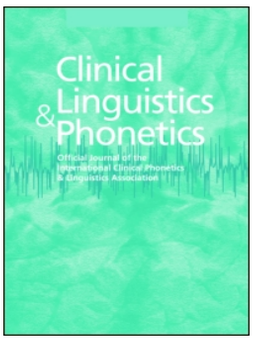
Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol

Laganaro, Marina; Fougeron, Cécile; Pernon, Michaëla; Levêque, Nathalie; Borel, Stéphanie; Fournet, Maryll; Catalano Chiuvé, Sabina; Lopez, Ursula; Trouville, Roland; Ménard, Lucie; Burkhard, Pierre; Assal, Frédéric; Delvaux, Véronique

How to cite

LAGANARO, Marina et al. Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol. In: Clinical Linguistics and Phonetics, 2021, p. 1–16. doi: 10.1080/02699206.2020.1865460

This publication URL: https://archive-ouverte.unige.ch/unige:151633
Publication DOI: 10.1080/02699206.2020.1865460



Clinical Linguistics & Phonetics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iclp20

Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol

Marina Laganaro , Cécile Fougeron , Michaela Pernon , Nathalie Levêque , Stéphanie Borel , Maryll Fournet , Sabina Catalano Chiuvé , Ursula Lopez , Roland Trouville , Lucie Ménard , Pierre R. Burkhard , Frédéric Assal & Véronique Delvaux

To cite this article: Marina Laganaro , Cécile Fougeron , Michaela Pernon , Nathalie Levêque , Stéphanie Borel , Maryll Fournet , Sabina Catalano Chiuvé , Ursula Lopez , Roland Trouville , Lucie Ménard , Pierre R. Burkhard , Frédéric Assal & Véronique Delvaux (2021): Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol, Clinical Linguistics & Phonetics, DOI: 10.1080/02699206.2020.1865460

To link to this article: https://doi.org/10.1080/02699206.2020.1865460

CLINICAL LINGUISTICS & PHONETICS
https://doi.org/10.1080/02699206.2020.1865460

Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol

Marina Laganaro[a], Cécile Fougeron[b], Michaela Pernon[b,c,d], Nathalie Levêque[b,e], Stéphanie Borel[f], Maryll Fournet[a,c], Sabina Catalano Chiuvé[c], Ursula Lopez[c], Roland Trouville[b], Lucie Ménard[g], Pierre R. Burkhard[c], Frédéric Assal[c], and Véronique Delvaux[h]

[a]Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland; [b]Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/University Sorbonne Nouvelle, Paris, France; [c]Geneva University Hospitals and Faculty of Medicine, University of Geneva, Geneva, Switzerland; [d]Centre de Référence Maladie Rare Pour la Maladie de Wilson, Service de Neurologie, Fondation Rothschild, Paris, France; [e]APHP, Département de Neurologie, Hôpital Pitié-Salpêtrière, Centre Référent SLA, Paris, France; [f]APHP Sorbonne Université, GHU Pitié-Salpêtrière, Service d'ORL & Institut du Cerveau et de la Moelle (ICM), Inserm/CNRS, Paris, France; [g]Laboratoire de Phonétique, Université de Québec à Montréal, Montréal, Canada; [h]FNRS & IRSTL (Language Sciences and Technologies Research Institute), University of Mons, Mons, Belgium

ABSTRACT

To respond to the need of objective screening tools for motor speech disorders (MSD), we present the screening version of a speech assessment protocol (*MonPaGe-2.0.s)*, which is based on semi-automated acoustic and perceptual measures on several speech dimensions in French. We validate the screening tool by testing its sensitivity and specificity and comparing its outcome with external standard assessment tools. The data from 80 patients diagnosed with different types of mild to moderate MSD and 62 healthy test controls were assessed against the normative data obtained on 404 neurotypical speakers, with *Deviance Scores* computed on seven speech dimensions (voice, speech rate, articulation, prosody, pneumophonatory control, diadochokinetic rate, intelligibility) based on acoustic and perceptual measures. A cut-off of the *MonPaGe total deviance score (TotDevS)* >2 allowed MSD to be diagnosed with *specificity* of 95% and an overall *sensitivity* of 83.8% on all patients pulled, reaching 91% when very mildly impaired patients were excluded. A strong correlation was found between the *MonPaGe TotDevS* and an external composite perceptual score of MSD provided by six experts. The MonPaGe screening protocol has proven its sensitivity and specificity for diagnosing presence and severity of MSD. Further implementations are needed to complement the characterization of impaired dimensions in order to distinguish subtypes of MSD.

ARTICLE HISTORY
Received 3 July 2020
Revised 12 December 2020
Accepted 14 December 2020

KEYWORDS
Motor speech disorders; screening; French; sensitivity; specificity

Introduction

Although both impaired language and impaired speech can be associated with a variety of neurodegenerative pathologies and of acquired brain lesions, clinicians have at one's

CONTACT Marina Laganaro marina.laganaro@unige.ch FPSE, Université de Genève, 40 bd du Pont-d'Arve, Genève CH-1211.

Supplemental data for this article can be accessed on the publisher's website.

disposal a large choice of language assessment and screening tools in many languages, while very limited tools are available for the assessment of speech.[1] In addition, the few assessment tools available for motor speech disorders (MSD) in clinical practice are mostly based on perceptual judgments. To complement the field of perceptual assessment tools for MSD, a consortium made up of researchers and clinicians created the MonPaGe speech screening battery, which is aimed at going beyond the perceptual-only judgments of speech and enabling more objective, mainly acoustic-based, multiparametric quantification of speech disorders. Here, we test whether the screening version of the protocol (*MonPaGe-2.0.s*) can be reliably used as a screening tool for the detection of MSD by assessing its specificity and sensitivity on a group of 80 patients with various patterns of MSD and a group of neurotypical test controls against a normative database gathered from 404 French-speaking healthy adults. Crucially, the potential of the protocol to diagnose the presence of MSD in the patient group is tested by including speakers with mild MSD in the patient cohort. The severity score obtained with the screening protocol is also tested against an external assessment of the speech disorder severity provided by a perceptually based severity assessment made by six expert judges using a tool widely used in French-speaking clinical settings.

