Contents lists available at ScienceDirect

EClinicalMedicine

Research Paper

The effects of intensive speech treatment on intelligibility in Parkinson's disease: A randomised controlled trial

Erika S. Levy[a,*], Gemma Moya-Galé[b], Young Hwa M. Chang[a], Katherine Freeman[c], Karen Forrest[d], Mitchell F. Brin[e,f], Lorraine A. Ramig[a,g,h,i]

[a] Department of Biobehavioral Sciences, Teachers College, Columbia University, New York, NY, United States
[b] Department of Communication Sciences and Disorders, Long Island University, Brooklyn, NY, United States
[c] Department of Biomedical Sciences, Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, FL, United States
[d] Department of Speech and Hearing Sciences, Indiana University, Bloomington, IN, United States
[e] Department of Neurology, University of California, Irvine, Allergan LLC, Irvine, CA, United States
[f] Research and Development, Allergan, an AbbVie Company, Irvine, CA, United States
[g] National Center for Voice and Speech, Denver, CO, United States
[h] Department of Speech-Language and Hearing Sciences, University of Colorado, Boulder, CO, United States
[i] LSVT Global, Inc., Tucson, AZ, United States

ARTICLE INFO

ABSTRACT

Background: More than 6,000,000 individuals worldwide are diagnosed with Parkinson's disease (PD). Nearly 90% develop speech signs that may substantially impair their speech intelligibility, resulting in losses in their communication and quality of life. Benefits of intensive speech treatment have been documented for a range of speech signs. However, the critical question of whether speech is more intelligible after treatment has not been investigated in a randomised controlled trial (RCT). We hypothesised that intensive speech treatment would improve speech intelligibility in PD.

Method: Sixty-four patients with hypokinetic dysarthria secondary to PD participated in this single-centre, parallel arm, statistically-powered RCT. Reporting follows CONSORT guidelines for non-pharmacological treatment. Patients were recruited from US clinics and randomised using a statistician-derived minimisation algorithm, to intensive speech treatment (16 1-hour sessions/1 month) targeting voice (voice group) or targeting articulation (articulation group) or to an untreated group (no treatment group). Speech treatments were delivered by speech clinicians who specialised in treating patients with PD. Trial design minimised bias and supported equipoise. For intelligibility assessment, blinded listeners (n = 117) orthographically transcribed 57 patients' recorded, self-generated narrative speech samples, randomly presented in multi-talker babble noise. Listeners were American-English speakers, ages 18−35 years, with normal hearing. The primary outcome was baseline (pre-treatment) to post-treatment change in transcription accuracy (TA), recognised as the most objective measure of intelligibility. TA was defined as the percentage of words transcribed correctly. Listeners, data collectors, and data managers were blinded to treatment conditions and groups. Reliability was evaluated using intraclass correlation coefficients and differences amongst groups were evaluated by mixed-effects models, in accordance with the intention-to-treat approach.

This trial was registered with ClinicalTrials.gov Identifier: NCT00123084.

Findings: Between June 23, 2016 and August 14, 2017, blinded listeners transcribed baseline and post-treatment speech samples for intelligibility assessment of 57 patients in the voice (n = 19), articulation (n = 19) and no treatment (n = 19) groups. Between-group differences (d) in changes from baseline to post-treatment in TA indicated significantly greater increases following treatment targeting voice than treatment targeting articulation (d = 26·2%, 95% CI 1·5 − 51·0; p = 0·04; ES=1·0). Differences between TA changes in the treatment targeting voice and in the no treatment group were significant (d = 42·8%, 95% CI 22·4 − 63·2; p = 0·0002; ES=1·8). Differences between TA changes in the treatment targeting articulation and in the no treatment group were not significant (d = 16·5%, 95% CI -6·1 − 39·2; p = 0·147; ES=0·9).

* Corresponding author.
    E-mail address: elevy@tc.columbia.edu (E.S. Levy).

Interpretation: These findings provide the first RCT evidence that intensive speech treatment targeting voice improves speech intelligibility in PD. Thus, this evidence-based treatment may positively impact health-related quality of life for patients with PD globally when it is included in patient management.

Research in Context

Evidence before this study

Using the key words "Parkinson's disease," "intelligibility," and "speech treatment," we searched the Cochrane Database and PubMed for relevant articles from Jan. 1, 1988 to May 15, 2020. Three randomised controlled trials (RCTs) on speech treatment in PD were identified, all of which revealed benefits of intensive speech treatment targeting voice. While these studies were the first to generate short- and long-term efficacy data for speech treatment in PD, they were primarily confined to acoustic and physiologic measures and patient and family reports; no RCT examined whether speech treatment improves speech intelligibility in PD.

Added value of this study

Results are presented from the first RCT documenting that intensive speech treatment targeting voice improves speech intelligibility in PD. Further scientific advances include: 1) two active speech treatment comparators and an inactive comparator to dissociate the effects of speech treatment target and dosage; 2) orthographic transcription accuracy data collected from blinded listeners (*n* = 117), implementing the gold standard and most objective measure of intelligibility; 3) inclusion of the dimension of patients' vocal loudness in intelligibility assessment for the first time in an RCT.

Implications of all the available evidence

Previous evidence indicates that intensive speech treatment targeting voice can make short- and long-term improvements across a range of variables in PD, including vocal loudness, vocal quality, articulation, intonation, swallowing disorders, and patient-reported communication effectiveness, consistent with neural findings (positron emission tomography and functional magnetic resonance imaging). Our findings demonstrate that intensive speech treatment targeting voice improves speech intelligibility in PD. Taken together with previously reported gains following treatment targeting voice, there is substantial evidence that this treatment results in meaningful improvements in communication. Data from implementation studies in "real world" clinical settings, including with treatment delivered via telemedicine, are consistent with these findings; thus, this evidence-based treatment may positively impact health-related quality of life for patients with PD globally when it is included in patient management.

1. Introduction

The ability to communicate is fundamental to quality of life. Speech intelligibility is essential to communicative success [1]. Of the more than 6,000,000 patients worldwide who are diagnosed with

Parkinson's Disease (PD) [2], nearly 90% develop speech signs that may reduce intelligibility [3,4]. Although less visible than the limb motor signs of PD (e.g., tremor, loss of balance), the speech signs in PD often impair patients' communication and reduce their quality of life [5,6]. Reductions in speech intelligibility in PD are characterised by decreased vocal loudness and imprecise articulation, among the salient signs of the motor speech disorder of hypokinetic dysarthria [7]. It has been reported that deficits in sensory feedback and internal cueing may also play a significant role in the speech disorder associated with PD [8]. These concomitant deficits may help explain why neither medical treatments (e.g., neuropharmacological or neurosurgical) nor traditional speech treatments, which focus on motor speech symptoms alone, have consistently or significantly improved the degenerating speech in PD [9–13].

An intensive behavioural speech treatment targeting voice, Lee Silverman Voice Treatment (LSVT LOUD®), has produced the first evidence of short- and long-term (two-year) efficacy of speech treatment in PD, as demonstrated in three randomised controlled trials (RCTs) [14–16]. This treatment is unlike traditional speech treatments in that it: 1) focuses on a single treatment target of voice, as opposed to utilising multiple targets, such as respiration, loudness, rate, and articulation, 2) follows principles that promote motor learning and activity-dependent neuroplasticity [17], including intensive dosage (16 1-hour sessions over one month), high effort treatment, and 3) retrains sensory feedback and internal cueing.

These RCTs show benefits of intensive speech treatment targeting voice across a range of speech variables, including objective acoustic measures of vocal loudness, articulation, intonation, and perceptual measures of voice quality and patients' self-reported improvements in communicative effectiveness [4,14-16,18]. These observations are consistent with neural findings (positron emission tomography and functional magnetic resonance imaging) following this intensive voice treatment [19–21]. In addition, a positive treatment impact on swallowing disorders (dysphagia), frequently comorbid with dysarthria in patients with PD, has been reported [22,23]. However, while the ability to communicate is fundamental to quality of life and speech intelligibility is essential to communicative success [1], to date there have been no RCT data to support the impact of speech treatment on intelligibility in this population.

