Automatic quantification of prosodic measures in Dutch frontotemporal dementia populations and presymptomatic mutation carriers

Andrea Alicehajic, 10422021
University of Amsterdam
13-05-2024
Supervisor at the Erasmus MC and first examiner: dr. L.C. (Lize) Jiskoot
Supervisor at the University of Amsterdam: Dr. A.T. (Titia) Benders
Word count: 9148

Abstract

The present study examined abnormalities in prosody among Dutch populations with frontotemporal dementia (FTD), a neurodegenerative disorder presenting as a behavioural variant (bvFTD) or primary progressive aphasia (PPA), which can be further classified into three main subtypes: a semantic variant (svPPA), a non-fluent variant (nfvPPA), and a logopenic variant (lvPPA). These clinical subtypes of FTD were compared to healthy controls and to each other, with the aim of identifying potential clinical biomarkers that could aid in future diagnosis. Additionally, we investigated prosody in three presymptomatic mutation carrier subtypes, namely chromosome 9 open reading frame (*C9orf72*), progranulin (*GRN*), and microtubule-associated protein tau (*MAPT*), as compared to healthy controls and between subtypes. Using an automated speech analysis protocol, 210 semi-structured speech samples were analysed, including patients (17 bvFTD, 18 nfvPPA, 24 svPPA, 32 lvPPA, 8 mixed/atypical PPA), controls (58), and presymptomatic mutation carriers (23 *GRN*, 7 *MAPT*, 23 *C9orf72*). Four acoustic measures were extracted (f0 range, speech duration, pause duration, and pause rate) and compared between groups. Our findings showed that patients with bvFTD exhibited a wider f0 range than controls, contrary to previous findings. Pause duration was higher for all FTD patients irrespective of subtype as compared to controls; speech duration was shorter for patients with bvFTD, nfvPPA, and lvPPA; pause rate was higher for patients with nfvPPA and lvPPA compared to controls. Between FTD subtypes, significant differences were observed in speech duration and pause rate for nfvPPA. Differences between presymptomatic mutation carriers and controls and between mutation carrier subtypes were not significant. The findings of the study support abnormal prosody in patients with variants of FTD to varying degrees, with pause duration as the most sensitive clinical biomarker that could aid in the diagnosis of FTD, followed by speech duration and pause rate.

Table of contents

1. Introduction

Frontotemporal dementia (FTD) is a heterogeneous group of neurodegenerative brain disorders that are characterised by atrophy of the frontal and temporal lobes of the brain (Ahmed, Hodges & Piguet, 2021; Riedl et al., 2014; Seelaar, Pijnenburg, & van Swieten, 2009). FTD is the second most common type of early-onset dementia. Its manifestations are generally divided into two main clinical syndromes: the predominant behavioural variant (bvFTD) and primary progressive aphasia (PPA), with three different subtypes; progressive non-fluent aphasia (nfvPPA), logopenic aphasia (lvPPA), and semantic dementia (svPPA). These main presentations are found to overlap considerably with atypical parkinsonism: a brain condition causing slowed movements, rigidity, and tremors, accompanied by one or more atypical parkinsonian features, such as postural instability or supranuclear gaze palsy (Hogan et al, 2016; Litvan, 2005; Riedl et al., 2014).

When differentiating between the variants of FTD, we look out for particular symptoms or characteristics: bvFTD is characterised by a progressive decline in social cognitive functions and changes in personality and behaviour. Its diagnostic criteria include early behavioural disinhibition; apathy or inertia; loss of sympathy or empathy; perseverative and stereotyped behaviours, and hyperorality (Geraudie et al., 2021; Rascovsky et al., 2011; Riedl et al, 2014). In terms of language impairments, patients with early-stage bvFTD have been noted to have subtle linguistic deficits that do not qualify as traditional aphasia, such as reduced number of words per minute, reduced narrative organisation and impaired comprehension, and expression of abstract words and propositional speech (Nevler et al., 2017). The different manifestations of language impairments captured under PPA are grouped based on the pattern of language breakdown. In svPPA, a patient's speech remains fluent, with syntax, prosody, and motor speech intact. The patient's semantic memory, however, begins to decline, leading to loss of knowledge of words, concepts, and objects. Eventually, speech becomes empty and meaningless. In contrast to svPPA, primarily fluency is affected in nfvPPA; the patient's speech becomes agrammatic, slow, and effortful, the latter being caused by an articulation planning deficit (apraxia of speech) (Gorno-Tempini et al., 2011; Mesulam et al., 2021; Riedl et al., 2014; West et al., 2005). Lastly, lvPPA is marked by word-finding difficulties, impaired sentence repetition, and comprehension, causing the patient to speak in a slow and halting manner. Unlike nfvPPA, lvPPA patients do not demonstrate agrammatism (Gorno-Tempini et al., 2011; Josephs et al., 2006; Mesulam et al., 2021; Rohrer et al., 2012).