Motor speech disorders and their assessment

Following acquired or degenerative brain disorders, speech can be impaired at different degrees of severity and at different levels of the speech production process. Sub-types of MSD have been determined at distinct levels. At a first branching, a distinction is made between dysarthria and apraxia of speech (AoS). Dysarthria further groups several syndromes of impaired speech which are the consequences of impaired control of the neuromuscular commands involved in speech production, giving rise to altered voice and/or articulation and/or prosody and/or speech rate (J. Duffy, 2005; McNeil et al., 2008). Several subtypes of dysarthria are traditionally distinguished (spastic, flaccid, hypokinetic, hyperkinetic, ataxic and mixed, Darley et al., 1975; to which unilateral upper motor neuron dysarthria and undetermined dysarthria have been added in more recent; J. Duffy, 2005; J. R. Duffy, 2013). AoS on the other hand, refers to MSD that is not attributed to motor execution and control, but to impaired phonetic planning (Blumstein, 1990; Code, 1998; Darley et al., 1975; McNeil et al., 2008; Varley & Whiteside, 2001; Wertz et al., 1984; Ziegler, 2008, 2009; 2012). AoS and dysarthria both result in altered articulation, speech rate and prosody, with an overlap in the distortions observed in the dysarthric and AoS speech profiles.

Despite the fact that the two types of MSD share speech distortions and need to be differentiated, most assessment tools have been developed separately for dysarthria and for AoS (Dabul, 2000; Feiken et al., 2008; Feiken & Jonkers, 2012; Strand et al., 2014). For instance, the Apraxia Battery for Adults (ABA-2, Dabul, 2000) has been conceived to assess AoS and is composed of tasks focusing on articulation, diadochokinetic rate and non-verbal oral movements, but does not assess other parameters such as voice, prosody or intelligibility; similarly, the AoS Rating Scale (ASRS, Strand et al., 2014) has been conceived to specifically identify and describe the perceptual speech characteristics of AoS. On the other

[1]Unless otherwise specified, voice is considered to be part of speech.

hand, assessment tools have been developed specifically for dysarthria (among others, for English: the Frenchay dysarthria assessment, Enderby & Palmer, 2008, which has been adapted in several languages; for German: BoDys, Ziegler et al., 2015; for Dutch: The Radboud Dysarthria Assessment (RDA), Knuijt et al. (2017); for French: The Batterie d'Evaluation Clinique de la Dysarthrie-BECD, Auzou and Rolland-Monnoury, 2019, the DIADOLAB, Menin-Sicard and Sicard (2020), see also the partial adaptation of the Frenchay by Ghio et al., 2019). These tools may be used to assess AoS also, but they do not always include diadochokinetic rate or a comprehensive assessment of articulation.

Whether conceived to assess dysarthria or AoS, speech assessment batteries available in clinical practice mostly rely on perceptual classification of speech and voice disorders, often using a large set of feature descriptors, which are well mastered only by SLP with an expertise in motor speech disorders. There are some well-known downsides to the perceptual description of speech features. First, the reliability and reproducibility of perceptual judgment of speech parameters in MSD have often been questioned (Bunton et al., 2007; Haley et al., 2012). This is in particular the case when speech is assessed perceptually based on specific dimensions, whereas better interrater agreement is observed on overall severity scores, at least for trained raters (Ziegler et al., 2017). Although assessment is based on a balance of subjective and objective evaluation in many domains, subjective evaluation is preponderant in the perceptual assessment of specific speech parameters, as they rely on individual internal representations that are based on the natural variability of these speech parameters and that become somewhat shared with training, but not ascertained.

In the light of the acknowledged limitations of perceptual evaluation, acoustic analyses can be of great help to quantify disordered dimensions in a clinical setting. Acoustic analyses have been implemented in *voice* assessment clinical tools (e.g., AVQI: Maryn et al., 2010, 2014; DSI: Wuyts et al., 2000), but have not systematically been applied to other speech parameters such as speech rate, articulation or prosody, nor generalised to assessment tools used in clinical practice. Hence, although global, holistic perceptual impression represents the very first information available to the clinician in front of the patients, tools are needed to further guide the assessment, combining functional, perceptual judgments with objective acoustic descriptors of impaired speech parameters. One challenge of this approach is to select a limited number of clinically relevant descriptors, with no claim to be exhaustive with regards to all potential measures. Indeed, there is a constantly growing amount of measures, acoustic analyses and statistical techniques for characterizing a diversity of aspects of the acoustics of dysarthric speech reported in the scientific literature. However, they are typically very specific, quite sophisticated, not implemented as stand-up procedures in the most popular softwares for voice and speech analysis, such as Praat (Boersma & David Weenink, 2018) or the Multi-Dimensional Voice Program (MDVP, from Kay Elemetrics Corporation); in short, they are not readily available for clinical purpose.

The MonPaGe speech assessment protocol and the MonPaGe screening tool

The MonPaGe speech assessment protocol (Fougeron et al., 2018, 2016; Pernon et al., 2020) has been developed (1) to provide a comprehensive speech assessment tool for French-speaking adults and (2) to integrate objective descriptors of potentially deviant speech parameters obtained with acoustic analyses. The MonPaGe protocol integrates the main

speech parameters used to assess both AoS and dysarthria in clinical assessment tools. These parameters are quantified either perceptually or acoustically, in a way that can be carried out by SLT or other professionals qualified to assess speech but who do not necessarily have an expert competence in acoustic phonetics or speech engineering. The *MonPaGe-2.0.s* presented here is a screening version based on a subset of speech dimensions and parameters from the MonPaGe protocol and aimed at detecting speech patterns that deviate from normal conditions. A speech screening tool should allow to determine whether the patient's speech deviates from typical speech, how much it deviates, and on which speech parameters. This usually corresponds to the first step in clinical assessment. The second step, namely an in-depth speech assessment contributing to the differential diagnosis of subtypes of MSD, goes beyond the aims of the screening tool.

In order to consider *MonPaGe-2.0.s* as a valid screening tool for MSD, the tool should be sensitive enough to detect most patients with a clinical diagnosis of MSD, even with mildly disordered speech (high sensitivity), and at the same time it should be specific enough to avoid over diagnosis, i.e. to avert diagnosing healthy speech as disordered (high specificity). In addition, the assessment given with *MonPaGe-2.0.s* should present similar trends to the ones observed with other assessment methods which are already widely used in clinical practice, such as perceptual-only evaluations.

Method

Population

Reference group: healthy speakers

For the reference values on the speech dimensions tested in MonPaGe, a group including 404 neurotypical speakers aged 20–93 years was selected from the *MonPaGe_HA* (for Healthy Adults) database of spoken French (Fougeron et al., 2018). Participants were recruited in four French-speaking locations: Mons, Belgium (n = 104); Montréal, Canada (n = 103); Paris, France (n = 101); Geneva, Switzerland (n = 97). All participants spoke French as their primary language (mother tongue) and currently used language. The population recruited for the *MonPaGe_HA* database is balanced within and between countries on sex (209 women and 195 men) and age groups ([20–39], [40–49], [50–59], [60–74], [75+] (see Fougeron et al., 2018 for more details)).

