaker to produce /a:/ at gradually increasing and decreasing loudness and gradually rising and lowering pitch levels.

Speech diadochokinetic tasks (Ackermann, Hertrich & Hehr, 1995; Gadesmann & Miller, 2008; Ziegler, 2002) gauge tongue and lip movement parameters. The sound in the syllable is chosen to challenge a given movement, e.g. /pa/ for lips, /ta/ for tongue tip, /ka/ for tongue dorsum. One can measure time to produce 10 repetitions or number of repetitions in 5 seconds. Qualitative observations record how well the individual is able to remain on target – for instance do repetitions of /ba/ drift to what is heard as /ma/ because of velopharyngeal insufficiency? Does /pa/ drift to and from /ba/ from misco-ordination between oral and laryngeal gestures, or to /fa/ from decreased excursion or strength of lip movements?

Important information (co-ordination of movements; speech planning; apraxic difficulties v. neuromuscular, dysarthric impairment) can be gained through alternating syllable tasks (see diadochokinetic tasks above). The individual repeats sounds as fast as possible that contrast in place of articulation, e.g. /pa-ta-ka/. Time to produce 5 or 10 repetitions and ability to remain fluent and on target are measured. Ideally real words are used (as done in the protocol with Swahili words paa, taa, kaa), (i) since this aids understanding of the task, and (ii) because they relate more closely to real speech performance (Clark, 2003; Kent, 2004).

Activity limitation

Impairment measures do not necessarily relate to how far changes affect communication (Hartelius & Miller, 2010). To gauge the impact of changes on day-to-day activity other assessments are required. For this purpose we included a diagnostic intelligibility screen and speech naturalness rating (Appendices B & C).

Diagnostic intelligibility tests (DIT) (Kent, Weismer, Kent & Rosenbek, 1989; Weismer & Martin, 1992) address the problem of extremely poor intra- and inter-rater reliability of rating scales for assessment of intelligibility (Schiavetti, 1992). In DIT patients repeat a list of words and a listener (with no knowledge of what the intended words are) responds with what they believe has been said. Depending on availability and/or aims of the assessment, listeners can be clinical colleagues, family members, or untrained strangers unfamiliar with the person's speech. Since scores can differ between listener groups it is essential that on retest (e.g. after therapy) the same scorers are used. By totalling words recognised and analysing the pattern of mishearings one achieves a measure of intelligibility, as well as suggestions for sound contrasts that a speaker might have difficulty signalling.

DIT depend on devising matched parallel lists of words that differ by one sound from each other (minimal pairs, e.g. tea-pea, pay-pie, coat-code), and reflect the sound distributions, combinations and range of sound frequencies of the language. The protocol offers four parallel lists for Swahili following these principles. For administration, the examiner can

either select one of the lists, or to minimise rater learning effects or retest familiarity effects where the test is frequently applied, select one word randomly from each row to arrive at varied but matched sets of 25 words.

To estimate the overall impression of speech acceptability in the context of the gender, age and cultural expectations of the community the screening test employs a 1 - 5 naturalness/ disorderedness rating scale (1 – definitely a problem with speech; 5 – definitely no problem with speech) (Appendix C). Evaluation may be based on impressions from speech during general case history taking. The sentences included in Appendix B provide a more controlled task for comparisons across time and persons which are attuned to syllable (articulatorily simple v. complex), word (frequency), phrase length (shorter v. longer) and grammar and associated intonation (commands v. questions v. statements) patterns that constitute prominent variables in change perception. The sentences can also provide data to supplement ratings of pitch, loudness, stability of speech/voice (see above). Time to say the sentence(s) can be used as a reliable measure of speaking rate alongside previous diadochokinetic timing for maximum syllable rate.

In many neurological disorders speech output is more greatly affected when a speaker has to formulate responses themselves rather than repeat or read a prepared sentence (Bunton & Keintz, 2008; Ho, Iansek, & Bradshaw, 2002). Accordingly, speech examinations commonly include having a person describe an everyday activity (Kent, 2009). This affords a more realistic appraisal of the impact of the underlying impairment on day-to-day communication, as well as how speech production interacts with broader language and cognitive status. In the protocol here contrasting tasks offer the possibility of examining contrasts in loudness, pitch, stability, voice quality and naturalness between simple repetition tasks (saying 'paa', repeating single words in the intelligibility test) and the self-formulated speech while describing a common activity (Appendix A, making porridge).