In order for a patient to receive an FTD diagnosis, comprehensive assessment is essential (Rosness et al., 2008). Presently, FTD diagnosis is ascertained by means of clinical examination (Rascovsky & Grossman, 2013; Verhey & Pijnenburg, 2009) consisting of an *anamnesis* (informal conversation between the patient and the neurologist/clinical geriatrician), *heteroanamnesis* (including a patient's relative or carer), and *neuropsychological testing* (of cognitive skills, e.g. language, executive functioning, and memory); along with neuroimaging (e.g., brain MRI, PET) and DNA testing (e.g. bloodwork for detecting gene mutations).

It remains challenging to diagnose FTD. This is due to its insidious and early onset; symptoms presenting in FTD largely overlap with other dementias or primary psychiatric disorders, and are therefore not easily differentiated (Hogan et al., 2016; de Koning & Schmand, 2009; Rascovsky et al., 2011; Rascovsky & Grossman, 2013; Riedl et al., 2014; Rosness et al., 2008; Varma et al., 1999). For instance, in the early stages of PPA, a patient might report difficulties that do not yet correspond to imaging retrieved via MRI and PET scans, leading to unwarranted referrals to otolaryngologists or psychiatrists (Mesulam et al., 2021). What is more; although changes in language are often the presenting symptom in all variants of FTD, reliable assessment of these changes is not readily captured by traditional neuropsychological tests, as these often require verbal instructions and verbal responses, and a patient's inability to produce correct responses can point to multiple underlying causes. For example, patients who have impoverished speech because of impaired executive function are not necessarily aphasic and patients who cannot verbalise the nature of an object do not necessarily lack semantic knowledge of the object (Mesulam et al., 2021). Furthermore, traditional neuropsychological tests do not gauge a patient's linguistic performance in a natural setting (i.e. spontaneous speech), where the patient has to combine various linguistic competences, that simultaneously influence one another, such as syntax, semantics, and prosody, in order to form sentences and carry conversations. Lastly, another issue with the use of standardised neuropsychological tests concerns the lack of consistency in test selection and scoring across different clinical sites (Pakhomov et al., 2010).

One aspect of language that is particularly difficult to assess directly is prosody. Prosody reflects the suprasegmental combination of pitch, rhythm, and amplitude characteristics of a speech pattern (Nevler & Ash, 2017), and is crucial for communication as it conveys emotional and linguistic information in day-to-day social situations (Nevler et al., 2019). Within prosody, pitch represents the attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from low to high (Plack, Oxenham & Fay, 2006);

it is a perceptual variable that is closely correlated to the physical measure of fundamental frequency (f0), which can be defined as the inverse of the longest period (repeated waveform) in a complex periodic signal, and is measured in Hertz (Hz) (Nevler & Ash, 2017). When a person produces speech, they rarely produce sounds in one particular frequency; the frequency fluctuates across the production of a sentence. An example of such variance in height (intonation) can be found in interrogative sentences, which are generally produced by means of a rising frequency. The range within which a person produces higher and lower sound frequencies can be described as their *f0 range*. Other prosodic components of speech concern the length and ratio of speech and pauses in conversation.

Prosody is often found to be abnormal in patients diagnosed with a variant of FTD (Josephs et al., 2006), although this has typically been estimated qualitatively (e.g. impressions derived from a patient's speech) (Rohrer et al., 2012; Ross et al., 2008). In the past few years, technological advancements have brought forth a more reliable, objective, and quantitative protocol for the analysis of prosody: an automated speech analysis protocol based on a Speech Activity Detector (SAD) (Ryant, 2013; Ryant, 2023) used to time-segment speech files, combined with Praat pitch tracking (Boersma & Weenink, 2022) and an open source script (v 4.7, 2003) which are used together in order to extract f0 and speech and pause measures from speech samples (Nevler & Ash, 2017; Nevler et al., 2019). Subsequently, the summarised acoustic measures of patients with an FTD variant can be calculated and compared to those of healthy controls and other FTD variants. An advantage of the quantitative protocol lies in its applicability to (semi-)spontaneous speech recordings, which was previously not as readily available within traditional neuropsychological testing.