In order to provide normative data based on the largest possible relevant groups of speakers, the entire database was split into three age groups: younger speakers aged 20 to 59 years old [20–59] (N = 231, 119 women, mean age = 41.81, SD = 11.44), and older speakers further split into two groups [60–74] (N = 94, 48 women, mean age = 66.01, SD = 4.31) and [75–93] (N = 79, 42 women, mean age = 81.71, SD = 4.30). This choice was driven by a first screening of the data which confirmed findings in the literature showing little evidence of age effects on speech between 20 and 60. The two oldest speakers groups include less speakers but were kept as subgroups in response to clinical needs, where appropriate reference values based on such partition are needed for the elderlies.

Participants were recruited in the local communities through public advertisements and among relatives. Distribution of educational level (age of finishing school) was wide and similar across countries, with a median age of finishing school at 22 years old (range: 13 to

48). Speakers in the oldest group (>75) were also screened for language and cognitive deficits (with either the e-GeBAS, Chicherio et al., 2019, or MMSE, Folstein et al., 1975).

Test group 1: speakers with MSD

Eighty patients aged 24 to 81 (mean age = 57.1) diagnosed with mild to moderate MSD recruited in France (Pitié-Salpêtrière University Hospital and Lariboisière Hospital, Paris), in French-speaking Switzerland (Geneva University Hospitals) and French-speaking Belgium (Hôpital André Vésale, Charleroi University Hospital) were included for the validation. In order to vary the speech profiles, patients with MSD associated with six different underlying pathologies were included, namely Parkinson disease (PD), Amyotrophic Lateral Sclerosis (ALS), Friedreich ataxia (FA), Kennedy disease (spinal and bulbar muscular atrophy, SBMA), Wilson disease (WD), and post-stroke AoS. The neurological diagnosis was established by neurologists in the hospitals where patients were recruited. To be included in the present study, the patients had to be French native speakers, present mild or moderate acquired or progressive speech difficulty noticed by the patient and a SLP, no or only very mild language impairment. Patients with diagnosed dementia or psychiatric disorders and patients with history of developmental speech and language disorders or hearing impairment were excluded. Table 1 presents the distribution and description of the patients over the six clinical groups with associated basic descriptors.

The speech of each of the 80 patients was assessed by six experts (either SLP or clinical phoneticians) on a perceptual basis. A Composite Perceptual (hereby "CP") score, equivalent to the BECD's perceptual score (Auzou & Rolland-Monnoury, 2019) was computed based on a perceptual rating of the participant's speech on five dimensions: voice quality, segmental realization, prosody, intelligibility and naturalness of speech. Each dimension was assessed on a 5-point scale (0: normal to 4: severely impaired) on the recording of a sample of about 2 min. of continuous read speech. The composite perceptual score ranging from 0 to 20 was obtained by summing the five scores.

Test group 2: healthy control speakers

Sixty-two neurotypical French-speaking control participants who were not included in the reference group were recruited in Geneva. They were aged 22 to 88 (mean age = 53.4, 21 men) and had no history of neurological disease or speech and language disorders.

The study was approved by the local ethic committees at the institutions where the participants were recruited. All participants signed informed consent before participating in the study.

Table 1. Patient distribution over the six clinical groups and associated descriptors: sex, age, origin and severity measured with the Composite Perceptual (CP) score.

Group	*N* (*male*)	Mean age	City of recrutement	CP score, mean over 6 raters (*range*)
Parkinson disease	20 (*12*)	67.9	Geneva (10), Mons (10)	4.5 (*1.0–11.0*)
Amyotrophic Lateral Sclerosis	20 (*11*)	64.7	Paris (10), Geneva (10)	5.3 (*0.7–10.5*)
Kennedy disease	10 (*10*)	66.1	Paris	5.8 (*1.8–12.7*)
Friedreich Ataxia	10 (5)	37.6	Paris	9.8 (*2.3–15.3*)
Wilson disease	10 (*9*)	35.2	Paris	9.0 (*5.0, 13.5*)
Post-stroke AoS	10 (*4*)	52.5	Geneva	8.5 (*4.8–13.7*)

Speech material and speech descriptors

The MonPaGe protocol assesses different speech dimensions based on a variety of tasks (a detailed description of the MonPaGe protocol is provided in Fougeron et al., 2018). Here, a subset of the original speech descriptors was retained for the *MonPaGe-2.0.s* screening tool for MSD, providing information on seven dimensions, namely: intelligibility, articulation, pneumophonatory control, voice, prosody, speech rate and diadochokinetic rate. These dimensions have been selected in agreement with the expert SLT associated with the project, based on (i) dimensions that are most used in clinical practice and in other assessment tests and (ii) a balance between the time necessary to assess each dimension and its informativity. Note that the speech descriptors resulting from this procedure, although each relating to a specific dimension, have not the ambition to reflect all the potential disorders on that dimension of speech. Ultimately, they have been selected for their typicality, because they represent the participant's level of competence on that dimension, and/or for their informativity (with respect to the aim of developing a screening tool for MSD), not for the sake of comprehensiveness. For example, prosody is assessed in the extensive MonPaGe protocol via several acoustic and perceptual measures of the distinctive, demarcating and expressive functions of intonation in French, as exhibited by the reading of a tale and the production of selected sentences. In the end, only the ability to express an interrogation via melodic modulations is incorporated in the screening tool. Similarly, pneumophonatory control is indexed by, but by no means reducible to, the indicator of maximum phonation time, which is very commonly used in clinical practice, although its relationship with laryngeal and respiratory physiology is admittedly highly complex (e.g., Solomon et al., 2000).

The seven speech dimensions, the associated speech tasks and the resulting selective descriptors are detailed below. The measurements were carried out off-line using a computerized tool integrating perceptual tasks for scoring intelligibility, coding pseudowords errors, etc. and a customized Praat script dedicated to human supervision of acoustic measurements.