Swallowing assessment

Swallowing and speech disturbances are not directly related to each other, but they do frequently co-occur and management of both often falls to the same person, in westernised countries typically an SLT. Hence we included as part of the screen the 150 ml water swallow test (Nathadwarawala, Nicklin, & Wiles, 1992) that has been shown to be valid and reliable at quantifying swallowing efficiency.

Validation procedure

The speech-motor and intelligibility measures chosen for the test have proven validity as measures of speech performance (Duffy, 2005; Kent, 2009; Kent *et al.*, 1989; Ziegler, 2002). The current protocol was also reviewed independently by six SLTs specialising in acquired neurological speech disorders to judge its suitability as a screen for assessing acquired motor-speech disorders.

Our aim was also to examine whether conducting the test was feasible in a community with minimal training (given the lack of SLTs and requirement to conduct brief training of other professionals); whether it was acceptable to users; and whether the newly devised materials could potentially detect differences in performance across individuals with and without a neurological disorder. To this end we piloted the tasks on a group of people with Parkinson's disease (PD) and a group matched overall for age and gender who were non-neurologically impaired. This was to establish the feasibility of the materials and tests, not specifically to examine differences between people with and without PD, which is the subject of a separate report.

Participants

We assessed people with PD and control participants. They were recruited from a community-based prevalence survey (Dotchin *et al.*, 2008) in Hai District, Tanzania. Participation was by voluntary informed consent following UK and Tanzanian Ethics Committee approved procedures.

Results are based on 26 people with PD (7 female) and the overall matched group without PD from the same district. The people with PD

were assessed before they commenced medical therapy, and 19 of them again 12 months later, after 3 months on medication. Four had died and 3 were too ill for reassessment.

Median age of individuals without neurological illness was 76 years (interquartile range (IQR) 67 - 78) and for those with PD median 78 years (IQR 70 - 84) ($p=0.10$). For the latter, mean estimated duration of symptoms was 5 years (range 0.25 - 19). As regards overall motor status their mean Unified Parkinson's Disease Rating Scale (UPDRS) (Goetz *et al.*, 2003) score was 50 (range 24 - 97).

Procedures

A Tanzanian PD nurse specialist received 3 hours' induction and training from an experienced UK SLT covering the rationale and procedures for the test, how to make the sound recordings (Edirol R1 and AKG C420 headmounted microphone) and scoring of items. Laminated directions sheets for field use for all tasks and for live scoring were provided.

People with PD were assessed before they commenced pharmacotherapy and approximately 12 months later after 3 months of levodopa treatment. Speech assessment by the nurse specialist took place in the participant's home at the same time as assessments of their motor, cognitive, mood and social status (often in the presence of other family members). Recordings were downloaded to laptop computer on site and returned on compact discs for cross-checking and analysis in the UK.

Data processing

Time for sustained 'ah' was noted at the time of assessment. Counts for number of repetitions in 5 seconds of 'paa' (roof in Swahili), 'kaa' (charcoal) and 'paa-taa (light)-kaa' were made at the time of assessment. Following standard practice speakers attempted each maximum performance task twice. The better performance was taken as their score. Results for /a:/ and syllable repetitions were compared across groups and time.

Speech rate for the sentence 'Wale watoto wanafanya kazi kwa bidii shambani' (those children are working hard in the field) was calculated from the acoustic waveform in syllables per second using PRAAT (Boersma & Wennink, 2011). Sound pressure level variability (standard deviation (SD) of mean fundamental frequency) was measured from PRAAT based on the same sentence.

To complete intelligibility test scoring, six native Swahili-speaking Tanzanian medical elective students studying in the UK heard recordings in random order of participants with and without PD saying words from the four parallel word lists (Appendix B). Recordings were played free field (Dell Inspiron laptop connected to Fostex Personal Monitor 6301B loudspeaker) in a quiet clinic office with volume setting the same for all tracks. They were blind to word-list number, speaker identity and group. Half the listeners heard tracks in reverse order. For each speaker they wrote down which word they believed they heard. The derived intelligibility score was the percentage of words correctly recognised across all listeners.

