The treatment of apraxia of speech

Hurkmans, Josephus Johannes Stephanus

Chapter 8

The Modified Diadochokinesis Test, an evaluation instrument for the treatment of Apraxia of Speech

This chapter was adapted from a peer-reviewed publication

Hurkmans, J., Jonkers, R., Boonstra, A., Stewart, R., Reinders-Messelink, H. (2012) Assessing treatment effects in Apraxia of Speech: Introduction and evaluation of the Modified Diadochokinesis Test, *International Journal of Language and Communications Disorders, 47*, 427-436.

8.1 | Introduction

The present study describes the development and evaluation of a new instrument to evaluate treatment in Apraxia of Speech: the Modified Diadochokinesis test (MDT).

Although the debate about the underlying deficit in Apraxia of Speech (AoS) is ongoing, there is agreement on at least some salient symptoms of this disorder. Ziegler (2008) characterised AoS as dysfluent, with groping and effortful speech, phonetic distortions, phonemic paraphasias, and a frequent occurrence of false starts. The key symptoms can be divided into three categories: (1) segmental impairments, (2) error variability and (3) prosodic impairment. Segmental deficits concern phonetic distortions (awkward-sounding speech sounds) and phonemic paraphasias (deletion, substitution and additions of speech sounds). The inconsistency in the production of errors tends to be large: a patient may produce the same sound accurately or inaccurately, while multiple inaccurate productions may have different qualities. Recent research, however, has suggested that errors may not be variable. Distortions have been found to be the predominant error type (Mauszycki, Wambaugh, & Cameron, 2010). Prosodic impairments concern disturbances in the flow and melody of speech. Speech is hesitant and halting with pauses between syllables, false starts, repairs and repetitive attempts at initiating speech (Ziegler, 2008, McNeil, Robin, & Schmidt, 2009).

AoS is often accompanied by other linguistic and motor execution disorders, such as aphasia and dysarthria (West, Hesketh, Vail, & Bowen, 2005). In both atactic dysarthria and conduction aphasia sound errors result in non-fluent speech and this may cause considerable problems in clinical diagnosis. The speech patterns of dysarthric patients are characterised by predictable and constant disturbances of speech phonation and articulation (Ziegler, 2008). This is an important distinction with AoS with variable articulatory skills. There is too little empirical evidence for the distinctions between phonological impairment and AoS. There are, however, some indications. It has been speculated that phonological impairment afflicts word and syllable endings more than word and syllable onsets, whereas the reverse is true for AoS (Aichert & Ziegler, 2004).

In clinical practice a variety of techniques are applied to treat patients with persisting AoS. The majority of objective evidence supporting treatment for AoS addresses articulatory-kinematic therapeutic approaches (Wambaugh, Duffy, Mc Neill, Robin, & Rogers, 2006b, Wambaugh & Shuster, 2008). Recent research has suggested that rate/rhythm treatments also may improve articulation in AoS (Brendel & Ziegler, 2008, Wambaugh & Shuster, 2008).The Melodic Intonation Therapy (MIT, Albert, Sparks, & Helm, 1973) is probably the best-known therapy in the rate/rhythm category. An important characteristic of the MIT is the assumption that melody and rhythm can support fluency of speech. Particularly in the field of music therapy there is great variety in approaches (see e.g. Modified MIT, Baker, 2000; SIPARI, Jungblut & Aldridge, 2004). Recently, Hurkmans, de Bruin, Boonstra, Jonkers, Arendzen, & Reinders-Messelink (2012) reviewed more therapy programmes that make use of musical elements in the treatment of non-fluent aphasic speakers and in speakers with AoS, among which Speech-Music Therapy for Aphasia (SMTA; De Bruijn, Zielman, & Hurkmans, 2005), a programme combining elements of speech therapy with music-based components. Like in the MIT, melody and rhythm play an important role here, but SMTA also uses other musical elements, such as dynamics, metre and tonality.