Intelligibility

A short intelligibility test was administered in the form of an interactive task between the experimenter and the participant in a face-to-face setting. The participant was asked to instruct the experimenter to place some test-words on a 5 × 5 grid combining icons of various shapes and colors, using a pre-learned carrier sentence ("Place the word [target_word] on the [color] [shape]" (e.g., 'Place the word "dog" on the red circle')). The color-shapes combinations were of limited number, while target words were drawn from a database of 437 picturable French words: each target word had one to six competitors (phonological minimal pairs, along five types of phonological contrast, namely place of articulation, voice, manner of articulation, nasality/cluster and vowel quality) within the database and possibly more in the French lexicon. For each session/speaker, a randomization procedure randomly extracted 15 times a target and a specific color-shape combination (corresponding to a particular location on the grid). The participant, but not the experimenter, saw each target picture and corresponding

written word on the appropriate grid location on the computer screen and gave instructions to the experimenter, who had to write each target word on a paper grid.

The final *intelligibility score* of the participant was computed based on examination of the paper grid, as the number of test words not understood correctly by the experimenter during the interaction (independently of the location on the grid). A rating of 1 was given to incorrect responses, and 0.5 when two responses (including the correct response, e.g., "desert"/"dessert" for "dessert") were provided.

Articulation

Articulatory precision was assessed on the production of a set of 50 pseudo-words, covering the articulation of most of the French consonants and vowels as well as consonant clusters.

Pseudoword production was elicited via the bimodal presentation of each target (in an orthographic form on a screen and in an audio form via headphones), to minimise errors due to pseudo-word reading difficulties or to misperception. Regarding the assessment of articulation accuracy, errors on targeted consonants, vowel and syllables were listed based on perceptual assessment following a guided coding procedure for each pseudo-word, where questions were targeted on specific phoneme and on potential types of errors (distortion, substitution, omission, insertion). The computerized tool allowed the raters to play each pseudoword as needed and score targeted phonemes or syllables as correct or incorrect. Overall, 151 targeted phonemes or syllables included in the 50 pseudowords were scored per speaker and a descriptor related to articulation accuracy is expressed in terms of a *number of errors* (from 0 to a maximum of 151). The pseudo-word production is used as a screening for articulatory precision in *MonPaGe-2.0.* as only the total number of errors is considered and can be easily extracted via the guided computerized scoring tool, but it has been conceived to be used also for further more detailed assessments.

Maximum phonation time (MPT)

As a standard measure of pneumo-phonatory control, a descriptor representing maximum phonation time over a sustained vowel was computed. Participants were instructed to produce a sustained vowel/a/as long as possible after taking a maximal inhalation, at a comfortable pitch and at their habitual loudness. The task was repeated as many times as needed and two productions were recorded. The duration of the two trials of the sustained/a/production was measured with Praat and the best performance selected as the *maximum phonation time (MPT).*

Voice

Voice-related measures were based on a sustained production for 2–3 seconds of the vowel/a/at a comfortable height and loudness and on the reading of a 7-syllable sentence composed of only voiced sounds ("Mélanie vend du lilas" – [melanivɑ̃dylila], '*Melanie sells lilac*'). All measurements were computed with Praat, using a semi-automatic procedure: a customized Praat script was developed to guide the non-expert user so that s/he can adjust the relevant settings for optimal results at each stage of the procedure.

A first set of standard descriptors of voice quality was taken on the first 2 seconds of the sustained/a/vowel. These included the two short-term (cycle-to-cycle) measures of vocal instability in terms of frequency and amplitude: jitter and shimmer, respectively. These were

computed with Praat as the 5-point Period Perturbation Quotient (*a_Jitter-PPQ5*) and the 11-point Amplitude Perturbation Quotient (*a_Shimmer-APQ11*). Instability in vocal fold vibration over the whole 2-second window was assessed in terms of f0 standard deviation (*a_SDf0*), potentially allowing for the detection of vocal tremor. Presence of a noise component during the vowel was measured in terms of a harmonic-to-noise ratio (*a_HNR*) and possible dysphonia was also assessed with the smoothed cepstral peak prominence measure (*a_CPPs*) (Hillenbrand et al., 1994).

A second set of descriptors related to the speaking voice was measured on the longer continuous read speech sample "Mélanie vend du lilas". On the f0 time series computed over the whole sentence, mean (*speaking_meanf0*) and standard deviation (*speaking_SDf0*) were computed. The smoothed cepstral peak prominence was also taken from the whole voiced sentence (*speaking_CPPs*).

Prosodic contrast

The production of an assertive vs. interrogative prosodic contrast was tested on a four-syllable fully voiced sentence 'Laurie l'a lu' ([loʁilaly], *'Laurie read it'*). The sentence was first presented on the screen as a declarative and read by the speaker. Then the speaker was asked to say the same sentence again but asking a question, while the sentence was presented with a question mark.

The prosodic contrast between the two modalities was computed in terms of a difference in f0 modulation over the sentence. In the interrogative condition marked by a large final rise of f0 in French, expected a large delta in f0 range between the first half and the second half of the sentence is expected. This ambitus was expected to be smaller in the assertive condition, although it was not inexistent if the speaker produces a large f0 fall at the end of the assertion, or an initial rise at the beginning of it. Therefore, the prosodic contrast was expressed as the difference between the two modalities in these f0-range deltas between the beginning and the end of the sentence. The recorded sentences were automatically split into two parts, and f0 range was computed in semitones for the first and second parts. The descriptor for prosodic contrast achievement was thus computed as:

$$(\text{f0range@end-f0range@begin})_{\text{question}} - (\text{f0range@end-f0range@begin})_{\text{assertion}}$$

Speech rate

Speech rate was assessed on a short sentence reading task. The duration of the sentence "Mélanie vend du lilas" ([melanivɑ̃dylila], 'Melanie sells lilac') was measured with Praat and a simple measure of speech rate was obtained by dividing the expected number of phonemes (14 phonemes) by this duration.

Diadochokinetic (DDK) rate

Maximum repetition rates with oral *diadochokinetic* tasks are often used in clinical practice to test the ability to make alternating articulatory movements in quick and accurate succession. Seven items, which vary in terms of phonological complexity, were used here. They included standard sequences used to compute alternative motion rate (AMR) with the repetition of a CV syllable (AMR_{CV}) or a CCV syllable (AMR_{CCV}). Different CV and CCV syllables were used to target alternating movements with different articulators: jaw/lips with/ba/, front part of the tongue with/de/, tongue body with/go/, constrictions with/kla/, and front to back with/tʁa/. Finally, a repetitive sequence/badego/was used to compute

a sequential motion rate (SMR_{CV}). Participants were instructed to produce these sequences in a continuous manner for at least five second as fast and as accurately as possible.