The adoption of this method has provided some evidence in support of the hypothesis that a reduction in prosodic measures of f0 range and speech duration and an increase in pause duration and pause rate (number of pauses per minute of speech) indeed reflects impaired prosody in both bvFTD and PPA, and, consequently, that this particular method of automatic speech quantification analysis may become feasible as an ancillary tool for the clinical diagnosis of FTD. This evidence stems from two studies in which speech data of native (American) English patients diagnosed with probable bvFTD in one study (Nevler & Ash, 2017), and PPA in the other (Nevler et al., 2019) was analysed. As the proposed method constitutes a recent development yielding limited evidence as of yet and a rather small population sample, and as prosody can vary between different languages, more data (from other languages) should be included in automatic speech quantification analysis. The primary focus of the present study therefore lies in the analysis of differences in prosody in

semi-structured digitised speech samples of Dutch bvFTD and PPA patients as compared to healthy controls. The following research questions relate to the primary objective:

(1) To what extent is prosody, as measured by f0 range, speech duration, pause duration, and pause rate, abnormal in Dutch bvFTD patients as compared to healthy controls?
(2) To what extent is prosody, as measured by f0 range, speech duration, pause duration, and pause rate, abnormal in Dutch patients with variants of PPA as compared to healthy controls?
(3) To what extent is prosody as measured by f0 range, speech duration, pause duration, and pause rate different between Dutch patients with bvFTD and Dutch patients with variants of PPA?

What is more, up to 40 percent of FTD cases are linked to the familial transmission of a genetic mutation (Riedl et al., 2014). The mutations or gene expansions majoritarily occur in three genes: chromosome 9 open reading frame (*C9orf72*), progranulin (*GRN*), and microtubule-associated protein tau (*MAPT*) (Rohrer & Boxer, 2021). Additionally, close to 75% of the familial FTD cases show an *autosomal dominant pattern of inheritance*, meaning that the descendants of an individual carrying one of the genetic mutations causing FTD have a 50 percent chance of inheriting the same genetic mutation and developing FTD later in life. Such individuals are termed *presymptomatic*[1] *mutation carriers* (Breteler & Schrijvers, 2009; Rohrer et al., 2015). Research on populations of (presymptomatic) mutation carriers has provided an important insight into the earliest stages of the disease progression of FTD, namely that reductions in brain volume can occur as far as 10 years before the onset of clinical symptoms (Rohrer et al, 2013; Rohrer et al., 2015). This means that, by the time a patient demonstrates their first noticeable symptoms, a significant amount of irreversible neuronal loss has already occurred. Current research on presymptomatic mutation carriers thus aims to identify early subclinical changes to language and behaviour in order to formulate treatments that may delay or prevent disease onset (Jiskoot et al., 2016; Jiskoot et al., 2018; Nevler et al., 2024).

As part of a cohort study published earlier this year, Nevler et al. (2024) applied their automated protocol to speech data of a population of native English presymptomatic mutation carriers and compared their derived measures to those of non-carriers including, but not limited to, prosodic measures of f0 range, pause and speech durations, and pause rate. The

[1] Or asymptomatic (Nevler et al., 2024)

analysis revealed that the prosodic measure of speech duration was different between the two groups; with presymptomatic mutation carriers on average producing shorter speech segments than non-carriers, indicating a potentially valuable outcome. However, given the limited evidence available regarding the population and language sample, further investigation is warranted. For that reason, a cohort of Dutch presymptomatic FTD mutation carriers will serve as the secondary focus of the present study, and will be examined by means of the following research questions:

(4) To what extent is prosody, as measured by f0 range, speech duration, pause duration, and pause rate, abnormal in Dutch presymptomatic carriers of one of the most common types of gene mutation or expansion causing FTD (*C9orf72*, *GRN*, and *MAPT*) as compared to healthy controls?
(5) To what extent is prosody, as measured by f0 range, speech duration, pause duration, and pause rate, different between Dutch FTD presymptomatic carriers of one of one of the most common types of gene mutation or expansion causing FTD (*C9orf72*, *GRN*, and *MAPT*) as compared to healthy controls?