We report the first RCT aimed to evaluate whether intensive speech treatment can improve intelligibility in PD. To dissociate the specific contributions of the intensive dosage of treatment and the target of treatment, our design included two active treatment comparators: intensive treatment targeting voice (voice group, LSVT LOUD) and intensive treatment targeting articulation (articulation group, LSVT ARTIC™) [16]. Articulation was selected as a treatment target to address the well-established articulatory deficits in PD [7]. Articulation disorders in PD have been treated, though not intensively, and with moderate success [7,24,25]. The intensive treatment targeting articulation was developed by the clinical research team to impact the articulatory subsystem, which has, in fact, been identified as the strongest contributor to speech intelligibility in PD dysarthria [26]. All other elements of the two treatments (e.g., dosage and effort) were matched (Table 1). Patients were also randomised to an inactive comparator (no treatment group). Our outcome measure was blinded listeners' transcription accuracy (TA), the gold standard and most objective measure of intelligibility [27].

Table 1

Comparison of treatment targeting voice (Voice) and treatment targeting articulation (Articulation) for patients with Parkinson's disease.

	Voice	Articulation
Focus of treatment	Loudness	Enunciation[*]
Dosage	Increased movement amplitude directed predominately to respiratory-laryngeal systems	Increased movement amplitude directed predominately to orofacial-articulatory system
	Individual treatment session of one hour, four consecutive days per week over a 4-week period	Individual treatment session of one hour, four consecutive days per week over a 4-week period
Effort	Push for maximum participant perceived effort	Push for maximum participant perceived effort
Daily Exercises		
Maximum sustained movements completing multiple repetitions of tasks, minutes 1−12	Sustain the vowel "ah" in a good quality, loud voice, for as long as possible.	Sustain articulatory placement for "p" (lips closed) and "t" (tongue tip behind upper teeth) with Iowa Oral Pressure Instrument (IOPI); hold for 4-seconds for each trial.
Directional movements completing multiple repetitions of tasks, minutes 13−23	Say the vowel "ah" in a good quality, loud voice gliding high in pitch; hold for 5 s.	Repeat as many as possible in 5 s trials, each of the following single consonants with precise articulation (voiceless productions): /p/ /t/ /k/
	Say the vowel "ah" in a good quality, loud voice gliding low in pitch; hold for 5 s.	Repeat as many as possible in 5 s trials, each of the following minimal pair combinations with precise articulation: /t-k/, /n-g/, "oo-ee" and "oo-ah"
Functional movements, minutes 24−30	Participant reads 10 self-generated phrases he/she says daily in functional living (e.g., "Good morning") using the same effort and loudness as he/she did during the maximum sustained movements exercise.	Participant reads 10 self-generated phrases he/she says daily in functional living (e.g., "Good morning") using same effort for enunciation as he/she did during the maximum sustained movements exercise.
Hierarchy Exercises, minutes 31−55		
Purpose	Train rescaled vocal loudness achieved in the Daily Exercises into context specific and variable speaking activities.	Train rescaled enunciation achieved in the Daily Exercises into context specific and variable speaking activities.
Method	Incorporate multiple repetitions of reading and conversation tasks with a focus on vocal loudness.	Incorporate multiple repetitions of reading and conversation tasks with a focus on enunciate.
Tasks	Tasks increase in length of utterance and difficulty across weeks, progressing from words to phrases to sentences to reading to conversation, and can be tailored to each participant's goals (e.g., communicate at work or with caregivers) and interests (e.g., speak on topics of golf, cooking).	Tasks increase in length of utterance and difficulty across weeks, progressing from words to phrases to sentences to reading to conversation, and can be tailored to each participant's goals (e.g., communicate at work or with caregivers) and interests (e.g., speak on topics of golf, cooking).
Assign Homework Exercises to be completed outside of the therapy room, minutes 56−60		
Duration and repetitions on treatment days (4 days/week)	Subset of the Daily Exercises and Hierarchy Exercises; 10 min, performed once per day	Subset of the Daily Exercises and Hierarchy Exercises; 10 min, performed once per day
Duration and repetitions on non-treatment days (3 days/week)	Subset of the Daily Exercises and Hierarchy Exercises; 15 min, performed twice per day	Subset of the Daily Exercises and Hierarchy Exercises; 15 min, performed twice per day
Conversational Carryover Assignment	Participant is to use the louder voice practiced in exercises in a real-world communication situation	Participant is to use enunciated speech practiced in exercises in a real-world communication situation
Difficulty level	Matched to the level of the hierarchy where the participant is in treatment	Matched to the level of the hierarchy where the participant is in treatment
Shaping Techniques		
Purpose and approach	Train vocal loudness that is healthy (i.e., no unwanted vocal strain) through use of modeling ("do what I do") or tactile/visual cues.	Train speech enunciation that is within normal limits (i.e., no excessive movements) through use of modeling ("do what I do") or tactile/visual cues.
Sensory Calibration	Focus attention on how it feels and sounds to talk with increased vocal loudness (self-monitoring) and to internally cue (self-generate) new loudness effort in speech.	Focus attention on how it feels and sounds to talk with increased enunciation (self-monitoring) and to internally cue (self-generate) new enunciation effort in speech.
Objective and subjective clinical data collected during each treatment session	Measures of duration, frequency, and sound pressure level Documentation of percentage of cueing required to implement vocal loudness strategy Observations of perceptual voice quality Participant's self-reported comments about successful use of the improved loudness in daily communication Participant self-reported perceived effort	Measures of oral pressure and precise articulatory productions Documentation of percentage of cueing required to implement enunciation strategy Observations of perceptual speech intelligibility Participant's self-reported comments about successful use of the improved enunciation in daily communication Participant self-reported perceived effort

[*] The instruction 'Enunciate' is used to train articulatory effort and 'Speak loud' is used to train healthy vocal effort.

Both therapies are standardized with respect to intensive dosage. Effort in treatment targeting voice and treatment targeting articulation are based on the participant's self-perceived effort during treatment tasks, on a scale of 1−10, with 10 being highest perceived effort.

It was hypothesised that both intensive speech treatments would improve speech intelligibility in PD. Intelligibility benefits from speech treatment targeting voice would stem from an increase in vocal loudness, as well as from the impact of increased amplitude and coordination of movement gained across speech production subsystems. This would result in improved articulation, intonation, and voice quality, as well as reduced speech rate; thus, increasing intelligibility [4,14−16,18,28]. In contrast, the speech treatment targeting articulation was expected to enhance intelligibility primarily by increasing articulatory displacement and reducing speech rate [29,30]. Given the relationship reported between articulation and speech intelligibility, adjustments made at the articulatory level might be greater in the articulation group than in the voice group, likely yielding particular benefits to intelligibility [26]. No intelligibility changes were hypothesised for the no treatment group. Findings from this RCT were expected to advance knowledge regarding the impact of intensive speech treatment on speech intelligibility in the growing population of patients with PD [31].

2. Methods

2.1. Study design

The current single-centre, parallel arm RCT, was reported following CONSORT guidelines for non-pharmacological treatment [32,33]. The study was designed to assess whether speech in patients with PD is more intelligible following intensive speech treatment targeting voice or targeting articulation (ClinicalTrials.gov Identifier: NCT00123084). Thus, this RCT included two active treatment comparators, permitting dissociation of the effects of speech treatment target (voice versus articulation) and speech treatment dosage. The inactive comparator, the no treatment group, evaluated on the same schedule as treated patients, represented the natural progression of the speech disorder in PD. A decision was made to not include a sham treatment so as to adhere to the principle of equipoise and to not place an undue burden of time and effort on the patients who would have received a treatment with little potential for a therapeutic effect [34,35]. Additionally, given questions regarding the appropriateness of shams as treatment comparators in behavioural research [35], in this study, the active comparators and the inactive comparator (no treatment group) were judged to provide more helpful contrasts than a sham group.

This study is considered a single-centre, parallel arm RCT because all patient speech sample data were collected at the National centre for Voice and Speech (NCVS), Denver, Colorado. As these speech samples were subsequently analysed at other centres, the study protocols were approved by the Institutional Review Boards in the United States at the University of Colorado Health Sciences centre Denver, the University of Colorado, Boulder, Columbia University, New York City (NYC), New York, and Indiana University (IU), Bloomington. Protocols are available at https://clinicaltrials.gov/ct2/show/NCT00123084 and from the corresponding author by request.