Measurement of the 150 ml swallowing test (ml per second) was calculated from volume drunk and time taken from records at the time of testing.

Results

Validity and feasibility

The screening test was deemed to have sufficient face validity as independently judged by health and other workers in the community where it was to be applied. Content validity was independently confirmed by review from a panel of speech-language pathologists experienced in neurological speech disorders asked to judge whether the test adequately and appropriately screened key dimensions for a speech-voice assessment in people with neurological disturbances. Neither group recommended any changes to the content or delivery of the protocol.

Feasibility was confirmed. Time to complete all sections typically took around 15 - 25 minutes. With the exception of one control speaker who did not understand the nature of the syllable repetition task, all participants were able to comprehend instructions and carry out the tasks correctly.

The nurse specialist was able to detail performance counts/times for items requiring live scoring. Ten per cent of counts and timing for 'paa'/'kaa' and 'paa-taa-kaa' were randomly selected and calculated from the audio-recordings by an experienced SLT blind to initial measurements. There was a high correlation (Spearman's r 0.96) for counts between raters with no significant difference for either control speaker recordings or people with PD. The time to repeat 'paa-taa-kaa' correlation of measures was similarly high (r 0.95) and there was no statistically significant difference between raters. There were some issues around audio-recording (see Discussion) which impacted on the quality and completeness of some data sets. For this reason there were variable numbers of individual scores employed for the analyses that follow. Group comparisons were conducted only on pairs where there were valid matched recordings available.

Differences between groups

Table 1 displays results obtained from the participants with and without PD. Columns 2, 3, and 5 present the descriptive summaries for the different groups/times while columns 4 and 6 record results for statistical tests looking at possible differences between groups/times. Statistically significant results are shown in bold.

On the prolonged /a:/ task, single syllable repetition rates and overall speech rate people with and without PD as groups did not perform statistically significantly differently to each other. On the more taxing multisyllable alternation task (paa-taa-kaa) there was a statistically significant difference between people with PD and controls and between baseline and follow-up for people with PD.

The effects of altered voice and articulation were also clear in the intelligibility and disorderedness ratings. People with versus without

PD scored significantly differently ($p=0.03$) on words correctly recognised by naïve listeners. People with PD were perceived by listeners to be significantly more disordered ($p=0.007$) in their speech than those without, based on perceptual rating of the sentence repetitions.

Differences between people with and without PD were statistically significant for the water swallow test (ml/sec). There was no statistically significant change ($p=0.09$) in performance between baseline and follow-up for the people with PD.

Discussion

We have developed the first preliminarily validated screening test in Swahili for speech changes in neurological disorders that is not simply an unadapted translation from English and that addresses activity limitation measures (intelligibility; naturalness) as well as impairment performance. The nurse specialist was able to acquire the skills to apply and score the test after minimal training, indicated by the absence of any data loss due to misinstructions for tasks, misapplication of tasks or misscoring.

The test was acceptable to participants. There were no objections to or questioning of words and tasks used. No one refused to carry it out, whether for ethical, practical or comprehension reasons. From the large variability in performance on some tasks it appears participants occasionally appeared to give suboptimal responses. It remains to be established from further observation and analysis whether this relates to cultural influences in carrying out unaccustomed testing or whether it pertains to issues around examiner training in eliciting maximum performance.

A major problem encountered, that affected quality and analysis of data, concerned difficulties with audio-track labelling and with simultaneously controlling audio-recording equipment and attending to speech performance in order to deliver live scores. These point to issues in training and methods employed for detailing live performance that must be addressed in later training development.