As mentioned above, most of these programmes were developed in clinical practice, and empirical support for their efficacy and effectiveness is still scarce. The few available studies have poor methodological quality; in a Cochrane meta-analysis of AoS interventions none fulfilled the criteria of a randomised controlled trial (West et al., 2005). However, several studies with small patient groups using behavioural treatments based on speech motor exercises do show substantial improvements in the samples tested (Ziegler, 2008, Wambaugh, 2002, Brendel & Ziegler, 2008). Using a multiple-baseline-across-behaviours design, Fucetola, Tucker, Blank and Corbetta (2005) showed that effectiveness evaluations of evidence-based aphasia treatment were feasible in clinical practice in small patient groups. This design refers to treatments with an established outcome at early trial phases now progressing to the effectiveness phase. Treatment evaluation included pre-treatment tests, weekly evaluations during the experiment, and post-treatment and follow-up assessments.

To date, there are no clearly defined methods to evaluate rate and rhythm control therapies for the treatment of AoS, although symptom changes can be evaluated with various tests, such as the Aachen Aphasia Test (AAT; Graetz, de Bleser, & Willmes, 1992), gauging overall language functions and the Amsterdam-Nijmegen Everyday Language Test (ANELT; Blomert, Kean, Koster, & Schokker, 1995), assessing functional communication (Mumby, Bowen, & Hesketh, 2007). A specific diagnostic instrument for AoS in Dutch was still missing until Feiken and Jonkers (2012) recently developed the Diagnostic Instrument for Apraxia of Speech (DIAS), measuring planning and programming of speech movements. This means that a diagnostic test for AoS can also be used for evaluation. A sensitive test that allows the improvement in phonetic encoding to be assessed on a weekly basis is, however, lacking.

Rapid syllable repetitions require alternating articulatory movements, allowing oral diadochokinesis (DDK) to be tested (Ackermann, Hertrich, & Hehr, 1995). Also Ziegler (2002) considered repetitions of monosyllables a sensitive measure for the assessment of the motor performance of speech. A DDK task might then also be adequate to evaluate the effectiveness of speech therapy at the phonetic encoding level. Most diadochokinesis tests distinguish sequential DDK, in which the same syllable (e.g., “pa” – “pa” – “pa”) is to be repeated, from alternating DDK, in which different syllables (e.g., “pa” – “ta” –“ka”) are to be alternated. The literature shows that patients with AoS perform better on the sequential tests than they do on the alternating tests (Wertz, LaPointe, & Rosenbek, 1984, Deger & Ziegler, 2002, Ogar, Willock, Baldo, Wilcins, Ludy, & Dronkers, 2006). Speech rate is often used as a variable to measure DDK performance and has been shown to provide a sensitive indicator of the presence and severity of neurological impairment (Ackermann et al., 1995). Speech rate can be scored by the count-by-time procedure, in which the number of repetitions within a preset interval is recorded, or by the time-by-count method, where the time needed to repeat an item is measured (Prathanee, 1998). Gadesmann and Miller (2008), however, reported problems regarding rate measurements and the intra- and inter-rater reliabilities of DDK tests. They also claimed that the relationship between impaired DDK and other functional speech measures is not transparent. Goozée, Murdoch and Theodoros (2001) argued that DDK performance does not predict intelligibility, or overall functional communicative success in any transparent way. Attempts to improve the tests, entailing modifications to allow a more accurate assessment of the control and output characteristics of natural speech, have met with equivocal results (Lowit, Miller, & Poedjianto, 2003).

Seeking to preclude the mentioned problems with assessing DDK in patients with AoS, the present study describes the development of the Modified Diadochokinesis Test (MDT), a new instrument to help assess the effects of rate and rhythm therapies in clinical trials as well as daily practice. The MDT was designed with the assessment procedure proposed by Fucetola, Tucker, Blank, & Corbetta (2005) in mind and an effectiveness evaluation of the Speech-Music Therapy for Aphasia (SMTA; De Bruijn, Zielman, & Hurkmans, 2005).

First, to be suitable on a weekly basis, the test needed to be concise as well as sufficiently sensitive. Second, to measure motor performance in AoS we opted to test DDK in both sequential and alternating syllable structures. We did not make use of language-based tests as the aim was to use these in the final pre- and post-treatment and follow-up evaluations of the SMTA. Finally, one important modification to a classical feature of DDK measurement was made. Given the poor intra- and inter-rater reliability Gadesmann and Miller (2008) reported, it was decided not to consider speech rate with the test.