2.2. Participants

All patients had hypokinetic dysarthria due to PD. Patient speech samples were collected before and after treatment at NCVS, Denver, an internationally recognised clinical research facility affiliated with the University of Colorado, Boulder. Speech sample selection and piloting of speech intelligibility assessment occurred at the Speech Production and Perception Lab at Columbia University, NYC. Speech sample preparation and speech intelligibility assessment occurred at the Speech Acoustics Lab at IU, Bloomington. Both labs are internationally recognised for their research in perceptual and acoustic analysis of disordered speech. This multi-site collaboration allowed us to address our critical research question through expertise and resources at each site.

At NCVS, Denver, patients with PD were recruited from support groups, clinics, and physicians in the United States. Patients were eligible if they had a neurologist-obtained diagnosis of PD, showed clinical stability on their antiparkinsonian medication in the opinion of the referring neurologist, were within Hoehn and Yahr severity scale Stages I-IV [36], and were native speakers of American English, with mild, moderate or severe speech and voice disorder. Adults with atypical Parkinson symptoms at screening were excluded, as were those who had received intensive speech treatment in the past two years or had medical conditions such as moderate to severe dementia (Mini Mental Status Exam < 24/30) [37], untreated depression (Beck Depression Inventory II $\geq$ 5) [38], vocal fold pathology as diagnosed by an otolaryngologist, or any speech disorder or neurological condition that was unrelated to PD. Further information on recruitment, inclusion criteria, enrolment, and demographic and clinical characteristics, as well as sample size methodology, was reported by Ramig and colleagues [16], as this was part of a larger research project on speech production in PD.

Subsequently at IU, Bloomington, listeners were recruited to assess the speech intelligibility of the patient speech samples. Listeners were neurotypical young adults between the ages of 18 and 35 years who were recruited by flyers and website postings. Listeners were excluded if they failed a hearing screening at 20 dB HL at octave frequencies between 500 and 4000 Hz or were not native speakers of American English. A total of 57 patients and 117 listeners were included in the study. (See power analysis in Statistical analysis section below.) All patients and listeners provided written consent to participate and were compensated for their time [16].

2.3. Randomisation and masking

As summarised in Fig. 1, in this RCT, patients with PD generated speech samples at baseline and post-treatment. Subsequently, listeners performed intelligibility assessment of these samples. Randomisation and masking are described first with regard to patient speech samples and then with regard to listeners' intelligibility assessment.

At NCVS, Denver, randomisation of patients was performed by a statistician not otherwise involved in the study. A minimisation algorithm was derived using a 1:1:1 randomisation ratio to one-month voice or articulation or no treatment groups. Patients with PD were assigned based on demographic and clinical characteristics such as age, disease duration and severity, depression, and their respective weights. A minimisation algorithm was used as part of the randomisation process toward balancing baseline characteristics. The statistician generated a written allocation derived from the minimisation algorithm and forwarded it to an available treating clinician, who enroled the patient.

In keeping with informed consent procedures of the Institutional Review Board (IRB), patients were informed that they would be randomly placed in one of three groups, including possibly a no treatment group, with equal chance of being assigned to any group. Before randomisation, all patients were informed that if they were placed in the untreated group, they could receive complimentary treatment after the study was completed. As stated by van der Kolk and colleagues [39], "masking in non-pharmacological studies is virtually impossible because the intervention is obvious to those who receive it" (p. 1006). Nevertheless, all efforts were made to limit bias throughout the study. For example, the treatment names were never disclosed. Patients were not provided with information about the specific treatment tasks in other groups, nor did patients interact with the other groups. Patients' experimental speech samples were recorded following scripted protocols, by well-trained data collectors blinded to the treatment the patient received. Patients were not cued to modify their speech (e.g., to speak loudly or to enunciate) during
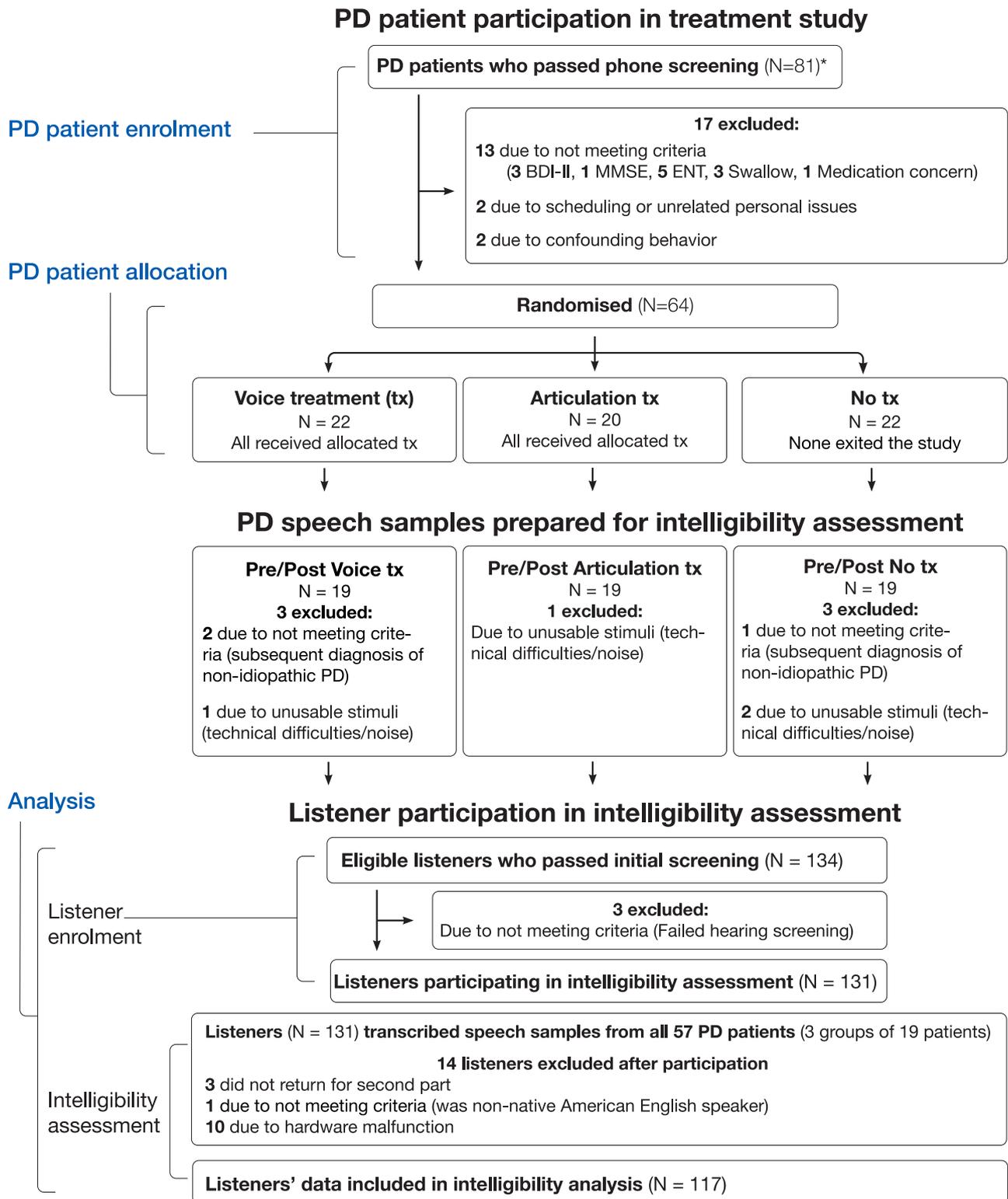
PD patient participation in treatment study

PD patients who passed phone screening (N=81)*

PD patient enrolment

17 excluded:
13 due to not meeting criteria
  (**3** BDI-II, **1** MMSE, **5** ENT, **3** Swallow, **1** Medication concern)
2 due to scheduling or unrelated personal issues
2 due to confounding behavior