In the next section, acoustic characteristics associated with prosody will be examined in greater detail, alongside an overview of studies pertinent to the primary and secondary objective of the present study. Subsequently, expectations regarding the analysis outcome will be outlined.

2. Theoretical Background

2.1 Prosody: fundamental frequency (f0) and duration

As briefly described in the introduction, prosody is concerned with elements of speech that are spanned over larger units (e.g. syllables and phrases) than individual phonetic segments (e.g. vowels and consonants). These suprasegmental elements of speech are generally divided into two types of variables, namely auditory and acoustic variables. The auditory variables consist of elements that designate the listener's impression, or physiological sensation, of acoustic variables, such as *pitch* (low and high), *length* (long and short), *loudness* (soft and loud) and *timbre* (quality of sound). Acoustic variables, on the other hand, constitute the measurable physical characteristics that are closely correlated to their auditory counterparts, namely *fundamental frequency* or *f0* (measured in Hertz), *duration* (measured in

milliseconds) *intensity* (measured in decibels), and *spectral characteristics* (the distribution of energy across different parts of the audible frequency range), respectively (Hirst & Di Christo, 1998; Niebuhr et al., 2021). Acoustic and auditory variables are often used interchangeably in the literature.

The most prominent of the acoustic variables is f0, arguably, which can be described as the frequency of the vibration of the vocal folds, that is measured in cycles per second and expressed in hertz (Hz) (Hirst & De Looze, 2021). F0 is typically not stagnant, but changes constantly within an utterance and can be used for expressive and emotive purposes: raising the fundamental frequency (or pitch) throughout a phrase, for example, can be used to express surprise in English. The f0 range of a person typically lies between 80 Hz and 450 Hz, where male voices occupy a lower part of the range, and female and children's voices are often towards the higher end of the range (Bäckström et al., 2022).

In both speech and music analysis, oftentimes not Hertz, but a logarithmic transformation to the semitone (ST) scale is applied in order to analyse the data of female and male voices at once: the semitone scale captures the perceptual equivalence of similar relative pitch changes, irrespective of the absolute pitch and its change in Hertz. This is important in order to be able to compare individual differences as well as female versus male differences (Nevler & Ash, 2017; Benders, StGeorge & Fletcher, 2021; Hirst & De Looze, 2021). The following example, taken from Benders et al. (2021) demonstrates why the semitone scale allows for clearer comparison: if a man who speaks at 100 Hz raises their voice by 50 Hz, and if a woman who speaks at 200 Hz raises their voice by 100 Hz, then on the Hertz scale the woman will have raised her voice by double the amount of Hertz that of the man. However, it is generally accepted that a pitch rise or fall is perceived as being smaller for a higher-pitched voice than for a lower-pitched voice (Hirst & De Loozen, 2021), and thus when considering this pitch rise relative to their starting points, both the woman and the man have raised their pitch by half an octave, or six semitones when transformed on the logarithmic semitone scale.

Besides f0 range, another acoustic variable that can be considered relevant to FTD speech analysis is *duration*. From the duration (in seconds) of speech segments and silent (or pause) segments in a semi-structured speech fragment, it is possible to calculate a number of measures of fluency, such as mean speech segment duration, mean pause segment duration, and pause rate (number of pauses per minute of speech → ppm). As a composite of prosody, fluency has also been hypothesised to be abnormal in FTD, and particularly in PPA variants

of FTD (Nevler et al., 2017, Nevler et al., 2019., Ossenwaarde et al., 2019). Therefore, both f0 range and duration are of interest for analysis by means of the previously described automated approach.