Table 1. Summary statistics for groups

Tasks and measurements	PD time 1 median/mean (SD/ IQR)	PD time 2 median/mean (SD/ IQR)	PD time 1 v. 2	Control median/mean (SD/ IQR)	PD1 v. control
Prolonged /a:/ (secs)	9.04 (4.47)	10.16 (4.12)	Z 1.06 <i>p</i> 0.29	9.71 (4.38)	Z 0.524 <i>p</i> 0.60
Paa in 5 secs	13.64 (5.57)	12.00 (4.45)	Z 0.34 <i>p</i> 0.74	12.44 (5.54)	Z 0.18 <i>p</i> 0.86
Kaa in 5 secs	14.05 (6.21)	12.67 (4.54)	Z 0.44 <i>p</i> 0.68	13.57 (4.84)	Z 0.03 <i>p</i> 0.97
Time to repeat paa-taa-kaa x5 (secs)	5.61 (1.62)	4.92 (0.43)	Z 4.71 <i>p</i> <0.001	4.5 (0.48)	Z 4.84 <i>p</i> <0.001
Speech rate (syllables/sec)	0.17 (0.034)	na	na	0.17 (0.022)	Z 0.82 <i>p</i> 0.41
Loudness/sound pressure level dB	<i>n</i> =24 48.21 (4.60)	na	na	<i>n</i> =24 46.0 (6.83)	Z 1.13 <i>p</i> 0.27
Loudness variability SD dB	<i>n</i> =24 4.22 (1.2)			<i>n</i> =24 3.86 (1.51)	Z 0.70 <i>p</i> 0.48
Intelligibility (% total all listeners)	<i>n</i> =26 72 (51-78)	na	na	<i>n</i> =14 75.5 (72-89)	Z 2.13 <i>p</i> 0.03
Disorderedness based on sentences (1 - 5, 5 normal)	<i>n</i> =18 1.75 (1-2.5)	na	na	<i>n</i> =26 4 (3.5-5)	<i>p</i> <0.007
Swallowing ml/sec	6.5 (3.73-10.89)	10.19 (6.44-12.38)	Z 1.7 <i>p</i> 0.089	19.68 (15.06-26.90)	Z 5.15 <i>p</i> <0.001

PD1 = people with PD before medication; PD2 = assessment after medication commenced; SD = standard deviation; IQR = interquartile range; na = task not assessed; *n* = number if not 26.

The test tasks were able to differentiate performance levels and correctly detected differences that were expected between people with and without PD ('paa-taa-kaa' repetitions, Ackermann *et al.*, 1995; Ho, Bradshaw, Cunnington, Phillips & Iansek, 1998; Ziegler, 2002), intelligibility (Miller *et al.*, 2007) and swallowing (Miller *et al.*, 2009). The fact that between- and within-group differences on the single-syllable repetition tasks did not reach statistical significance is unsurprising, given that the nature of speech changes in PD may not be sufficiently severe to register on impairment measures. The fact that the more challenging syllable-alternating task (Ho *et al.*, 1998) did detect significant differences supports this interpretation. Similarly, on the water swallow test, the time between baseline and follow-up may not have been sufficient to expect significant changes in people with PD, especially given that they received medical intervention in the interim. As PD is a progressive condition one would expect deterioration in function over 1 year but this is likely to be counter-balanced by the drug therapy they received for 3 months before the second assessment, which had a major impact on motor function in some cases (Dotchin, Jusbani, & Walker, 2011).

The protocol is ready to use. However, there are several features that ideally require further development or need local norms against which to interpret performance. The screen should also be tested on larger numbers and on other groups of people with neurological illness (e.g. stroke). Monitoring the ability of a wider group of people to apply the protocol would also be helpful rather than the one trained tester employed here. Next steps also include the development of culturally appropriate questionnaire measures of perceptions of change and perceived impact of speech changes.

Acknowledgements. This study was partly funded by a grant from the British Academy UK-Africa Academic Partnerships.