PD patient allocation

Randomised (N=64)

| **Voice treatment (tx)**
N = 22
All received allocated tx | **Articulation tx**
N = 20
All received allocated tx | **No tx**
N = 22
None exited the study |

PD speech samples prepared for intelligibility assessment

| **Pre/Post Voice tx**
N = 19
3 excluded:
2 due to not meeting criteria (subsequent diagnosis of non-idiopathic PD)

1 due to unusable stimuli (technical difficulties/noise) | **Pre/Post Articulation tx**
N = 19
1 excluded:
Due to unusable stimuli (technical difficulties/noise) | **Pre/Post No tx**
N = 19
3 excluded:
1 due to not meeting criteria (subsequent diagnosis of non-idiopathic PD)

2 due to unusable stimuli (technical difficulties/noise) |

Listener participation in intelligibility assessment

Analysis

Listener enrolment

Eligible listeners who passed initial screening (N = 134)

3 excluded:
Due to not meeting criteria (Failed hearing screening)

Listeners participating in intelligibility assessment (N = 131)

Intelligibility assessment

Listeners (N = 131) **transcribed speech samples from all 57 PD patients** (3 groups of 19 patients)
14 listeners excluded after participation
3 did not return for second part
1 due to not meeting criteria (was non-native American English speaker)
10 due to hardware malfunction

Listeners' data included in intelligibility analysis (N = 117)

Fig. 1. Trial Profile. In this RCT, patients with PD generated speech samples. Subsequently, listeners performed intelligibility assessment of these speech samples. Data are summarized in the Results. Please see Ramig and colleagues' previous work for details on patients with PD [16]. PD=Parkinson's disease; BDI-II=Beck Depression Inventory-II; MMSE=Mini Mental Status Exam; ENT=Ear Nose Throat examination. *For details on PD patient enrolment, see Ramig and colleagues' previous work [16].

data collection. To maintain independence between treating clinicians and patient experimental data, clinicians were not involved in data collection of patients they treated. Speech treatments were delivered by three speech clinicians who specialised in treating patients with PD [16]. The treating clinicians were trained to avoid imparting bias when administering treatment, with an emphasis on delivering treatments without preference.

The treatments began one to two weeks after randomisation. Treatments followed established standardised protocols, and a primary investigator provided oversight in order to ensure fidelity of both treatments. Patients' adherence to interventions was assessed by their attendance at all 16 sessions and submission of daily homework and carryover checklists. In addition, on treatment days, patients and family members responded to clinicians' questions

regarding how other people responded to the patients' speech outside of the treatment setting (to assess generalisation of treatment effects). Finally, patients and family members completed rating scales at baseline and at post-treatment to document the impact of the treatment on the patients' functional communication [16]. Adherence was comparable across both treatment groups.

Both treatments were provided by all clinicians. Clinicians provided the same positive reinforcement during both treatments and conferred frequently to promote fidelity and consistency in delivery of each treatment. In post-treatment interviews, clinicians reported equal investment and belief in the effectiveness of both treatments, thus supporting equipoise. Moreover, when treated patients were asked whether they received the best treatment to which they could have been randomised, 95% of the patients in the articulation group and 100% of the patients in the voice group responded positively.

At IU, Bloomington, listeners performed intelligibility assessment of the speech samples generated by patients. All listeners heard all speech samples.

The design of this study was an assessor-blind RCT. Assessors were data collectors, listeners, and data managers. Data were coded such that all assessors at both locations were blinded to group assignment and treatment conditions. For example, at Columbia University, NYC, the baseline and post-treatment sentence stimuli were selected from the patients' recordings by a blinded investigator. At IU, Bloomington, addition of noise, speech intelligibility assessment and data management were performed by the blinded research team. Listeners were presented with patients' stimuli with no information provided regarding group assignment or treatment conditions.

2.4. Procedures

At NCVS, Denver, one of two intensive speech treatments (targeting voice or targeting articulation) was delivered to two groups of patients, and the third group received no treatment. As indicated in Table 1 (previously published) [16], both treatments were matched on all variables (e.g., dosage and effort) except for the single target of voice or articulation. Both treatments followed three principles: 1) a single training target (voice or articulation), aiming for greater amplitude of speech output. (A key element of treatment targeting voice involves training vocal loudness; in the articulation group, a key component is training articulation.); 2) an intensive dosage of treatment (16 1-hour sessions/1 month) with daily carryover activities and homework (10−15 min on treatment days, twice on non-treatment days); 3) retraining of sensory feedback and internal cueing. Further details of the treatments, clinicians, and control of bias have been published previously and are included in Table 1 [16].

At NCVS, Denver, patients' acoustic data (speech samples) were collected by research staff within one week prior to the onset of treatment (baseline) and within one week after termination of one month of treatment (post-treatment). For each patient, timing of data collection was designed to maintain consistency in relation to the patient's medication levels. The patients were seated in an Industrial Acoustics Company (IAC) sound-treated booth. A head-mounted AKG 420 condenser microphone was positioned 8 cm from the patient's lips [16,40]. The microphone was calibrated to a sound-level metre (Bruel and Kjaer 2239), and sound pressure level (SPL) was extracted with a reference distance of 30 cm [40].

Narrative speech was collected, rather than the commonly-examined read or repeated speech [4,41], to increase external validity by approximating daily self-generated spontaneous conversation [42]. Consistent with previous work [43], all patients generated a 90-second personal narrative monologue ("narrative speech") describing a time when they felt extremely happy. These narratives were originally elicited for a study on facial expression in PD [44], but were subsequently analysed for this intelligibility study. Because emotional

content affects speech production and intelligibility [45], the effects of emotional content were thereby controlled [43]. The audio recordings of the speech samples were then restored to their originally-produced SPLs using published methods [14−16,40].

At Columbia University, NYC, sentences from the speech sample recordings were selected for speech intelligibility assessment. One baseline and one post-treatment sentence were selected (by an investigator blinded to group assignment) from the recordings of each patient. The decision to include one sentence at each timepoint was necessary in order to implement TA, the most objective, consistent, and reliable measure of intelligibility [27]. In contrast to more subjective measures, such as listener intelligibility ratings (e.g., on a scale of 0−5), orthographic transcription requires the listener considerable time and effort to decode, process, remember, and type each utterance. Thus, the choice to include only one baseline and one post-treatment sentence from each patient (126 sentences, with reliability sentences included) rendered it feasible to collect high-quality, reliable data on all patients. Each sentence selected was the first complete phrase that was at least 18 s from the start of the speech sample and contained 4−11 words [46]. No significant differences in the number of words per sentence were found amongst groups at baseline ($p = 0.9663$) or at post-treatment ($p = 0.7246$). There were also no significant baseline to post-treatment changes in the number of words per sentence within groups ($p = 0.8085$).

At IU, Bloomington, speech sample preparation, and intelligibility assessment and data management were performed. The research team was blinded to treatment conditions. Using standard procedures [47], ten-talker babble noise from the AzBio sentences was added to the sentence stimuli [48]. Noise was added to approximate "real-world" background noise conditions and to allow controlled measurement of the impact of patients' reduced vocal loudness on speech intelligibility [48]. We maintained the same noise level in the baseline and post-treatment stimuli of each patient because environmental noise would not be expected to change post-treatment.

To prepare the stimuli, sentences were first edited with Praat software to eliminate silences before and after the utterance [49]. These edited sentences and babble noise were input to a custom MATLAB programme that measured the root mean square (RMS)-voltage of each baseline sentence [50]. The MATLAB code then adjusted the RMS-voltage of the babble noise to produce a baseline sentence embedded in noise with a 0 dB signal-to-noise ratio (SNR) [51]. Next, the baseline noise (i.e., at the pre-treatment noise level) was mixed with the post-treatment sentence for each patient's speech. Finally, the MATLAB programme added 400 ms of babble noise prior to the onset of speech. The sentence offset was followed by 50 ms of babble noise to avoid abrupt stimulus onsets and offsets.

Once listeners passed screening, they participated in a familiarisation task, followed by the experimental speech intelligibility assessment task. During familiarisation, the listener sat alone in a double-walled IAC sound booth. The listener heard recorded sentences in noise from six patients with PD whose samples were not included as experimental stimuli. All listeners transcribed (i.e., typed) all familiarisation sentences with high accuracy and therefore proceeded to the experimental task.