The number of phonemes produced over an interval of about 4 seconds of continuous repetition was used as an index of DDK rate for AMR_{CV}, AMR_{CCV}, SMR_{CV}. This interval was selected from the onset of the speech waveform and was manually adjusted to the right in order not to cut the last syllable if needed. Also, in order to capture difficulties in the repetition of the same syllables (AMR) vs. the repetition of a sequence of three syllable (SMR) which could be found for speakers with AoS for instance (e.g., Ziegler, 2002), we also computed the difference between the sequential motion rate and the alternative motion rate averaged over all CV sequences: (SMR_{CV}- AMR_{CV}).

Procedure and analyses

All the participants underwent the MonPaGe protocol following the standard assessment procedure with speech and speech-language pathologists in clinical settings for the patients and in a standard room at the University for the controls. The MonPaGe protocol was run on a laptop and speech samples were recorded using either a head-mounted or a table microphone depending on the place of recruitment. This variability in recording reflects the variations across clinical settings and has also been introduced in the recordings of the reference group (see Fougeron et al., 2018). The assessment takes about 20 minutes with typical speakers and 30 minutes with patients with moderate MSD (Pernon et al., 2020), and the same duration is necessary for scoring.

Constitution of the norms from the reference healthy speakers group

The productions of the 404 healthy speakers reference group were analyzed in order to obtain normative values. For each descriptor, missing data and extreme values (below and above the 1^{st}/99^{th} percentiles) were first removed. In order to account for sex or age-group effects on the norms, linear models were first computed in R with Sex (male/female) and Age-group ([20–59], [60–74], [75+]) as between-subject factors and the various descriptors as dependent variables. Whenever an effect of Sex or an interaction with Age-group was found, further analyses were carried out for the male and female speakers separately. Inter-group differences for the Age-group predictor were tested with Tukey post-hoc tests. Results from these statistical analyses are provided in the supplementary material. Normative values for each descriptor was then computed by subgroups accordingly. Since groups split or merging were done according to the sex or age effects, the number of participants retained for the computation of normative values varied according to the descriptor considered.

Computation of deviance scores

For each descriptor, the speech of the participants in the two test groups (patients and healthy controls) was compared to the normative values of the reference population, according to its sex and age group.

A ***deviance score* (*DevS*)** spanning from 0 to 4 was defined for each descriptor depending on its distance from the reference value. The standard limit from the normative data was fixed at centile 5 (Brooks et al., 2011) and five degrees of *DevS* are defined, with *DevS* = 0 standing for no-deviant (within normal range) and *DevS* = 4 for excessively deviant,

according to the position of the speaker's descriptor value relative to the reference distribution. Cut-off values for defining these deviance degrees are determined according to either the tails of the distribution (c5/c95 and c1/c99, for the *DevS* = 1 and *DevS* = 2, respectively) and further severity, beyond c1/c99, is computed based on the inter-centile distance between c50 and c5/c95, as detailed in Table 2.

For the dimensions of voice quality and DDK performance, which rely on several descriptors, a composite deviance score was further computed as detailed in Table 3. Finally, each participant's speech is globally characterized by the *MonPaGe total score*

Table 2. Calculation of deviance scores (*DevS*) used in the MonPaGe screening protocol (c = centiles, SD = standard deviations from the reference distribution).

Descriptive severity	Quantitative distance from the normative data	*DevS*
Within normal range	> c5*	0
Mildly deviant	< c5 and ≥ c1*	1
Moderately deviant	< c1 and ≥ 1.5*(C50-C5)* or at the maximum error score**	2
Severely deviant	>1.5*(C50-C5) and ≤ 2*(C50-C5)* or beyond the maximum error score**	3
Excessively deviant	> 2*(C50-C5)* or well beyond the maximum error score (Intelligibility: > 7 errors; articulatory accuracy: > 20 errors)	4

*Or on the other end (< c95, etc.) depending on the direction of the measure.
**error scores for intelligibility and articulatory accuracy, see measures in text and Table 3.

Table 3. Descriptors and deviant scores for each of the seven speech dimensions.

Dimension	Descriptors	*DevS*	For the computation of the MonPaGe *TotalDevS*
(1) Voice	Sustained/a/: **a1**./a/F0 SD **b**./a/Jitter **c**./a/Shimmer **d1**./a/CPPS Sentence: **a2**. Speaking F0 SD **d2**. Speaking CPPs	Individual *DevS* (0 to 4) for each of the 6 voice measures	Composite Voice *DevS* (0 to 6) = ½*[max *DevS* (a1, a2) + max *DevS* (b, c) + max *DevS* (d1,d2)]
(2) Intelligibility	Number of errors (max 15)	*DevS* (0 to 4)	*DevS* (0 to 4)
(3) Articulatory accuracy	Number of segmental errors (max 151)	*DevS* (0 to 4)	*DevS* (0 to 4)
(4) Maximum phonation time	Max MPT (sec)	*DevS* (0 to 4)	*DevS* (0 to 4)
(5) Speech rate	Speech rate (phonemes/sec)	*DevS* (0 to 4)	*DevS* (0 to 4)
(6) Prosodic contrast	Interrogative minus assertive difference in f0 modulation (Hz)	*DevS* (0 to 4)	*DevS* (0 to 4)
(7) DDK rate	A. CCV AMR B. CV AMR C. CV SMR D. CV SMR – CV AMR difference (=C-B) (phonemes/sec)	Individual *DevS* (0 to 4) for each of the 4 DDK measures	Composite DDK *DevS* (0 to 6): = ½*[*DevS* (A) + maximum *DevS* (B,C) + *DevS* D)]
MonPaGeTotalDevS			Sum of *DevS* from the 7 dimensions (0 to 32)

(*TotalDevS*, ranging from 0 to 32); which corresponds to the sum of the deviant scores on the seven speech dimensions, where larger scores indicate more deviance in speech.

The detailed procedures for the assessment of sensitivity, specificity and external validity are described along with the results.

Results

Reference values

The normative values for each descriptor obtained from the 404 healthy speakers reference group are presented in Appendix A (Tables A.I and A.II). Normative values for each descriptor and for each age/sex subgroup are presented in terms of mean, median and standard deviation, except for intelligibility and segmental errors for which median and maximum error rates are provided. Distribution tails of the population's scores are also given with the $1^{st}/99^{th}$, $5^{th}/95^{th}$ and $10^{th}/90^{th}$ percentiles, in order to provide possible cut-off scores of the normal performance limits. According to the descriptor, the tails of interest for the determination of the cut-off scores are either on the upper or lower sides (e.g., for error values, alteration is to be found in the upper tail while 0 error is normal).