In accordance with Ziegler's (2008) AoS symptom classification, the MDT assesses consistency (i.e., error variability), accuracy (i.e., segmental errors) and fluency of speech (i.e., prosodic disturbances). Also, to determine articulatory complexity the influence of syllable structure and type of alternation on motor performance was charted. The MDT's internal consistency and reliability (test-retest, intra- and inter-rater reliabilities) were investigated and validity (convergent and discriminant validities) was constructed.

8.2 | Methods

8.2.1 | Participants

Twenty-four adults with AoS took part in the trial. Relevant demographic data are provided in Table 8.1. All patients were referred to the rehabilitation centre "Revalidatie Friesland" or the Center for Rehabilitation of the University Medical Center Groningen by speech therapists working in the northern region of the Netherlands. Eleven patients were male and 13 female, with ages ranging from 34 to 78 years. Educational levels were equally distributed across patients. All had suffered a stroke, with post-onset times varying from 1 to 29 months. The diagnosis of AoS was established by experienced speech therapists on the basis of clinical judgment. The patients were tested with the recently developed Diagnostic Instrument for Apraxia of Speech (Feiken & Jonkers, 2012) to confirm the original diagnosis. Three patients were also dysarthric, as was established with the Radboud Oral Assessment (ROO; Kalf & de Swart, 2007) and the Radboud Dysarthria Assessment (RDO; Knuijt & de Swart, 2007). Eighteen patients were concomitantly suffering from aphasia (conduction aphasia and Broca's aphasia) as assessed with the AAT (Graetz et al., 1992).

Twelve control speakers with unaffected speech, matched for age, gender and educational level, were recruited from among hospital staff and by posters (at the rehabilitation centre). None of the participants reported above-average hearing loss or visual problems.

The study was approved by the medical ethics committee of the University Medical Center Groningen and informed consent was obtained from all participants.

Table 8.1 **Demographics of the participants with AoS and the unimpaired control speakers**

Pat nr	Medical diagnosis	MPO	Diagnosis	Age	Gender	Education
1	ICVA	5	AoS+aphasia	65	male	Intermediate
2	ICVA	3	AoS+dysarthria	42	male	High
3	ICVA	3	AoS+dysarthria	51	female	Low
4	ICVA	8	AoS+aphasia	34	female	Intermediate
5	HCVA	27	AoS+aphasia	62	female	High
6	ICVA	5	AoS+aphasia	52	male	High
7	ICVA	29	AoS+aphasia	49	female	Low
8	ICVA	6	AoS+aphasia	51	female	Intermediate
9	ICVA	16	AoS+aphasia	47	male	Low
10	ICVA	19	AoS+aphasia	38	female	High
11	ICVA	3	AoS+aphasia	75	male	Low
12	ICVA	3	AoS+aphasia	63	female	Intermediate
13	ICVA	2	AoS+aphasia	69	female	High
14	HCVA	7	AoS+aphasia	78	male	Low
15	ICVA	24	AoS	50	male	High
16	ICVA	5	AoS+aphasia	36	female	Intermediate
17	ICVA	3	AoS+aphasia	70	male	Intermediate
18	ICVA	2	AoS	39	male	Low
19	ICVA	1	AoS+aphasia	64	female	High
20	ICVA	4	AoS+aphasia	71	male	Low
21	ICVA	1	AoS+aphasia	70	female	Low
22	ICVA	1	AoS+aphasia	62	male	Intermediate
23	ICVA	2	AoS	47	female	High
24	ICVA	3	AoS+dysarthria	64	female	Low
Control participants (N=12)	-	-	-	M 55.8 SD 11.28	male 5 female 7	High 6 Intermediate 2 Low 4

ICVA=Ischaemic cerebrovascular accident, HCVA=Haemorrhagic cerebrovascular accident, MPO=months post onset, AoS=apraxia of speech, M=mean, SD=standard deviation

8.2.2 | Materials

8.2.3 | The Modified Diadochokinesis Test (MDT)

The MDT consists of four blocks, each comprising four strings of three single-syllable non-words with 16 items in total. Table 8.2 provides an overview of all sequences presented. Each block starts with sequential diadochokinesis and then systematically alternates three distinctive features, i.e., *place* of articulation (e.g., "pa" - "ta" - "ka"), *manner* of articulation (e.g., "da" - "na" - "la") and *vowel change* (e.g., "pa" - "po" - "pu").