2.2 Previous studies

In 2017, Nevler et al. developed a method for the automatic analysis of prosody, using a speech activity detector to time-segment semi-spontaneous speech data of 32 patients with bvFTD and 17 matched healthy controls, after which the Praat pitch tracker combined with an open source script were used to extract f0 percentile estimates for each participant's speech segment (in semitones); with the 90th percentile constituting f0 range. The f0 range of patients with bvFTD was compared to that of the healthy controls. On top of that, mean speech duration and mean pause duration were analysed and compared between the two groups. The findings demonstrated that on average, the f0 range of patients with bvFTD was shallower (mean 4.3 ± 1.8 ST) as compared to the healthy controls (mean 5.8 ± 2.1 ST, U = 170, p = 0.03). Additionally, mean speech segment duration differed significantly between the healthy controls ((2.15 ± 0.64 seconds) and patients with bvFTD (1.33 ± 0.33 seconds, U = 476, p < 0.005). Lastly, mean pause segment duration also differed between healthy controls (0.94 ± 0.54 seconds) and patients with bvFTD (1.73 ± 0.86 seconds, U = 101, p < 0.0002). In this study, pause rate was not calculated as a measure, but rather total speech-to-pause ratio, which was 2.84 ± 1.51 seconds for healthy controls and 1.02 ± 0.58 seconds for patients with bvFTD (U = 477, p < 0.0001).

The above mentioned method was then also applied to patients with one of the PPA variants (Nevler et al., 2019): speech samples of 59 patients with PPA (nfvPPA = 15, svPPA = 21, lvPPA = 23) and 31 matched healthy controls were analysed using the same process of automatic speech quantification as in the prior study. This yielded the following results: a significantly reduced f0 range could be found for patients with nfvPPA (mean 3.86 ± 1.15 ST) compared to healthy controls (mean 6.06 ± 1.95 ST; p < 0.001) and patients with svPPA (mean 6.12 ± 1.77 ST; p = 0.001), but not for patients with lvPPA. Pause rate differed significantly between each group compared to healthy controls (mean 32.24 ± 9.75 ppm; $p \leq$ 0.002 per contrast), and between all groups: patients with nfvPPA (mean 61.36 ± 20.8 ppm) differed from patients with svPPA (mean 47.15 ± 14.34 ppm; p = 0.02), and patients with lvPPA (58.74 ± 16.41 ppm) also differed from patients with svPPA (p = 0.04). Additionally, mean speech duration was reduced in each

patient group as compared to healthy controls ($p < 0.001$ for each contrast), but no significant differences were found between patient groups. Finally, mean pause duration was similar in all groups.

Nevler et al. (2024) then applied the automated speech analysis protocol to data of 41 presymptomatic carriers of gene mutations or gene expansions associated with FTD, encompassing *GRN* ($n = 19$), *MAPT* ($n = 8$), *C9orf72* ($n = 10$), *TARDBP* ($n = 1$), and *VCP* ($n = 3$), alongside 64 non-carriers. Their analysis focused on eight acoustic measures and twenty-two lexical-semantic measures. Taking into account the scope of the present study, only results pertaining to one of the four acoustic measures listed above will be treated here. Nevler et al. discovered that on average, presymptomatic mutation carriers exhibited shorter speech segments compared to non-carriers ($\beta = -0.28$, 95% CI [−0.55 to −0.02], $p = 0.04$). In addition, differences in measures between the three most common mutations or expansions associated with FTD, *GRN*, *MAPT*, and *C9orf72*, were examined (Nevler et al., 2024, Supplement: Mutation specific sub-analyses), although the sub-analysis was underpowered due to the small sample size of each group. However, one measure stood out as significantly different between carriers of the *MAPT* mutation and healthy controls, namely mean pause segment duration, which was on average shorter for *MAPT* mutation carriers than for healthy controls ($p = 0.02$). Other than that, no differences were found between the three groups on the basis of f0 range, mean speech segment duration, mean pause segment duration, and pause rate.

3. Analysis outcome expectations

On the basis of the automated speech analysis outcomes of the listed studies and the described profiles of language breakdown in FTD variants, we expect to find **(1)** reduced f0 range, reduced mean speech segment duration, and increased mean pause segment duration in Dutch patients with bvFTD as compared to healthy controls. Based on the difference in speech-to-pause ratio between healthy controls and patients with bvFTD, we expect the pause rate to reflect a difference as well, with patients with bvFTD having a higher pause rate than healthy controls.

For Dutch patients with a variant of PPA **(2)** we expect to find reduced f0 range in Dutch patients with nfvPPA as compared to healthy controls, as well as reduced mean speech duration and an increase in pause rate. For Dutch patients with svPPA, we expect to find a similar f0 range to healthy controls, but reduced mean speech segment duration and increased

pause rate. For Dutch patients with lvPPA, we do not expect to find a reduced f0 range as compared to healthy controls, but we do predict a reduction in mean speech segment duration and an increase in pause rate. For mean pause segment duration we do not expect to find a difference for any of the patients with a variant of PPA as compared to healthy controls.