For the experimental speech intelligibility assessment task, the listener remained in the sound booth. Patients' recorded speech samples were delivered from a desktop computer (Dell Optiplex 9020). The listener heard the sentence stimuli through insert-earphones presented at the patients' originally-produced SPLs (mean SPL=76 dB; SD=4.1 dB; range=66−86 dB at 30 cm, across sentences). Output was set by the examiner and checked before each experimental session. Listeners were not permitted to adjust the volume. In classic approaches to measuring speech intelligibility, speech signal levels may be equalised and listeners may be permitted to adjust the playback volume to amplify inaudible speech [52]; thus, the ability to include the contribution of vocal loudness to speech intelligibility is

lost. Our recording and playback procedures, in contrast, maintained the relative vocal loudness variations within and across the patients' sentences, permitting patients' original vocal loudness to be replicated for listeners. Any reduced vocal loudness (diminished audibility) that may have contributed to reductions in intelligibility in PD was thereby captured.

Listeners were presented with a sentence and asked to type exactly what they heard. They were permitted to replay the sentence once if needed. Upon completion of typing, listeners pressed the enter key, and the next sentence began after a delay of 50% of the duration of the previous sentence. Customised software was used for both presentation and acquisition of the TA data in MATLAB (version R2015b) [50,53].

A total of 126 sentences were used in the speech intelligibility assessment, to include one baseline and one post-treatment sentence from each of the 57 patients. This included 12 sentences (approximately 10%) repeated to enable assessment of listener reliability. Customised presentation software was used for randomisation and presentation of stimuli [53]. For each patient, the baseline and post-treatment sentences were presented as a pair, with the sentences randomised within each pair. These pairs of baseline and post-treatment sentences were presented in random order. Each listener was presented with a different randomisation of the pairs. No information was provided about the patients or treatment conditions.

In preparation for scoring, two investigators, blinded to treatment conditions, orthographically transcribed the patients' sentences (i.e., typed the sentences they heard) without noise independently. They then compared each other's transcriptions, adjusting the playback volume as needed to check accuracy. When discrepancies were revealed (5% of stimuli, primarily "a" versus "the"), they were resolved through reanalysis and discussion. TA scoring involved comparing listeners' transcriptions to the investigators' transcription and determining the percentage of words transcribed correctly. Spelling errors and homonyms were scored as correct [52].

2.5. Outcomes

The primary outcome variable for this RCT on intelligibility was blinded listeners' orthographic transcription accuracy (TA), defined here as percentage of words transcribed correctly. Results of the RCT involving SPL outcomes, not reported here, were published previously [16].

2.6. Statistical analysis

To derive the required adult listener sample size for the intelligibility assessment, 16 listeners who were recruited in New York City, comprised a pilot sample. The 16 listeners transcribed baseline and post-treatment sentences from the 57 patients in the RCT, and TA was determined. Based on these data, the effect for the voice group relative to no treatment was computed as 13·5%, with an estimate of standard deviation of 12% from mixed effects models. A Bonferroni correction was applied for the three multiple comparisons to yield an overall alpha error of 0·05 for a two-tailed test. Given these assumptions, 117 listeners were required to transcribe the speech for intelligibility assessment. From these data we derived TA as a percentage for 19 patients in each of the three groups to yield 80% power to detect differences amongst groups [54]. Intraclass correlation coefficients (ICCs) were derived to assess intra- and inter-listener reliability.

Descriptive statistics for demographic and clinical variables at baseline are presented in the Results section by group as means and standard deviations or relative frequencies. Differences amongst PD groups at baseline were tested using chi-square or Exact tests for categorical variables and analysis-of-variance or Kruskal-Wallis tests accounting for three multiple comparisons (Bonferroni). Descriptive statistics for TA are presented as least squares means (LSM) and

adjusted 95% confidence intervals derived from mixed-effects models or generalised-estimating equations, accounting for the multiple listeners and time (i.e., baseline and post-treatment).

Analysis ascribed to the intention-to-treat approach. Significance was set as $p<0.05$. All tests were two-tailed and performed using Statistical Analysis System (SAS) Version 9·4, Cary NC. There was no Data and Safety Monitoring Board, and no interim analysis was performed.

2.7. Role of funding source

3. Results

At NCVS, in Denver, a total of 81 patients with PD passed a phone screening. At the baseline visit (pre-treatment), eligibility was confirmed. The 64 patients who were not excluded were randomised to the voice group ($n = 22$), the articulation group ($n = 20$), and the no treatment group ($n = 22$) (Fig. 1). The most common reason patients were excluded was not meeting medical criteria, as described in previous work [16]. The trial was terminated as per protocol. No adverse events were reported.

Fig. 1 shows that during preparation of speech samples for intelligibility assessment, samples from 7 of the 64 patients were excluded due to patients not meeting criteria (subsequent diagnosis of non-idiopathic PD) and unusable stimuli (technical difficulties/noise). Speech samples from 57 patients ($n = 19$ in each group) remained for intelligibility assessment. The study population had a mean age of 66·5 (SD=8·5). Of the 57 patients, 16 were women. The mean number of years since PD diagnosis was 4·9 (SD=5·4). Patients' baseline demographic and clinical characteristics are listed in Table 2. As also described in previous work [16], differences amongst groups at baseline with regard to each of the demographic and clinical characteristics were not statistically significant ($p>0.26$).

At IU, Bloomington, intelligibility assessment took place between June 23, 2016 and August 14, 2017. Neurotypical young adults ($n = 134$) were recruited as listeners and were blinded to study conditions. Three listeners were excluded from participating due to failing the hearing screening. Data from a total of 14 participants were excluded from analysis due to not returning for the second part of testing ($n = 3$), not meeting language criteria ($n = 1$), or unusable data because of hardware malfunction ($n = 10$). Analysable data remained from 117 listeners. High reliability was found between and within listeners. The ICC for inter-listener reliability by group and by baseline/post-treatment for TA ranged from 0·66 to 0·83. The ICC for intra-listener reliability across all listeners pooled for TA was 0·92 ($p<0.0001$).

Table 3 presents mean TA for baseline, post-treatment and difference pre- to post-treatment in TA, as well as corresponding 95% confidence intervals. Fig. 2 shows changes in TA by treatment group. The voice group, articulation group, and no treatment group did not differ significantly in TA at baseline ($p = 0.20$), as indicated in Table 2. Specifically, pairwise differences amongst the groups were not statistically significant (voice vs. articulation $p = 1.0000$; articulation vs. no treatment $p = 0.2947$; voice vs. no treatment $p = 0.5202$).

For the voice group, within group (w) increases from baseline to post-treatment were statistically significant ($w = 31.5\%$, CI $19.6 - 43.5$; $p<0.0001$; effect size [ES]=1·2). For the articulation

Table 2
Patients' demographic and clinical characteristics and transcription accuracy at baseline by treatment group.

Characteristics (Weights)	Voice (N = 19)	Articulation (N = 19)	No treatment (N = 19)	All PD combined (N = 57)
Males (0·5)				
N (%)	14 (73·7)	15 (79·0)	12 (63·2)	41 (71·9)
Females (0·5)				
N (%)	5 (26·3)	4 (21·1)	7 (36·8)	16 (28·1)
Age in years (0·5)				
Mean (SD)	67·9 (7·2)	67·5 (9·0)	64·2 (9·2)	66·5 (8·5)
Years since diagnosis (0·5)				
Mean (SD)	4·9 (6·9)	5·0 (5·2)	4·9 (4·2)	4·9 (5·4)
Hoehn & Yahr stage w/ med. (0·5)				
Mean (SD)	2·1 (0·6)	2·3 (0·7)	2·0 (0·6)	2·1 (0·6)
Swallow (1)				
Mean (SD)	1·3 (0·9)	1·2 (1·0)	1·0 (0·6)	1·1 (0·8)
Voice (1)				
Mean (SD)	1·7 (0·7)	1·7 (0·8)	1·6 (0·7)	1·7 (0·7)
Articulation (1)				
Mean (SD)	0·8 (0·7)	0·8 (0·5)	0·6 (0·6)	0·7 (0·6)
BDI-II (0·25)				
Mean (SD)	9·6 (5·8)	9·2 (5·7)	7·6 (4·8)	8·8 (5·4)
Levadopa equivalent med. (mg/d)				
Mean (SD)	649 (380)	667 (468)	722 (418)	679 (417)
MMSE (0·25)				
Mean (SD)	28·8 (1·4)	28·7 (1·2)	29·0 (0·8)	28·8 (1·2)
Transcription accuracy*				
Mean (SD)	53·6 (31·1)	44·8 (31·9)	64·4 (26·9)	54·3 (30·1)