The participants also completed related subtests of the AAT (articulation/prosody and phonological structure of spontaneous speech; the repetition subtest (repetition of phonemes, words and sentences)), the intelligibility measure of the ANELT, and the subtests DDK, articulation of phonemes and articulation of words of the DIAS.

Table 8.2 **The consonant/vowel (C/V) syllable sequences of the Modified Diadochokinesis Test (MDT) per block**

Block 1 (CV)	Block 2 (CVC)	Block 3 (CVCC)	Block 4 (CCVC)
pa pa pa (1.1)	paf paf paf (2.1)	paks paks paks (3.1)	spag spag spag (4.1)
pa ta ka (1.2)	paf taf kaf (2.2)	paps pats paks (3.2)	spag stag skag (4.2)
da na la (1.3)	daf naf laf (2.3)	pats pans pals (3.3)	stag snag slag (4.3)
pa po pu (1.4)	paf pof puf (2.4)	paks poks puks (3.4)	spag spog spug (4.4)

C=consonant, V=vowel

8.2.4 | Test and scoring procedures

All participants were tested by the same examiner: an experienced speech therapist. They were offered a visual representation of the MDT syllable sequences and were requested to repeat the examiner's model of each string five times as accurately as possible. If requested, the examiner repeated the string, but only once. The test started with two practice strings before the actual test sequences were presented. No particular instructions regarding speech rate were given. The tests were recorded on videotape.

Consistency was scored by comparing all five repetitions of the syllables in each item with each other. A repetition was awarded a score of 1 when it was identical to and 0 when it differed from the other repetition. The total consistency score per syllable string consequently ranged from 0 (all repetitions being different) to 4 (all repetitions being identical).

Accuracy was established by rating segmental impairments, i.e., errors concerning segments (phonemes), as reflected by (1) phonemic paraphasias such as deletions, additions and substitutions and (2) phonetic distortions. Each repetition of a syllable was scored as follows: 3 = response identical to target; 2 = one or two segmental errors; 1 = three to five segmental errors; 0 = more than five segmental errors. Total accuracy scores ranged from 0 to 15 per syllable sequence.

Fluency was assessed by scoring interruptions in the flow of speech, i.e., hesitations, halting (pauses between syllables), false starts, repairs, and repetitive attempts at initiating a syllable. The separate repetitions of the syllable strings were scored as follows: 1 = fluent; 0 = non-fluent, with the total fluency score ranging from 0 (each repetition of all or one of the three syllables being dysfluent) to 5 (each repetition being fluent).

To determine the test-retest reliability of the MDT, ten randomly selected speakers from the AoS group were tested twice and the same examiner scored the recorded sessions (the author) at a 4-week interval. To exclude spontaneous recovery as much as possible, it was verified that the post-onset time for all ten patients was at least 3 months. To establish the intra-rater reliability, the same examiner scored the videotapes of ten again randomly selected patients after an interval of at least 3 months. Inter-rater reliability was tested by comparing the ratings of ten sessions independently conducted by two examiners (the author and another experienced speech therapist and clinical linguist). The construct validity of the MDT was evaluated by testing its convergent and discriminant validities. The scores of the MDT were compared with the scores obtained with allied subtests of the AAT, the ANELT and the DIAS to determine the convergent validity. The scores of the MDT of the patient group were compared with the scores of the control group to establish the discriminant validity.

8.2.5 | Statistical analysis

Internal consistency was determined in two ways. Using a confirmatory factor analysis (CFA) the loading of the syllable sequences on the outcome measures consistency, accuracy and fluency was inspected. The correlations were subsequently tested as a function of the total scores on the three outcome measures (consistency, accuracy and fluency). The data were not normally distributed; a non-parametric test, i.e., Spearman's rho was therefore used.