In terms of differentiation between patients with a variant of FTD **(3)**; on f0 range, we predict to be able to differentiate patients with bvFTD and patients with nfvPPA from patients with svPPA and patients with lvPPA, but not between patients with bvFTD and patients with nfvPPA and neither between patients with svPPA and patients with lvPPA. For mean speech segment duration, we do not expect to find significant differences between patients with bvFTD, nfvPPA, svPPA, or lvPPA, but we do expect to be able to find differences in mean pause segment duration between patients with bvFTD and patients with each variant of PPA. Finally, pause rate is expected to differ significantly between all variants of PPA, but it is not yet possible to formulate a prediction for patients with bvFTD in comparison to the patients with one of the PPA variants, since the measure of pause rate was not calculated for patients with bvFTD in Nevler et al., 2017).

In terms of comparisons of presymptomatic mutation carriers to healthy controls **(4)**, we expect to find a reduction in the measure of mean speech segment duration for all three types of presymptomatic mutation carriers as compared to healthy controls, and an increase in mean pause segment duration only for the *MAPT* group compared to healthy controls. In terms of differentiation between presymptomatic *MAPT*, *GRN*, and *C9orf72* mutation carriers **(5)** on f0 range, mean speech segment duration, mean pause segment duration, and pause rate, we do not expect to find any significant differences on the basis of previous studies.

4. Method

4.1 Participants

A total of 210 digitised Dutch speech samples were analysed for the purpose of the present study: 99 FTD samples, 58 healthy control samples, and 53 presymptomatic mutation carrier samples. Within the FTD population, patients were further categorised under bvFTD (n = 17) on the one hand, and nfvPPA (n = 18), svPPA (n = 24), and lvPPA (n = 32) on the other. Additionally, a number of samples from patients diagnosed with mixed or atypical PPA (n = 8) were included in the analysis linked to research questions two and three. As this subgroup was not part of previous analyses by Nevler et al. (2017; 2019) and sample sizes are small, outcomes will be treated exploratorily. In the selection process for the samples of patients

with FTD, patients with concomitant motor disorders, such as amyotrophic lateral sclerosis (ALS), corticobasal syndrome (CBS), and progressive supranuclear palsy (PSP), were excluded in order to minimise possible motor confounds. Samples of patients with (potential) underlying pathologies, such as Alzheimer's disease (AD), vascular dementia (VaD), and chronic traumatic encephalopathy[2] (CTE) were not excluded. Lastly, speech samples of presymptomatic mutation carriers (PMC), collected from the Dutch familial FTD risk cohort, were grouped according to their respective gene mutation or gene expansion: *GRN* (*n* = 23) , *MAPT* (*n* = 7), and *C9orf72* (*n* = 23). In Table 1 and 2 below, demographic data per population and subpopulation is listed.

Table 1. *Demographic data per population*

	FTD	**PMC**	**HC**
n	**99**	**53**	**58**
Age (mean)	**67**	**53**	**55**
Sex = Female (%)	**47 (46.5%)**	**34 (64.2%)**	**37 (63.8%)**
Level of Education* (mean)	**5**	**6**	**6**

*According to the Dutch educational system, categorised from level 1 to 7, where 1 equals less than 6 years of primary education and 7 equals academic education (Duits & Kessels, 2014). Abbreviations: FTD = frontotemporal dementia; PMC = presymptomatic mutation carriers; HC = healthy controls

Table 2. *Demographic data per subpopulation*

	bvFTD	**nfvPPA**	**svPPA**	**lvPPA**	**m/aPPA***	***GRN***	***MAPT***	***C9orf72***
n	**17**	**18**	**25**	**31**	**8**	**23**	**7**	**23**
Age (y, mean)	**66**	**69**	**62**	**69**	**71**	**59**	**54**	**46**
Sex = Female (%)	**7 (41.2%)**	**12 (66.7%)**	**9 (36.0%)**	**14 (45.2%)**	**5 (62.5%)**	**16 (69.6%)**	**4 (57.1%)**	**14 (60.9%)**
Level of Education (mean)	**5**	**5**	**5**	**5**	**5**	**6**	**5**	**6**

Abbreviations: bvFTD = behavioural variant frontotemporal dementia; nfvPPA = non-fluent variant primary progressive aphasia; semantic variant primary progressive aphasia; logopenic variant primary progressive aphasia; mixed or atypical primary progressive aphasia; *GRN* = progranulin; *MAPT* = microtubule-associated protein tau; *C9orf72* = chromosome 9 open reading frame.