References

- Ackermann, H., Hertrich, I., & Hehr, T. (1995). Oral diadochokinesis in neurological dysarthrias. *Folia Phoniatrica Et Logopaedica*, 47(1), 15-23. doi.org/10.1159/000266338
- Altenberg, E. P., & Ferrand, C. T. (2006). Perception of individuals with voice disorders by monolingual English, bilingual Cantonese-English, and bilingual Russian-English women. *Journal of Speech Language and Hearing Research*, 49(4), 879-887. doi.org/10.1044/1092-4388(2006/063)
- Ball, L. J., Willis, A., Beukelman, D. R., & Pattee, G. L. (2001). A protocol for identification of early bulbar signs in amyotrophic lateral sclerosis. *Journal of Neurological Sciences*, 191(1-2), 43-53. doi.org/10.1016/S0022-510X(01)00623-2
- Bebout, L., & Arthur, B. (1992). Cross-cultural attitudes toward speech disorders. *Journal of Speech and Hearing Research*, 35(1), 45-52.
- Boersma, P., & Wennink, D. (2011). Doing phonetics by computer. Available from <http://www.fon.hum.uva.nl/praat/> 30 September 2011.
- Bunton, K., & Keintz, C. (2008). Use of dual task paradigm for assessing speech intelligibility in clients with Parkinson's disease. *Journal of Medical Speech-Language Pathology*, 16(3), 141-155.
- Clark, H. M. (2003). Neuromuscular treatments for speech and swallowing: a tutorial. *American Journal of Speech-Language Pathology*, 12, 400-415. doi.org/10.1044/1058-0360(2003/086)
- Dotchin, C., Jusbani, A., & Walker, R. (2011). Three year follow up of levodopa plus carbidopa treatment in a prevalent cohort of patients with Parkinson's disease in Hai, Tanzania. *Journal of Neurology*, 258(9), 1649-1656. doi.org/10.1007/s00415-011-5988-y
- Dotchin, C., Mtsuya, O., Kissima, J., Massawe, J., Mhina, A., Moshy, A., Walker, R. (2008). The prevalence of Parkinson's disease in rural Tanzania. *Movement Disorders*, 23(11), 1567-1572. doi.org/10.1002/mds.21898
- Duffy, J. (2005). *Motor speech disorders* (2nd ed.). St Louis: Mosby.
- Gadesmann, M., & Miller, N. (2008). Reliability of speech diadochokinetic test measurement. *International Journal of Language & Communication Disorders*, 43(1), 41 - 54. doi.org/10.1080/13682820701546813
- Goetz, C. G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G. T., Fahn, S., Seidl, L. (2003). The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. *Movement Disorders*, 18(7), 738-750.
- Harel, B. T., Cannizzaro, M. S., Cohen, H., Reilly, N., & Snyder, P. J. (2004). Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment. *Journal of Neurolinguistics*, 17(6), 439-453. doi.org/10.1016/j.jneuroling.2004.06.001
- Hartelius, L., & Miller, N. (2010). The ICF framework and its relevance to the assessment of people with motor speech disorders. In L. Lowit & R. Kent (Eds.), *Assessment of motor speech disorders*. San Diego: Plural.
- Ho, A. K., Bradshaw, J. L., Cunnington, R., Phillips, J. G., & Iansek, R. (1998). Sequence heterogeneity in Parkinsonian speech. *Brain and Language*, 64(1), 122-145. doi.org/10.1006/brln.1998.1959
- Ho, A. K., Iansek, R., & Bradshaw, J. L. (2002). The effect of a concurrent task on Parkinsonian speech. *Journal of Clinical and Experimental Neuropsychology*, 24(1), 36-47. doi.org/10.1076/jcen.24.1.36.972
- Kent, R. (2004). The uniqueness of speech among motor systems. *Clinical Linguistics & Phonetics*, 18(6-8), 495-505. doi.org/10.1080/02699200410001703600
- Kent, R. (2009). Perceptual sensorimotor speech examination for motor speech disorders. In M. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 19-29). New York: Thieme.
- Kent, R., & Kent, J. (2000). Task-based profiles of the dysarthrias. *Folia Phoniatrica Et Logopaedica*, 52(1-3), 48-53. doi.org/10.1159/000021512
- Kent, R., Kent, J., & Rosenbek, J. C. (1987). Maximum performance tests of speech production. *Journal of Speech and Hearing Disorders*, 52(4), 367-387.
- Kent, R., Weismer, G., Kent, J., & Rosenbek, J. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, 54(4), 482-499.
- Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, 24(2), 145 - 167.
- Mackey, L. S., Finn, P., & Ingham, R. J. (1997). Effect of speech dialect on speech naturalness ratings: A systematic replication of Martin, Haroldson, and Triden (1984). *Journal of Speech Language and Hearing Research*, 40(2), 349-360.
- Miller, N., Alcock, L., Hildreth, A. J., Jones, D., Noble, E., & Burn, D. J. (2009). Swallowing problems in Parkinson disease: frequency and clinical correlates. *Journal of Neurology, Neurosurgery and Psychiatry*, 80(9), 1047-1049. doi.org/10.1136/jnnp.2008.157701
- Miller, N., Alcock, L., Jones, D., Noble, E., Hildreth, A. J., & Burn, D. J. (2007). Prevalence and pattern of perceived intelligibility changes in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, 78(11), 1188-1190. doi.org/10.1136/jnnp.2006.110171
- Miller, N., Noble, E., Jones, D., Alcock, L., & Burn, D. J. (2008). How do I sound to me? Perceived changes in communication in Parkinson's disease. *Clinical Rehabilitation*, 22(1), 14-22. doi.org/10.1177/0269215507079096
- Nathadwarawala, K. M., Nicklin, J., & Wiles, C. M. (1992). A timed test of swallowing capacity for neurological patients. *Journal of Neurology, Neurosurgery and Psychiatry*, 55(9), 822-825. doi.org/10.1136/jnnp.55.9.822
- Pascoe, M., & Norman, V. (2011). Contextually relevant resources in speech-language therapy and audiology in South Africa - are there any? *South African Journal of Communication Disorders*, 58, 2-5.
- Schiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In R. Kent (Ed.), *Intelligibility in speech disorders* (pp. 11-34). Amsterdam: Benjamins.
- Weismer, G., & Martin, R. (1992). Acoustic and perceptual approaches to the study of intelligibility. In R. Kent (Ed.), *Intelligibility in speech disorders* (pp. 67-118). Amsterdam: Benjamins.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for translation and cultural adaptation. *Value in Health*, 8(2), 94-104. doi.org/10.1111/j.1524-4733.2005.04054.x
- Wong, P. C. M., Perrachione, T. K., Gunasekera, G., & Chandrasekaran, B. (2009). Communication disorders in speakers of tone languages: Etiological bases and clinical considerations. *Seminars in Speech and Language*, 30(3), 162-173. doi.org/10.1055/s-0029-1225953
- Yiu, E. M. L., Murdoch, B., Hird, K., Lau, P., & Ho, E. M. (2008). Cultural and language differences in voice quality perception: A preliminary investigation using synthesized signals. *Folia Phoniatrica Et Logopaedica*, 60(3), 107-119. doi.org/10.1159/000119746
- 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.
- Ziegler, W. (2002). Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech. *Brain and Language*, 80(3), 556-575. doi.org/10.1006/brln.2001.2614