The effects of syllable structure and distinctive features of articulation (place, manner and vowel change) were analysed with repeated measures ANOVAs. Correlation coefficients were determined using Spearman's rho for test-retest reliability, intra- and inter-rater reliability and convergent validity. Mann-Whitney *U*-tests were used to compare the scores of the patients and the controls. All data were analysed using SAS 9.2 or SPSS 16.0. The level of significance was set at $p \leq 0.05$, two-tailed. For an overview of the statistical analyses, see Table 8.3.

Table 8.3 **Overview of all statistical analyses conducted**

	Character	Analysis	Table
Internal consistency	syllable sequence	CFA	8.4
	total	Spearman	8.5
Complexity	syllable structure	ANOVA	-
	distinctive feature	ANOVA	-
Reliability	test-retest	Spearman	8.6
	intra-rater	Spearman	8.6
	Inter-rater	Spearman	8.6
Validity	convergent	Spearman	8.7
	discriminant	Mann-Whitney U	-

CFA=confirmatory factor analysis

8.3 | Results

The individual scores of the patients are provided in Appendix B.1.

8.3.1 | Internal consistency

The loadings of sequential DKK in the CV syllables were < 0.35. The loadings of all other CV and CVC structures ranged between 0.4 and 0.7. The loadings in the syllable structures with a cluster were > 0.7. The coefficients varied considerably but certain trends could be observed. The coefficients of the sequential DDKs (1.1, 2.1, 3.1 and 4.1) were the lowest for all three factors and the correlations became stronger in each block (see Table 8.4).

The total scores of the three MDT outcome measures correlated highly (see Table 8.5).

Table 8.4 **Loadings on the three Modified Diadochokinesis Test outcome measures**

	Consistency	Accuracy	Fluency
CV 1.1	0.34	0.30	0.12
CV 1.2	0.74	0.71	0.65
CV 1.3	0.61	0.78	0.55
CV 1.4	0.68	0.62	0.63
CVC 2.1	0.51	0.56	0.16
CVC 2.2	0.61	0.83	0.47
CVC 2.3	0.84	0.84	0.80
CVC 2.4	0.79	0.82	0.70
CVCC 3.1	0.87	0.87	0.77
CVCC 3.2	0.87	0.95	0.90
CVCC 3.3	0.97	0.99	0.94
CVCC 3.4	0.89	0.97	0.89
CCVC 4.1	0.85	0.74	0.86
CCVC 4.2	0.91	0.95	0.79
CCVC 4.3	0.95	0.98	0.95
CCVC 4.4	0.98	0.99	0.93

C=consonant, V=vowel

Table 8.5 Correlation coefficients for internal consistency for the three Modified Diadochokinesis Test outcome measures.

	Consistency	Accuracy	Fluency
Consistency	X	0.95 *	0.93 *
Accuracy	0.95 *	X	0.91 *
Fluency	0.93 *	0.91 *	X

* = $p < 0.05$

8.3.2 | Complexity

A significant effect for syllable structure ($F(3,195)=67.5$; $p<0.001$) was found. Post-hoc analyses revealed an effect on all structures, with mean differences of 9.3 between CV and CVC (CI 95%=6.2-12.3), 14.0 for the CVC-CVCC comparison (CI 95%=10.2-17.9), and 4.8 for CV-CC-CCVC (CI 95%=2.3-7.3). None of the three distinctive features of articulation had significantly affected the scores ($F(2,130)=2.18$; $p=0.118$).

8.3.3 | Reliability

The results of the reliability tests are summarised in Table 8.6. The test-retest reliability for all three MDT outcome measures was high ($r = 0.74$ - 0.97). The same holds for MDT intra-rater reliability ($r = 0.97$ - 0.99) and inter-rater reliability ($r = 0.96$ to 0.98).

Table 8.6 Correlation coefficients for the Modified Diadochokinesis Test outcome measures

	Intra-rater reliability	Inter-rater reliability	Test-retest reliability
Consistency	0.97 *	0.96 *	0.80 *
Accuracy	0.99 *	0.97 *	0.98 *
Fluency	0.98 *	0.98 *	0.84*

* = $p<0.05$

8.3.4 | Validity

The results of the construct validity tests are summarised in Table 8.7.