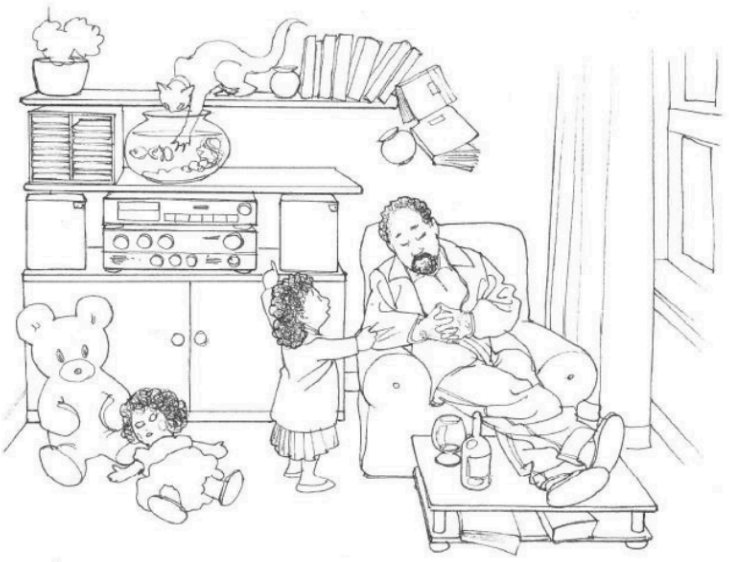
4.2 Speech samples

All speech samples were collected between 2016 and 2024 at the Department of Neurology and Alzheimer Centre of the Erasmus Medical Centre (Rotterdam, the Netherlands). For each person, we selected the most recently available recording of the Dutch Comprehensive

[2] A neurodegenerative disorder associated with repeated head trauma (McKee, Stein, Kiernan & Alvarez, 2015).

Aphasia Test's (CAT-NL) situation description task (Visch-Brink, Vandenborre, De Smet, & Mariën, 2014; adapted to Dutch from Swinburn, Porter & Howard, 2014). This task elicits semi-structured narrative speech and is part of the standard diagnostic procedure for FTD at the Department of Neurology and Alzheimer Centre of the Erasmus Medical Centre. For the situation picture that is used for elicitation in the task, see Figure 1.

Figure 1. *CAT-NL situation description task*



4.3 Sound processing

An automated speech quantification protocol developed at the University of Pennsylvania (Nevler et al., 2017) and retrieved from Nevler (the UPenn-FTDC-Acoustic-pipeline, [GitHub repository], 2023), was adopted for sound processing. This protocol consists of four main parts, which will be described below, as well as any adaptations that were made to the protocol for successful completion of the analysis. All speech samples were converted to .WAV format and adapted from stereo to single channel in Praat (v.6.3.18) beforehand. The first step of the protocol entailed running the LDC Speech Activity Detector (Ryant, 2013) in order to time-segment the acoustic signal into speech and non-speech fragments. The SAD is independent of a specific language (Nevler et al., 2019) and could therefore be applied to Dutch data without preprogramming. The SAD python script required libraries to be installed which were not available outside of the UPenn environment. Therefore, a more recent version of the same SAD (Ryant, 2023) taken from an open source (Linguistic Data Consortium, n.d.) was installed, and as in the protocol, boundaries for speech segments were set for 250 milliseconds, and 150 milliseconds for non-speech segments. The output of the SAD in the

original protocol consisted of a number of .lab files that needed to be converted to TextGrids for Praat. The current SAD made it possible to include this step in the coding in order to create TextGrids as the direct output.

The second step consisted of manually validating the time-segmentation and labelling (performed by the SAD) in Praat, and labelling interviewer speech and other noises that could negatively influence pitch estimates and segment durations, such as coughing, laughing, and thumps on the microphone. One element that was not described in the protocol, but turned out to be the most laborious aspect of the manual validation, concerned filler words such as 'uhm', 'en' (*and*), and 'dus' (*so*). These segments were all labelled as non-speech, as they were uttered at moments when the speaker was thinking of what to say, as opposed to actually describing the events in the situation picture.