Specificity and sensitivity of MonPaGe

Appendix B and C present the deviant scores per speech dimension and *TotalDevS* for each of the 80 patients and 62 healthy test controls, respectively. The *MonPaGe TotalDevS* ranges from 0 to 12 in the patient group and from 0 to 4 in the healthy test control group.

The cut-off score for a diagnosis of MSD with MonPaGe was defined by plotting sensitivity versus specificity for the few possible cut-off scores (from 1 to the maximum *MonPaGeTotalDevS* in the test control group). The best cut-off score was a ***MonPaGeTotalDevS* >2** with a *specificity* of 95.2% and *sensitivity* of 83.8%. Figure 1 illustrates how sensitivity varies according to patients' severity as assessed externally with the Composite Perceptual Score. It can be observed that sensitivity raises to 92% when excluding patients with very mild impairments (CP score ≤4/20).

For the *external validity*, the composite perceptual score of severity was considered the clinical gold standard. Pearson correlation computed between the *MonPaGe TotalDevS* and the external CP score on the 80 patients is $r(80) = .737$, $p < .0001$.

Discussion

The sensitivity and specificity in the diagnosis of MSD of the MonPaGe screening protocol based on a set of semi-automated acoustic measures and on targeted perceptual measures was assessed on the speech from 80 French-speaking patients diagnosed with a diversity of mild to moderate MSD and 62 healthy test controls against the normative data of 404 neurotypical adult speakers. On each of the seven speech dimensions (voice, speech rate, articulation, prosody, pneumophonatory control, diadochokinetic rate, intelligibility) a deviance score (*DevS*) was calculated for each participant with reference to the normative data and an overall deviance score, the *MonPaGe TotalDevS*, was computed as the sum of all *DevS*. With a cut-off score at *MonPaGe TotalDevS* > *2*, the *specificity* of the tool was

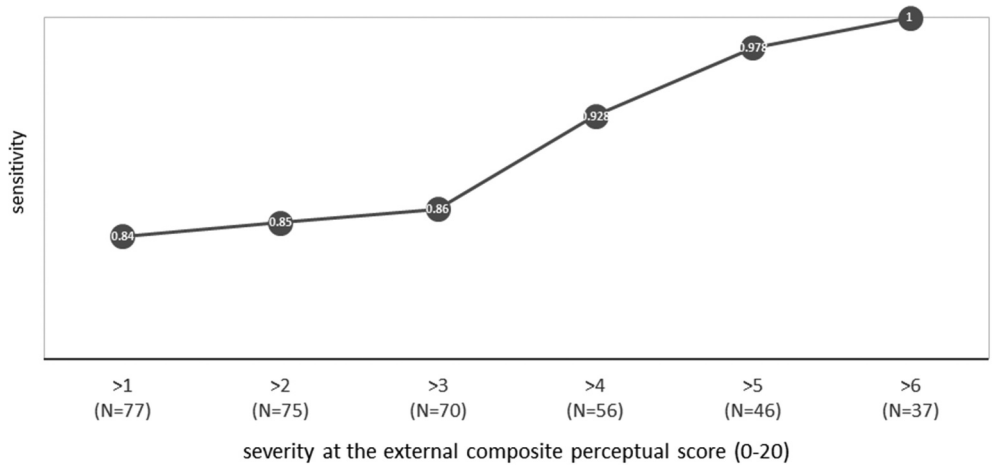


Figure 1. Sensitivity of the MonPaGe screening protocol, with a *MonPaGeTotalDevS* cut-off score at >2, according to severity level of the population assessed with the external composite perceptual score, ranging 0 to 20 (where 1 to 6 is considered mildly impaired, 7 to 13 moderate, 14 to 16 severe and >16 very severe).

excellent (95%), and its sensitivity was very high, in particular for the mild to moderate patients (sensitivity >92% when the patients with very mild MSD are excluded). Along with the strong correlation between the *MonPaGe TotalDevS* and an external composite perceptual assessment of severity obtained from six expert judges, the present results indicate that the MonPaGe protocol is a reliable screening tool for assessing the presence of MSD and its severity.

The current study has two crucial features related to the specificities of the analyses in the MonPaGe protocol and to the population involved in the study, which will be discussed in further detail below.

Guided objective analyses of impaired speech

As exposed in the introduction, the available clinical assessment tools for MSD are mostly based on perceptual judgment, which are bound to rely heavily on subjective evaluation. For these reasons the MonPaGe assessment protocol relies mainly on descriptors extracted from the acoustic signal for five out of the seven dimensions (voice, MPT, speech rate, prosodic contrast and DDK rate). For the other two dimensions (intelligibility and articulation accuracy), the descriptors are not objective in the sense of 'acoustic', but are not purely subjective either. In the functional intelligibility testing, the clinician writes down his/her understanding of the word pronounced by the patient and the matching with the intended word produced by the patient is done *off-line* during the scoring procedure, a procedure which has been used previously to assess intelligibility (Kent, 1992; Miller, 2013; Weismer, 2008; Yorkston & Beukelman, 1978). Here the intervention of a human listener is needed since the very notion of intelligibility refers to how adequately the intended targets are actually perceived by an interlocutor. In the assessment of articulation accuracy, the scoring of segmental errors on pseudo-words is essentially relying on perceptual judgments. It

should be noticed however that this assessment is also done off-line: the clinician follows a guided coding procedure for each pseudo-word, where questions are targeted on specific phonemes and on potential types of distortions. To answer these questions, the clinician/scorer can listen to each production as many times as necessary. Inter-rater agreement was evaluated for pseudo-word scoring, as this measure involves some amount of subjective assessment. Two different raters (an expert and a recently qualified SLP) independently scored the 151 target phoneme/syllables in the pseudo-word module for 20 patients (5 with PD, 5 ALS, 5 WD and 5 AoS). The mean inter-rater agreement on the 3020 observations was very high (98.94%, Cohens' Kappa = 0.89). With this off-line scoring procedure, we thus obtained a very high inter-rater agreement between two clinicians on a significant subpart of the validation data.