Appendix A. Speech and voice changes in people with Parkinson's disease

© Nick Miller, Gerry Mshana, Richard Walker

Speech and intelligibility assessment protocol (Swahili version available on request)

Ideally these tasks should be recorded on audio or video recorder for later scoring and comparison with later assessments. For the intelligibility test it is imperative. Use a new sound track for each task.

T = Time on stopwatch

DDK (Diadochokinetic repetition) repeat 'paa' for 5 secs	Say 'paa' (roof) as many times as you can until I say stop (i.e. after 5 seconds). Demonstrate with syllable /ma/ until person has understood the task. Elicit twice. Score: total number of repetitions of the syllable in the time on watch (as close to 5 secs as possible)	Repetitions in 5 secs Enter time on watch (e.g. 4.94) and total repetitions (e.g. 21)
T		1st attempt: 2nd attempt:

Appendix A. Speech and voice changes in people with Parkinson's disease (continued)

Repeat 'kaa' (charcoal) for 5 secs T	Instructions as for 'paa'	Repetitions in 5 secs Live 1 st attempt Live 2 nd attempt					
Repeat 'paa-taa-kaa' 5 times T	Can you say for me 'paa-taa-kaa' as fast as you can until I say stop? Elicit twice. Score time in seconds to say 5 complete sets	Time to repeat 5 times: 1st attempt: Time to repeat 5 times: 2nd attempt					
Intelligibility list	'I have some words I would like you to repeat after me. Just repeat them in your own time at your own speed'. Examiner: Remember not to record you saying which word it is. Remember to pause recording after person has said the previous word and restart after you have said the next word but before they reply.	Circle which word list was used Number: 1 2 3 4					
		Words correctly recognised by:					
		Listener					
		1	2	3	4	5	6
	Mean overall score for listeners 1 - 6						
Say /ah/ as long as possible T	'I want you to say /ah/ for as long as you can. Like this, look (demonstrate till person has understood what to do). OK, now take in as deep a breath as you can and then slowly let it out saying 'ah' as long as you can. Start watch as soon as person starts 'ah'. Elicit twice. Score in secs.					Trial 1 secs Trial 2 secs	
Sentences for repetition Describe how to carry out everyday activity	Person repeats sentences after examiner in their own time. Same instructions for recording as for intelligibility. This time it doesn't matter though if scorers can hear the stimulus sentences. Examiner: Make sure you say commands as commands, questions as questions; if person doesn't, no need to correct them. Use sentences below after 'intelligibility' words Please can you describe to me from start to finish how 'mtori' (banana porridge) is prepared?						