BDI-II= Beck Depression Inventory-II[38]; MMSE= Mini Mental Status Exam[37]; Swallow, voice and articulation were measured on a scale from 0 to 5, where 0 = no disorder and 5 = severe disorder. All variables in this Table except transcription accuracy (TA) were used in minimisation algorithm for RCT of patients with PD, who generated speech samples.[16] Randomisation ratio was 1:1:1 performed using the minimisation algorithm based on variables and weights ( ) chosen *a priori*. Subsequently, listeners performed intelligibility assessment of patient speech samples and TA was established. **There were no significant between-group differences at baseline for TA or any of the other variables.**

* Bonferroni adjusted p-values from mixed effects models across listeners for TA.

Table 3
Descriptive statistics for transcription accuracy (TA) at baseline (pre-treatment), post-treatment, and within-group changes across 117 listeners.

Variable	Voice (N = 19 patients)	Articulation (N = 19 patients)	No treatment (N = 19 patients)
	LSM [95% CI]	LSM [95% CI]	LSM [95% CI]
Baseline (pre-treatment)	53·6 [42·6, 64·5]	44·8 [37·0, 52·6]	64·4 [55·9, 72·9]
Post-treatment	85·1 [79·7, 90·5]	51·6 [45·1, 58·0]	52·5 [39·5, 65·5]
Difference Pre to Post	31·5 [19·6, 43·5]	6·8 [−3·2, 16·8]	−11·9 [−21·1, −2·7]

LSM = Least squares means; CI = Confidence intervals.

group, increases from baseline to post-treatment were not significant (w = 6·8%, CI −3·2 − 16·8; p = 0.18; ES=0·3). For the no treatment group, decreases in TA were significant (w=−11·9%, CI −21·1 − 2·7; p = 0·0115; ES=0·6).

Between-group (d) comparisons of changes in TA from baseline to post-treatment indicated that increases in the voice group were significantly greater than those for both the articulation group (d = 26·2%, CI 1·5 − 51·0; p = 0.04; ES=1·0) and the no treatment group (d = 42·8%, CI 22·4 − 63·2; p = 0.0002; ES=1·8). Differences in TA changes between the articulation group and the no treatment group were not significant (d = 16·5%, CI −6·1 − 39·2; p = 0·147; ES=0·9).

4. Discussion

While previous evidence indicates that intensive speech treatment targeting voice can make short- and long-term improvements across a range of variables in PD, [4,12,14−16,18] the present study provides the first RCT evidence that intensive speech treatment targeting voice significantly improves speech intelligibility in this neurodegenerative disorder. Two intensive speech treatments (active comparators), treatment targeting voice and treatment targeting articulation, were studied, along with a no treatment group (inactive comparator). Patients' sentence stimuli were presented to 117 listeners in multi-talker babble noise. Only treatment targeting voice yielded statistically significant gains in transcription accuracy (TA), the primary outcome variable, when compared to treatment targeting articulation and no treatment. Intra- and inter-listener reliability of TA was high.

Speech intelligibility gains had been hypothesised for both intensive speech treatments [14−16]. In previous studies in PD, when patients were simply asked to speak louder (i.e., not treated), a minimum gain of 5% in intelligibility for sentences in noise was used as a benchmark for determining clinical meaningfulness [55]. Thus, the post-treatment gain of 31·5% in TA reported here for the voice group strongly (ES 1·2) suggests that implementation of speech treatment targeting voice would generate a clinically meaningful improvement in intelligibility. These speech intelligibility outcomes are added to previously published gains following treatment targeting voice [4], [12,14−16,18−23,56]. Taken together, these data suggest that
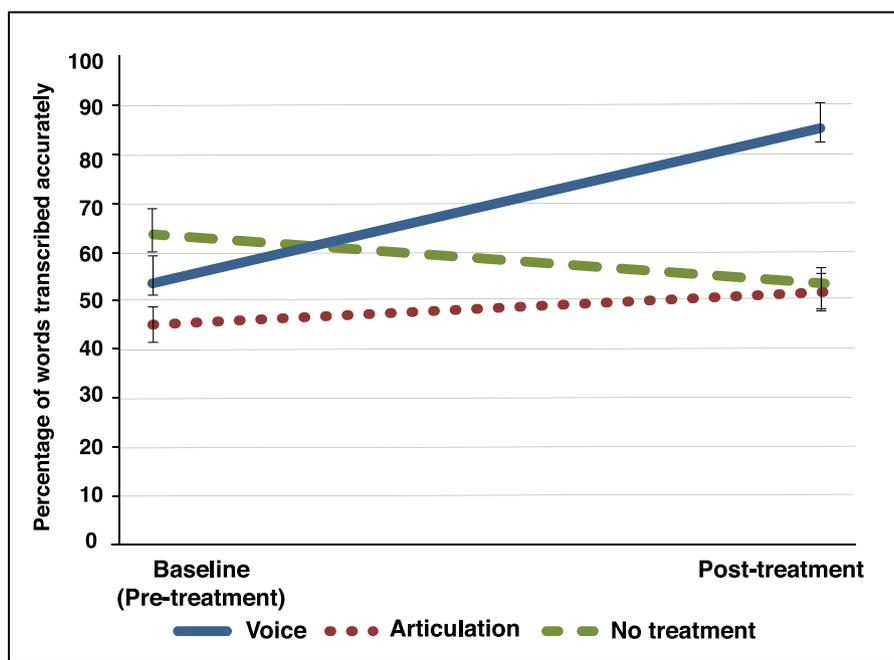
Fig. 2. Transcription (TA) accuracy by treatment. Error bars denote standard errors. For treatment targeting voice (voice group), within group increases in TA from baseline to post-treatment were significant ($p < 0.0001$). For treatment targeting articulation (articulation group), increases in TA from baseline to post-treatment were not significant ($p = 0.18$). For the no treatment group, decreases in TA were significant ($p = 0.0115$). Between-group comparisons of changes in TA from baseline to post-treatment indicated that increases in the voice group were significantly greater than those for both the articulation group ($p = 0.04$) and the no treatment group ($p = 0.0002$). Differences in TA changes between the articulation group and the no treatment group were not significant ($p = 0.147$).

meaningful improvements in communication and, thus, in health-related quality of life, would be expected when this treatment is included in patient management [57−60].

Although articulation has been considered the strongest contributor to intelligibility in PD dysarthria [26], the post-treatment gain of 6·8% in the articulation group was not statistically significant. However, potential benefits of treatment targeting articulation may be worth exploring for patient populations that typically exhibit more severe oral motor deficits.

The mechanisms underlying the intelligibility improvements following intensive speech treatment targeting voice could be twofold. First, a goal of treatment targeting voice is to increase vocal loudness. Previous outcome data from the current patient group indicated that increases in SPL, the acoustic correlate of vocal loudness, were greater for the voice group than for the articulation group [16]. Logically, increasing vocal loudness, and therefore, audibility, in patients with PD, many of whom present with decreased vocal loudness [7], would likely increase their intelligibility.

Secondly, beyond audibility, intelligibility benefits from intensive treatment targeting voice may stem from the impact of increased amplitude and coordination of movement gained across speech production subsystems when patients increase their vocal effort. That is, driving amplitude through the single target of voice may engage neurophysiological and biomechanical linkages between the vocal and the articulatory subsystems, thereby optimising treatment efficiency [4,61,62]. In support of this explanation, Neel found that while artificially amplifying habitual speech in patients with PD increased their intelligibility, having the patients speak loudly yielded even greater intelligibility improvements, when SPLs were matched [55,63]. These findings add support to the explanation of increased activity across motor systems, including in the phonatory and articulatory subsystems [4,14], contributing to improved intelligibility when voice is the target of intensive speech treatment.