Overall, the MonPaGe protocol takes advantage of objective descriptors of impaired speech parameters obtained via semi-automatic acoustic analyses or via targeted/guided coding of the intended production. Acoustic analysis has the advantage of releasing the clinicians from subjective descriptions, but it has the drawback of being time-consuming and relying on specific expertise. An assessment tool combining perceptual and easy-to-obtain acoustic information seems to represent a good balance taking advantage of the two approaches while minimizing their respective drawbacks. For *MonPaGe-2.0.s*, a semi-automatic acoustic analyses routine was developed to be easily performed with minimal intervention and minimal acoustic-phonetic knowledge. These interventions typically required to check the automatic segmentation of the onsets-offsets of analysis windows in the audio files, and to adjust some key parameters in the case of noisy recordings. Then, the automatic extraction of acoustic measures on the various speech parameters, followed by their automated comparison with normative data defined on a large reference population (404 speakers) free the clinicians from the many pitfalls of a heuristic, subjective approach to clinical evaluation. To date, the *MonPaGe-2.0.s* screening is based on a relatively limited set of well-defined acoustic descriptors and is quite performant.

Objective measures should also facilitate a fast and effective communication among clinicians and should ensure a good/quality screening of MSD even when done by clinicians without expert competence in acoustics and phonetics.

Validation on a variety of MSD

The few available MSD assessment tools often lack a validation procedure, or they have been validated on very specific subtypes of MSD (e.g., progressive AoS vs. progressive aphasia in Strand et al., 2014; dysarthria associated with Parkinson disease in Cardoso et al., 2017) or on cohorts of patients with a large variety of dysarthria subtypes but without a grouping approach (De Biagi et al., 2018; Knuijt et al., 2017). In the current study, the MonPaGe screening tool has been validated on a group of 80 patients encompassing a large diversity of types of MSD, including AoS following stroke, dysarthria associated to PD, to ALS, to FA, to WD and to Kennedy disease. Thus, as a screening tool, the MonPaGe protocol presents overall good performance in the diagnosis of the presence and severity of impaired speech for a variety of patterns of MSD. There are however some speech parameters which need to be improved in future implementations of MonPaGe. First, 25 patients with MSD displayed impaired intelligibility scores, but some patients with severe impairment of articulation displayed

intelligibility scores within the normal range. This observation may be related to the difficulty to measure intelligibility and call for further improvement, although the relationship between intelligibility and articulatory errors is not necessarily linear as articulatory errors may be predicted by the listener and therefore comprehensible (Coppens-Hofman et al., 2016). Second, prosodic contrasts and maximum phonation time scores were deviant in very few patients. This may be due to low sensitivity of these specific tasks or to the fact that these parameters were within normal range in our group of patients with mild to moderate MSD and should be specifically clarified in future. Finally, a further step will also be to achieve differential diagnosis for different subtypes of MSD by enriching the screening tool with specific speech descriptors and possibly additional speech tasks/materials.

Acknowledgments

This research was partly supported by Swiss National Science Foundation grants no. CRSII5_173711, by the Belgian FNRS grant 33663575 (CDR UltraAeroSpeech) and the program "Investissements d'Avenir" ANR-10-LABX-0083 (Labex EFL). The work was also possible thanks to funding from the French Friedreich Ataxia Association, and from Parkinson Suisse. The normative data could not have been constituted without the help of (a) other collaborators of the MonPaGe project: Pellet P., Bagou O., Kocjancic-Antolik T. and (b) our students: Blanchard C., Baradat Garcia A., Cohen T., Davat H., de Braquillanges A, de Laubier C., Dehon S., Leclerc-Poudrier A., Mabut C., Maysounave M., Miny A., Pochat-Baron M., Rary E., Vidou C., Van Malderen S., Visentini A.

Funding

This work was supported by the Swiss National Science Foundation - SNSF - [CRSII5_173711]; the Agence Nationale de la Recherche [ANR-10-LABX-0083]; Fonds De La Recherche Scientifique - FNRS [33663575].

ORCID

Marina Laganaro http://orcid.org/0000-0002-4054-0939

Declaration of interest statement

The authors report no conflict of interest.

References

Auzou, P., & Rolland-Monnoury, V. (2019). *BECD: Batterie d'Evaluation Clinique de la Dysarthrie*. Ortho-Edition.

Blumstein, S. (1990). Phonological deficits in aphasia: Theoretical perspectives. In A. Caramazza (Ed.), *Cognitive neuropsychology and neurolinguistics* (pp. 33–53). Lawrence Erlbaum.

Boersma, P., & David Weenink, D. (2018). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.37. http://www.praat.org/

Brooks, L. B., Sherman, E. M. S., Iverson, G. L., Slick, D. J., & Strauss, E. (2011). Psychometric foundations for the interpretation of neuropsychological test results. In M. R. Schoenberg & J. G. Scott (Eds.), *The little black book of neuropsychology*(pp. 893–922). New York: Springer.

Bunton, K., Kent, R. D., Duffy, J. R., Rosenbek, J. C., & Kent, J. F. (2007). Listener agreement for auditory-perceptual ratings of dysarthria. *Journal of Speech, Language, and Hearing Research*, *50* (6), 1481–1495. https://doi.org/10.1044/1092-4388(2007/102)

Cardoso, R., Guimarães, I., Santos, H., Loureiro, R., Domingos, J., de Abreu, D., Gonçalves, N., Pinto, S., & Ferreira, J. (2017). Frenchay dysarthria assessment (FDA-2) in Parkinson's disease: Cross-cultural adaptation and psychometric properties of the European Portuguese version. *Journal of Neurology*, *264*(1), 21–31. https://doi.org/10.1007/s00415-016-8298-6

Chicherio, C., Genoud-Prachex, T., Assal, F., & Laganaro, M. (2019). *E-GeBAS: Electronic Geneva bedside aphasia scale*. [Computer program]. https://www.unige.ch/fapse/logotools/fr/adultes

Code, C. (1998). Major review: Models, theories and heuristics in apraxia of speech. *Clinical Linguistics & Phonetics*, *12*(1), 47–65. https://doi.org/10.3109/02699209808985212

Coppens-Hofman, M. C., Terband, H., Snik, A. F., & Maassen, B. A. (2016). Speech characteristics and intelligibility in adults with mild and moderate intellectual disabilities. *Folia Phoniatrica Et Logopaedica*, *68*(4), 175–182. https://doi.org/10.1159/000450548

Dabul, B. (2000). *Apraxia battery for adults* (2nd ed.). Pro-Ed.

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

De Biagi, F., Frigo, A. C., Andrea, T., Sara, N., Berta, G., & Meneghello, F. (2018). Italian validation of a test to assess dysarthria in neurologic patients: A cross-sectional pilot study. *Otolaryngology*, *8* (1), 343. https://doi.org/10.4172/2161-119X.1000343

Duffy, J. (2005). *Motor speech disorders*. Mosby Elsevier.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). Elsevier Mosby.

Enderby, P., & Palmer, R. (2008). *FDA-2: Frenchay dysarthria assessment* (2nd ed.). PRO-ED.