Appendix B. Intelligibility test items

© N Miller, G Mshana, R Walker (Swahili version available on request)

Aim: The patient says the following words, one at a time, in their habitual voice/rate/loudness

Method: Either the individual reads one of the lists (can give them the list beforehand so they can check any words not familiar with), or the examiner says the word and the speaker repeats it after them. For retest purposes the same method (reading or repetition) must be employed.

If using repetition the examiner should take care not to record themselves saying the stimulus word. Easiest way to avoid that is to have recorder on PAUSE (not stop) when the examiner says the word; press start and give signal for speaker to say the word; put back on pause before next word by examiner.

NOTE: Examiner to say the word in natural, habitual way, making sure not to overemphasise any sounds/syllables or say the word in a deliberate fashion. Say the word only once unless the person has not heard it clearly and requests a repeat. If they repeat back the wrong word leave it and go onto next item – but record in field notes that Item X was given as 'ABC'. This needs to be taken into account when calculating scores.

Sets of words for Swahili diagnostic intelligibility screening test

Item number	List 1	List 2	List 3	List 4
1	Kana	Kaba	Kala	Kata
2	Piga	Pita	Pika	Piku
3	Pana	Pande	Panda	Panga
4	Baba	Babu	Bana	Basi
5	Cheza	Cheka	Chepe	Chema
6	Asali	Asili	Asile	Sifa
7	Nazi	Kazi	Vazi	Ngazi
8	Choka	Choma	Chora	Chota

Appendix B. Intelligibility test items (continued)

9	Njoo	Njaa	Njia	Njiwa
10	Sikio	Pitio	Salio	Kilio
11	Moja	Hoja	Ngoja	Onja
12	Paka	Baka	Taka	Waka
13	Duma	Dume	Dua	Dubu
14	Nyani	Nyati	Nyama	Nyavu
15	Tatu	Tano	Taka	Taga
16	Kaa	Paa	Zaa	Vaa
17	Waa	Twaa	Kwaa	Duwaa
18	Ninakula	Tunakula	Wanakula	Watakula
19	Bandua	Bangua	Pangua	Pungua
20	Walipata	Watapata	Ulipata	Utapata
21	Wewe	Hehe	Yeye	Nene
22	Pima	Kima	Hima	Lima
23	Kaja	Haja	Paja	Naja
24	Tuma	Tupa	Tua	Pua
25	Nawa	Tawa	Kawa	Sawa

Sentence repetition task

Aim: Person reads/repeats sentences below in their normal voice, loudness and rate.

Method: Either: individual reads the sentences out loud in normal voice (give them the sentences to look through first to check they know the words). Or: Speaker repeats the sentences after the examiner

Note on protocol sheet whether read or repeated. For retest use the same method.

Make sure the commands are said like commands, the questions like questions.

Instructions: I'm going to say some sentences/I want you to read some sentences. Just say the sentence in your normal voice in your normal way.

Sentences for people to repeat/read

- 1 Hapana, usiende pale!
- 2 Alisema uje au uende?
- 3 Wale watoto wanafanya kazi kwa bidii shambani.
- 4 Kaka ana kuku kumi na tano?
- 5 Ameenda sokoni kuuza kahawa yake.
- 6 Pita huku haraka!

Appendix C. Disorderedness/naturalness rating scale

'How sure are you that this person has a problem with their speech?'

1	2	3	4	5
Yes, definitely	Yes, maybe	I'm not sure	No, maybe not	Definitely not