8.3.5 | Convergent validity

Analysis of the MDT outcome measures and the DIAS yielded significant correlations (ranging between 0.55 to 0.91) except for the correlation between the DIAS DDK test and the MDT fluency outcome measure (*r*=0.44). Analysis of the three MDT outcome measures and aphasia tests (AAT and ANELT) yielded significant correlations between the AAT measure phonological structure in spontaneous speech (ranging between 0.63 to 0.80) and ANELT's intelligibility measure (ranging between 0.52 to 0.70). No significant correlations were found for the AAT subtest repetition and the AAT measure articulation/prosody in spontaneous speech.

Table 8.7 **Correlation coefficients for the construct validity of the measures consistency, accuracy and fluency with related subtests and measures of the AAT, ANELT and the DIAS**

	AAT/art	AAT/phon	AAT/rep	ANELT/INTEL	DIAS/DDK	DIAS/art ph	DIAS/art wrd
consistency	0.461	0.803 *	0.473	0.709 *	0.726 *	0.912 *	0.905 *
accuracy	0.468	0.633 *	0.424	0.556 *	0.853 *	0.828 *	0.888 *
fluency	0.172	0.645 *	0.221	0.528 *	0.448	0.828 *	0.766 *

AAT/art = articulation in spontaneous speech, AAT/phon = phonology in spontaneous speech, AAT/rep = repetition of phonemes, words and sentences, INTEL = intelligibility in verbal communication, DIAS/art ph = articulation of phonemes, DIAS/art wrd = articulation of words, *=$p<0.05$

8.3.6 | Discriminant validity

The scores of the control speakers were at ceiling. The patient group scored significantly lower than the control group on all outcome measures: *consistency*: $Z=-4.51$, $p<0.001$; *accuracy*: $Z=-4.51$, $p<0.001$; and *fluency*: $Z=-4.55$, $p<0.001$).

8.4 | Discussion

In search of a more accurate approach to assessing therapy outcome in Apraxia of Speech that would be suitable for both clinical trials and practice, the Modified Diadochokinesis Test (MDT) was developed. The high level of internal consistency founded in this first evaluation illustrates that each item of the MDT has a strong relationship with the three outcome measures tested, i.e., consistency, accuracy and fluency of speech. Sequential diadochokinesis showed the lowest correlation to all variables. These results correspond with those described by Wertz et al. (1984), Deger and Ziegler (2002) and Ogar et al. (2006), who argued that patients with AoS experience more difficulty alternating syllables than repeating the same syllables. Even though sequential DKK contributed the least to the discriminative power of the MDT, for the purpose of evaluating effects of speech therapy in patients with severe AoS we would suggest preserving these items.

The scores on the MDT revealed a clear relationship between the effects of complexity (syllable structure) and internal consistency. Post-hoc analyses showed the highest scores for simple consonant-vowel structures and the lowest for CCVC structures as a function of consistency, accuracy, and fluency. When the data of the loadings on the three measures were examined more closely, higher coefficients (up to 0.99) were found in the CCVC sequences of each block. In other words, the patients with AoS in this study experienced the most problems repeating phonemes in a syllable sequence with a consonant cluster in the initial position,with these items contributing the most to the test's discriminative power. The opposite trend was observed in the distinctive features of articulation examined. Contrary to what we expected, no effect of alternations in the place and manner of articulation or vowel change was found. The loadings on the outcome measures were very similar. The variations in the three articulatory aspects generated similar performance results, with these items contributing equally to the three outcome measures. As with the sequential items, for the purpose of speech therapy evaluations it is nevertheless important to reserve these articulation variations in DKK testing to detect possible possible changes in these outcome measures when changes in syllable structure are lacking.

MDT inter-rater reliability was good, implying that the results obtained were not influenced by the rater. This contrasts the findings of Gadesmann and Miller (2008) who reported poor inter-rater agreement in a classical DDK task requiring fast syllable repetitions. We accordingly propose that speech rate might be an important determinant in the reliability of DDK tasks.