Feiken, J., Hofstede, G., & Jonkers, R. (2008). De diagnostiek van verbale apraxia [The diagnosis of verbal apraxia]. *Logopedie En Foniatrie*, *80*(7/8), 228–234.

Feiken, J., & Jonkers, R. (2012). *Diagnostisch instrument voor apraxie van de spraak* [DIAS; Diagnostic instrument for apraxia of speech]. Bohn, Stafleu en Van Loghum.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189.198. https://doi.org/10.1016/0022-3956(75)90026-6

Fougeron, C., Delvaux, V., Pernon, M., Lévêque, N., Borel, S., Pellet, P., Bagou, O., Trouville, R., Ménard, L., Catalano, S., Lopez, U., Kocjancic-Antolik, T., & Laganaro, M. (2016). MonPaGe: Un protocole informatisé d'évaluation de la parole pathologique en langue française. In N. Joyeux & S. Topouzkhanian (Eds..), *Actes du colloque UNADREO Orthophonie et technologies innovantes* (Chapter 14, pp. 1–13). Isbergues: Orthoedition.

Fougeron, C., Delvaux, V., Ménard, L., & Laganaro, M. (2018). The MonPaGe_HA database for the documentation of spoken French throughout adulthood. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), L18-1679*.

Ghio, A., Giusti, L., Blanc, E., & Pinto, S. (2019). French adaptation of the "Frenchay dysarthria assessment 2" speech intelligibility test. *European Annals of Otorhinolaryngology, Head and Neck Diseases*, *137*(2), 111–116. https://doi.org/10.1016/j.anorl.2019.10.007

Haley, K. L., Jacks, A., de Riesthal, M., Abou-Khalil, R., & Roth, H. L. (2012). Toward a quantitative basis for assessment and diagnosis of apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *55*(Suppl), S1502–S1517.

Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech Language and Hearing Research*, *37*(4), 769–778. https://doi.org/10.1044/jshr.3704.769

Kent, R. D. (Ed.). (1992). *Intelligibility in speech disorders: Theory, measurement and management* (Vol. 1). John Benjamins Publishing.

Knuijt, S., Kalf, J. G., van Engelen, B. G. M., de Swart, B. J. M., & Geurts, A. C. H. (2017). The Radboud dysarthria assessment: Development and clinimetric evaluation. *Folia Phoniatrica Et Logopaedica*, *69*(4), 143–153. https://doi.org/10.1159/000484556

Maryn, Y., De Bodt, M., Barsties, B., & Roy, N. (2014). The value of the acoustic voice quality index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology*, *271*(6), 1609–1619. https://doi.org/10.1007/s00405-013-2730-7

Maryn, Y., De Bodt, M., & Roy, N. (2010). The acoustic voice quality index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, *43*(3), 161–174. https://doi.org/10.1016/j.jcomdis.2009.12.004

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (2008). Apraxia of speech: Definition and differential diagnosis. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (2nd ed., pp. 249–268). Thieme.

Menin-Sicard, A., & Sicard, E. (2020). Outils de mesures et d'évaluation objective de la phonologie, de la fluence et de la prosodie avec DIADOLAB3. *Rééducation orthophonique*, *281*, 131–168.

Miller, N. (2013). Measuring up to speech intelligibility. *International Journal of Language & Communication Disorders*, *48*(6), 601–612. https://doi.org/10.1111/1460-6984.12061

Pernon, M., Levêque, N., Delvaux, V., Assal, F., Borel, S., Fougeron, C., Trouville, R., & Laganaro, M. (2020). MonPaGe, un outil de screening francophone informatisé d'évaluation perceptive et acoustique des troubles moteurs de la parole (dysarthries, apraxie de la parole). *Rééducation Orthophonique*, *281*, 171–197.

Solomon, N. P., Garlitz, S. J., & Milbrath, R. L. (2000). Respiratory and laryngeal contributions to maximum phonation duration. *Journal of voice, 14*(3), 331–340.

Strand, E. A., Duffy, J. R., Clark, H. M., & Josephs, K. (2014). The apraxia of speech rating scale: A tool for diagnosis and description of apraxia of speech. *Journal of Communication Disorders*, *51*, 43–50. https://doi.org/10.1016/j.jcomdis.2014.06.008

Varley, R., & Whiteside, S. (2001). What is the underlying impairment in acquired apraxia of speech. *Aphasiology*, *15*(1), 39–49. https://doi.org/10.1080/02687040042000115

Weismer, G. (2008). Speech intelligibility. In M. J. Ball, M. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 568–582). Blackwell.

Wertz, R., LaPointe, L., & Rosenbek, J. (1984). *Apraxia of speech in adults. The disorder and its management*. Grune and Stratton.

Wuyts, L. R., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, K., Raes, J., & Van de Heyning, P. H. (2000). The dysphonia severity index: An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research*, *43*(3), 796–809. https://doi.org/10.1044/jslhr.4303.796

Yorkston, K. M., & Beukelman, D. R. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders*, *11*(6), 499–512. https://doi.org/10.1016/0021-9924(78)90024-2

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. https://doi.org/10.1006/brln.2001.2614

Ziegler, W. (2008). Apraxia of speech. In G. Goldenberg & B. L. Miller (Eds.), *Neuropsychology and Behavioral Neurology* (3rd ed., pp. 269–286). Elsevier.

Ziegler, W. (2009). Modelling the architecture of phonetic plans: Evidence from apraxia of speech. *Language and Cognitive Processes*, *24*(5), 631–661. https://doi.org/10.1080/01690960802327989

Ziegler, W., Aichert, I., & Staiger, A. (2012). Apraxia of speech: Concepts and controversies. *Journal of Speech, Language, and Hearing Research*, *55*(5), S1485–S1501. https://doi.org/10.1044/1092-4388(2012/12-0128)

Ziegler, W., Schölderle, T., Staiger, A., & Vogel, M. (2015). The Bogenhausen dysarthria scales (BoDyS): A standardized test for the clinical assessment of dysarthria in adults. *Sprache Stimme Gehör*, *39*(4), 171–175. https://doi.org/10.1055/s-0041-102792

Ziegler, W., Staiger, A., Schölderle, T., & Vogel, M. (2017). Gauging the auditory dimensions of dysarthric impairment: Reliability and construct validity of the Bogenhausen dysarthria scales (BoDyS). *Journal of Speech, Language, and Hearing Research*, *60*(6), 1516–1534. https://doi.org/10.1044/2017_JSLHR-S-16-0336