Comparing two treatment groups, matched on all variables except for target, demonstrates that the difference in outcomes here is not simply due to the intensive treatment dosage. The differential improvement in intelligibility related to the treatment target of voice provides important guidance for the implementation of speech treatment in PD. The significant decrease in TA in the no treatment group could reasonably be explained by the slightly higher (nonsignificant) TA levels at baseline relative to other groups and regression toward the mean [64].

Although the present study reports on patients' speech only at baseline and immediate post-treatment, narrative speech samples were also collected 6 months post-treatment. Previous RCTs on acoustic correlates of vocal loudness and intonation have found benefits of treatment targeting voice up to 2 years post-treatment [14,15]. Our further analyses and future RCTs should address long-term maintenance of intelligibility gains following intensive voice treatment and the potential need for continued treatment.

Scientific advances and rigorous methodology in behavioural treatment were implemented in this study. amongst the strengths and innovations, two active treatment comparators and an inactive comparator were included. Furthermore, the rigorous intelligibility measure of TA by 117 blinded listeners was reported. The study methodology permitted patients' originally-produced vocal loudness to be replicated for listeners, resulting in the dimension of vocal loudness being included in intelligibility assessment for the first time in an RCT. Finally, we examined the externally valid narrative speech generated by patients with PD, controlling for effects of emotional content and more closely approximating daily spontaneous communication than most previous work [45,56].

A limitation of narrative speech, studied here, however, is its inherent variability in length, linguistic complexity, and predictability. Still, the method for sentence selection was systematic and the sentences were comparable in number of words across groups and conditions. Moreover, the study was statistically powered for patient and listener sample sizes and reliability of the findings was high.

Another possible limitation is that, whereas speech treatment targeting voice is implemented internationally (by approximately 20,000 clinicians in 70 countries, including telemedicine applications) [65-68], this study involved only American English speakers.

Although these treatment studies report positive outcomes in languages other than English [65,66,69,70], replication of rigorous intelligibility research is needed across linguistic and cultural backgrounds to assess intelligibility gains as a function of treatment internationally.

The current findings are significant in that they present the first RCT data documenting that intensive speech treatment targeting voice improves speech intelligibility in PD. Research evaluating broad clinical translation of such findings to other populations, such as individuals with ataxia [71], cerebral palsy [72], or presbyphonia [73], is an important current and future direction.

Data sharing statement

De-identified participant data may be available from the corresponding author by request.

Declaration of Competing Interest

EL, GMG, KFo, MB, and YMC declare no competing interests. KFr has been a paid consultant for LSVT Global, Inc. for many years. LR is employed as Chief Scientific Officer and has ownership interest in the for-profit company LSVT Global, Inc. She is in full compliance with Federal Statute 42 C.F.R. Part 50, Subpart F (see https://grants.nih.gov/grants/policy/coi/index.htm). She has fully disclosed any conflict of interest and her conflict of interest management plan has been approved by the Office of Conflict of Interest and Commitment at the University of Colorado, Boulder and she is in full compliance.

LR reports grants from the National Institutes of Health during the conduct of the study.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2020.100429.

References

[1] Kent RD, Kim Y. The assessment of intelligibility in motor speech disorders. In: Lowit A, Kent RD, editors. Assessment of motor speech disorders. San Diego: Plural Publishing; 2011. p. 21–37.

[2] Dorsey ER, Elbaz A, Nichols E, et al. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the global burden of disease study 2016. Lancet Neurol 2018;17:939–53.

[3] Logemann JA, Fisher HB, Boshes B, Blonsky ER. Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. J Speech Hear Disord 1978;43:47–57.

[4] Sapir S, Spielman JL, Ramig LO, Story BH, Fox C. Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with individuals with idiopathic Parkinson disease: acoustic and perceptual findings. J Speech Lang Hear Res 2007;50:899–912.

[5] Miller N, Noble E, Jones D, Allcock L, Burn DJ. How do I sound to me? Perceived changes in communication in Parkinson's disease. Clin Rehabil 2008;22:14–22.

[6] Takahashi K, Kamide N, Suzuki M, Fukuda M. Quality of life in people with Parkinson's disease: the relevance of social relationships and communication. J Phys Ther Sci 2016;28:541–6.

[7] Duffy JR. Motor speech disorders: substrates, differential diagnosis, and management. St. Louis: Mosby; 2019.

[8] Sapir S. Multiple factors are involved in the dysarthria associated with Parkinson's disease: a review with implications for clinical practice and research. J Speech Lang Hear Res 2014;57:1330–43.

[9] Kompoliti K, Wang QE, Goetz CG, Leurgans S, Raman R. Effects of central dopaminergic stimulation by apomorphine on speech in Parkinson's disease. Neurology 2000;54:458–62.

[10] Baker KK, Ramig LO, Johnson AB, Freed CR. Preliminary voice and speech analysis following fetal dopamine transplants in 5 individuals with Parkinson disease. J Speech Lang Hear Res 1997;40:615–26.

[11] Schulz GM, Grant MK. Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: a review of the literature. J Commun Disord 2000;33:59–88.

[12] Pinto S, Ozsancak C, Tripoliti E, Thobois S, Limousin-Dowsey P, Auzou P. Treatments for dysarthria in Parkinson's disease. Lancet Neurol 2004;3:547–56.

[13] Tripoliti E, Zrinzo L, Martinez-Torres I, et al. Effects of contact location and voltage amplitude on speech and movement in bilateral subthalamic nucleus deep brain stimulation. Mov Disord 2008;23:2377–83.

[14] Ramig LO, Sapir S, Countryman S, et al. Intensive voice treatment (LSVT®) for patients with Parkinson's disease: a 2 year follow-up. J Neurol Neurosurg Psychiatry 2001;71:493–8.

[15] Ramig LO, Sapir S, Fox C, Countryman S. Changes in vocal loudness following intensive voice treatment (LSVT®) in individuals with Parkinson's disease: a comparison with untreated patients and normal age-matched controls. Mov Disord 2001;16:79–83.

[16] Ramig L, Halpern A, Spielman J, Fox C, Freeman K. Speech treatment in Parkinson's disease: randomized controlled trial (RCT). Mov Disord 2018;33:1777–91.

[17] Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008;51:225–39.

[18] Baumgartner CA, Sapir S, Ramig LO. Voice quality changes following phonatory-respiratory effort treatment (LSVT) versus respiratory effort treatment for individuals with Parkinson disease. J Voice 2001;15:105–14.

[19] Baumann A, Nebel A, Granert O, et al. Neural correlates of hypokinetic dysarthria and mechanisms of effective voice treatment in Parkinson disease. Neurorehabil Neural Repair 2018;32:1055–66.

[20] Liotti M, Ramig LO, Vogel D, et al. Hypophonia in Parkinson's disease: neural correlates of voice treatment revealed by PET. Neurology 2003;60:432–40.

[21] Narayana S, Fox PT, Zhang W, et al. Neural correlates of efficacy of voice therapy in Parkinson's disease identified by performance–Correlation analysis. Hum Brain Mapp 2010;31:222–36.

[22] El Sharkawi A, Ramig L, Logemann JA, et al. Swallowing and voice effects of Lee Silverman Voice Treatment (LSVT): a pilot study. J Neurol Neurosurg Psychiatry 2002;72:31–6.

[23] Miles A, Jardine M, Johnston F, de Lisle M, Friary P, Allen J. Effect of Lee Silverman Voice Treatment (LSVT LOUD®) on swallowing and cough in Parkinson's disease: a pilot study. J Neurol Sci 2017;383:180–7.

[24] Yorkston KM, Beukelman DR, Bell KR. Clinical management of dysarthric speakers. Boston: College Hill Press; 1988.

[25] Rosenbek JC, LaPointe LL. The dysarthrias: description, diagnosis and treatment. In: Johns D, editor. Clinical management of neurogenic communicative disorders. Boston: Little, Brown, and Company; 1985. p. 97–152.

[26] Kim Y, Kent RD, Weismer G. An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. J Speech Lang Hear Res 2011;54:417–29.