The results for convergent validity were more varied. Although some caution should be taken in the interpretation of the correlations, considering the number of analyses that were performed, it is clear that the MDT outcome measures correlated significantly with most of the related measures of the AAT, ANELT and DIAS, with three exceptions, one of which was the articulation/prosody measure of the AAT. Arguably, the MDT and AAT measures gauge different domains of speech. In the AAT, the articulation/prosody measure is designed to assess symptoms of dysarthria, prosody and speech rate. In this study, only three AoS patients also displayed dysarthria symptoms. Prosody influences speech fluency and is thus suitable to assess speech at the phonetic encoding level. In the MDT, however, all items concern single syllables. Disturbances concerning the flow of speech can hereby be assessed. Melody and rate of speech can only be assessed in sentences and spontaneous speech, as is the case in the AAT. Given that with the MDT speech rate is not gauged and sentences are not used, the absence of a correlation between the AAT subtest and the MDT measures is not surprising.

The repetition subtest of the AAT also did not correlate with the MDT. The first important difference between the two tests is that in the AAT repetitions involve phonemes, words (containing up to nine syllables; *wa-pen-stil-stands-on-der-han-de-ling*) and sentences, whereas the MDT exclusively uses syllables. The MDT scores did correlate significantly with the scores on the DIAS, which requires the repetition of phonemes and short words (up to three syllables). It is hence probable to assume that the repetition of long words and sentences largely determines the final score in the AAT. In addition, the AAT subtest requires one response, whereas the MDT, like the DIAS, requires multiple repetitions, calling for complex motor programming and planning.

Finally, the MDT fluency scores did not correlate with the DDK scores on the DIAS. In the DIAS subtest, repetitions are scored by counting the number of repetitions of the same or alternating syllables the speaker can produce in 8 seconds, while in the MDT speech rate does not play a role. Also, in the DIAS-DDK the presence of additions and distortions is measured, while in the MDT these errors are gauged in the accuracy domain, which measure did correlate significantly with the DIAS-DDK.

The MDT exploits the repetitive production of meaningless or pseudo-syllables to evaluate the effects of speech therapy and not words or sentences. Gadesmann and Miller (2008) stated the lack of association between para-speech tasks (such as DDK-based tests) and speech tasks. The present study, however, found significant correlations between MDT outcomes and those obtained with 'classical' speech tasks, the DIAS (articulation of phonemes and words) and the ANELT assessing functional language skills. And although Goozée et al. (2001) argued that DDK performance does not predict intelligibility, the MDT scores correlated significantly with the Intelligibility measure of the ANELT. Still, evaluating speech therapy solely by means of a para-speech task is, in our view, insufficient and inaccurate. Tasks such as the AAT, ANELT and DIAS should be used to supplement the MDT. Thus, to study the effectiveness of speech therapies (in this case the SMTA) appropriate comprehensive and well-developed assessment batteries should be employed as pre- and post-treatment and follow-up assessments, while the MDT can be used for the baseline measurements and the weekly assessments during the therapy period.

8.5 | Conclusion

The syllable sequences in the MDT gauge the test construct: the loadings on the three outcome measures were high and showed significant correlations. The test's internal consistency can therefore be regarded as strong. Syllable structure affected the performance of the patients in that they had more difficulty remaining consistent, accurate and fluent with increasing within-syllable phonemes. The location of the cluster also affected performance: outcomes on initial clusters were poorer than those for final clusters. MDT scores were not affected by any distinctive feature of articulation examined.

The results also showed the MDT to have strong test-retest, and intra- and inter-rater reliability. Furthermore, its discriminant validity was adequate; the test differentiated the healthy speakers from the speakers with AoS on all three outcome measures. With respect to convergent validity the analyses yielded diverse results. The MDT outcome measures correlated with the AAT phonological structure of spontaneous speech measure, the ANELT intelligibility measure, and all subtests of the DIAS except for the fluency-DDK comparison. No significant correlations emerged from the MDT scores and the AAT scores for articulation in spontaneous speech and repetition.

The study shows that the MDT has adequate psychometric properties, implying that it can be used to measure changes in speech motor control during treatment for AoS. The results demonstrate the validity and utility of the instrument as a supplement to speech tasks in assessing speech improvement aimed at the level of planning and programming of speech.