[27] Hustad KC. Estimating the intelligibility of speakers with dysarthria. Folia Phoniatr Logop 2006;58:217–28.

[28] Ramig LO, Countryman S, Thompson LL, Horii Y. Comparison of two forms of intensive speech treatment for Parkinson disease. J Speech Hear Res 1995;38:1232–51.

[29] Forrest K, Weismer G, Turner GS. Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults. J Acoust Soc Am 1989;85:2608–22.

[30] Tjaden K, Sussman JE, Wilding GE. Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis. J Speech Lang Hear Res 2014;57:779–92.

[31] Dorsey ER, Bloem BR. The Parkinson pandemic—A call to action. JAMA Neurol 2018;75:9–10.

[32] Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008;148:295–309.

[33] Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Ann Intern Med 2017;167:40–7.

[34] Schwartz CE, Chesney MA, Irvine MJ, Keefe FJ. The control group dilemma in clinical research: applications for psychosocial and behavioral medicine trials. Psychosom Med 1997;59:362–71.

[35] Rains JC, Penzien DB. Behavioral research and the double-blind placebo-controlled methodology: challenges in applying biomedical standard to behavioral headache research. Headache 2005;45:479–86.

[36] Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology 1967;17:427–42.

[37] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. J Psychiat Res 1975;12:189–98.

[38] Beck AT, Steer RA, Brown GK. Manual for the beck depression inventory-II. San Antonio: Psychological Corporation; 1996.

[39] van der Kolk NM, de Vries NM, Kessels RPC, et al. Effectiveness of home-based and remotely supervised aerobic exercise in Parkinson's disease: a double-blind, randomised controlled trial. Lancet Neurol 2019;18:998–1008.

[40] Švec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. Logoped Phoniatr Vocol 2003;28:181–92.

[41] Tjaden K, Wilding GE. Rate and loudness manipulations in dysarthria: acoustic and perceptual findings. J Speech Lang Hear Res 2004;47:766–83.

[42] Yorkston KM, Dowden PA, Beukelman DR. Intelligibility measurement as a tool in the clinical management of dysarthric speakers. In: Kent RD, editor. Intelligibility in speech disorders: theory, measurement, and management. Philadelphia: John Benjamins Publishing; 1992. p. 265–85.

[43] Borod JC, Tabert MH, Santschi C, Strauss EH. Neuropsychological assessment of emotional processing in brain-damaged patients. In: Borod JC, editor. The neuropsychology of emotion. New York: Oxford University Press; 2000. p. 80–105.

[44] Dumer AI, Oster H, McCabe D, et al. Effects of the Lee Silverman Voice Treatment (LSVT® LOUD) on hypomimia in Parkinson's disease. J Int Neuropsychol Soc 2014;20:302–12.

[45] Dupuis K, Pichora-Fuller MK. Intelligibility of emotional speech in younger and older adults. Ear Hear 2014;35:695–707.

[46] Beijer LJ, Clapham RP, Rietveld ACM. Evaluating the suitability of orthographic transcription and intelligibility scale rating of semantically unpredictable sentences (SUS) for speech training efficacy research in dysarthric speakers with Parkinson's disease. J Med Speech-Lang Pathol 2012;20:17–34.

[47] American National Standards Institute. Method for measuring the intelligibility of speech over communication systems (ANSI S3.6). New York: American National Standards Institute; 1989.

[48] Schafer EC, Pogue J, Milrany T. List equivalency of the AzBio sentence test in noise for listeners with normal-hearing sensitivity or cochlear implants. J Am Acad Audiol 2012;23:501–9.

[49] Boersma P, Weenink D. Praat: doing phonetics by computer (Version 6.0.21) [computer software]. 2017.

[50] Natick M. MATLAB and statistics toolbox release (r2015b) [computer software]. The MathWorks, Inc 2015.

[51] Duquesnoy AJ. Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. J Acoust Soc Am 1983;74:739–43.

[52] Cannito MP, Suiter DM, Beverly D, Chorna L, Wolf T, Pfeiffer RM. Sentence intelligibility before and after voice treatment in speakers with idiopathic Parkinson's disease. J Voice 2012;26:214–9.

[53] Chang E, Chang Y. Word transcription VAS & sentence VAS (Version 1.0) [computer software] 2015. Contact authors for retrieval information.

[54] Muller KE, LaVange LM, Ramey SL, Ramey CT. Power calculations for general linear multivariate models including repeated measures applications. J Am Stat Assoc 1992;87:1209–26.

[55] Stipancic KL, Tjaden K, Wilding G. Comparison of intelligibility measures for adults with Parkinson's disease, adults with multiple sclerosis, and healthy controls. J Speech Lang Hear Res 2016;59:230–8.

[56] Moya-Galé G, Levy ES. Parkinson's disease-associated dysarthria: prevalence, impact and management strategies. Res Rev Parkinsonism 2019;9:9–16.

[57] Abell RV, Baird AD, Chalmers KA. Group singing and health-related quality of life in Parkinson's disease. Health Psychol 2017;36:55–64.

[58] Stegemöller EL, Radig H, Hibbing P, Wingate J, Sapienza C. Effects of singing on voice, respiratory control and quality of life in persons with Parkinson's disease. Disabil Rehabil 2017;39:594–600.

[59] Tickle−Degnen L, Ellis T, Saint−Hilaire MH, Thomas CA, Wagenaar RC. Self−management rehabilitation and health−related quality of life in Parkinson's disease: a randomized controlled trial. Mov Disord 2010;25:194–204.

[60] Trend P, Kaye J, Gage H, Owen C, Wade D. Short-term effectiveness of intensive multidisciplinary rehabilitation for people with Parkinson's disease and their carers. Clin Rehabil 2002;16:717–25.

[61] McClean MD, Tasko SM. Association of orofacial with laryngeal and respiratory motor output during speech. Exp Brain Res 2002;146:481–9.

[62] McFarland DH, Tremblay P. Clinical implications of cross-system interactions. Semin Speech Lang 2006;27:300–9.

[63] Neel AT. Effects of loud and amplified speech on sentence and word intelligibility in Parkinson Disease. J Speech Lang Hear Res 2009;52:1021–33.

[64] Morton V, Torgerson DJ. Effect of regression to the mean on decision making in health care. BMJ 2003;326:1083–4.

[65] Moya-Galé G, Goudarzi A, Bayés A, McAuliffe M, Bulté B, Levy ES. The effects of intensive speech treatment on conversational intelligibility in Spanish speakers with Parkinson's disease. Am J Speech Lang Pathol 2018;27:154–65.

[66] LSVT Global. About Us. 2020. https://www.lsvtglobal.com/About_Us (accessed April 5, 2020).

[67] Theodoros DG, Hill AJ, Russell TG. Clinical and quality of life outcomes of speech treatment for Parkinson's disease delivered to the home via telerehabilitation: a noninferiority randomized controlled trial. Am J Speech Lang Pathol 2016;25:214–32.

[68] Nakayama K, Yamamoto T, Oda C, Sato M, Murakami T, Horiguchi S. Effectiveness of Lee Silverman Voice Treatment® LOUD on Japanese-speaking patients with Parkinson's disease. Rehabil Res Pract 2020;2020:1–7.

[69] Martel-Sauvageau V, Roy JP, Langlois M, Macoir J. Impact of the LSVT on vowel articulation and coarticulation in Parkinson's disease. Clin Linguist Phon 2015;29:424–40.

[70] Whitehill TL, Kwan L, Lee FP-H, Chow MM-N. Effect of LSVT on lexical tone in speakers with Parkinson's disease. Parkinsons Dis 2011;2011:1–9.

[71] Sapir S, Spielman J, Ramig LO, et al. Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on ataxic dysarthria: a case study. Am J Speech Lang Pathol 2003;12:387–99.

[72] Fox CM, Boliek CA. Intensive voice treatment (LSVT LOUD) for children with spastic cerebral palsy and dysarthria. J Speech Lang Hear Res 2012;55:930–45.

[73] Lu FL, Presley S, Lammers B. Efficacy of intensive phonatory-respiratory treatment (LSVT) for presbyphonia: two case reports. J Voice 2013;27(786):e11–23.