In the third part of the protocol, pitch and duration data were extracted from the TextGrids with the 'get_pitch_quantiles_duration_5.praat' Praat script. Following Nevler et al. (2017; 2019), pitch tracking limits were set at 75 - 300 Hz. The script yielded a data frame where each row designated a segment with f0 estimates in 10 percentile bins ranging from the 10th to the 90th percentile, as well as start time, end time, and duration of the segment. An issue that arose during the examination of the table, was that 124 speech segments returned –undefined– as values for the entire row. This was the case for 13 patients, 10 healthy controls, and 7 presymptomatic mutation carriers. After analysing these segments in Praat, it was concluded that pitch contours were sufficient for the segments to be labelled as speech, and an additional script (Vet, 2024) was written in order to obtain the pitch estimates and durations of these segments as well. The second script yielded a data frame containing the pitch estimates and durations of the previously undefined segments, which were then combined with the original output.

The pitch and duration data provided the input for the fourth and final part of the protocol which was carried out with the R script 'process_SAD_f0_durations_20220419' (v1). F0 percentiles were converted from the Hertz scale to the semitone scale and each subject's 10th percentile bin was taken as the reference for the semitone scale, as this helps control for individual differences in pitch. Pauses were flagged and removed if they were the first or last segment of a recording, or if they followed interviewer speech. Additionally f0 range outliers were identified and removed by assigning a cutoff at 10.5 ST, after which summarised measures were calculated for f0 percentiles, mean speech segment duration, mean pause segment duration, and pause rate for each recording (or subject). The summarised acoustic measures provided the input for the statistical analysis.

Upon inspection of the summarised acoustic measures, it is worth noting that, despite a pitch cutoff assigned at 10.5 ST in the data preprocessing stage, several individual f0 range means stood out within the dataset, as they were located on the higher end of the spectrum, ranging between 6.3 and 9 ST. The samples (n = 13) were distributed over various groups: control (n = 4), bvFTD (n = 2), mixed/atypical PPA (n = 2), nfvPPA (n = 3), *GRN* (n = 1), and *C9orf72* (n = 1). Upon qualitative examination of the pitch variability per recording it was concluded that this variability was not due to potential methodological confounds. However, for two subjects, one patient with bvFTD (f0 range = 6.5 ST) and one patient with mixed/atypical PPA (f0 range 6.6 ST), a creaky quality was observed in some parts of the recording. These did not appear to affect the lowest and highest bound of pitch, and therefore the recordings were included in the analysis.

4.4 Statistical analysis

All calculations were conducted with R (v.4.3.3) in RStudio (RStudio-2023.12.1-402), and additional R packages tidyverse (v2.0.0), ggplot2 (v3.5.0), dplyr (v1.1.4), ggpubr (v0.6.0), RColorBrewer (v1.1.3), ggrain (v0.0.4), and rstatix (v0.7.2). Q-Q plots and Kernel density plots were used to inspect distributions in speech variables. As many of the speech variables diverged from normal distribution, a non-parametric Mann-Whitney test with Bonferroni correction was used for the group comparison between patients with bvFTD and healthy controls on measures of f0 range, speech segment duration, and pause rate (RQ1). For multiple comparisons between more than two groups (RQ 2-5), non-parametric Kruskal-Wallis tests were performed, combined with post hoc pairwise Wilcoxon tests and Bonferroni corrections for the adjustment of p-values. Within these tests, medians and measures of variability are calculated and compared, rather than means and standard deviations. Nevertheless, individual means and standard deviations can provide a clearer and more intuitive representation of the data's central tendency and variability, and therefore these values will also be listed in the results along with the U-statistic and p-values of the non-parametric tests, as well as visualisation of the results in the shape of error bars and rain plots. Please note that Figures were plotted for each acoustic measure pertaining to each research question. A selection of these is treated alongside the most relevant results in the following section. For the complete List of Figures: see Appendix.

5. Results

RQ1: F0 range differed significantly between the control group (mean 4.44 ± 1.27 ST) and the bvFTD group (mean 5.16 ± 0.98 ST, U = 669, $p < 0.03$); f0 range being shallower on average for the control group, as demonstrated in Figure 2. Here, we can also observe that the f0 range distribution of the control group is fairly normally distributed, whereas the bvFTD f0 range is most concentrated between 5-6 ST, approximately.

Figure 2. *F0 percentiles (left) and f0 range distribution (right)*

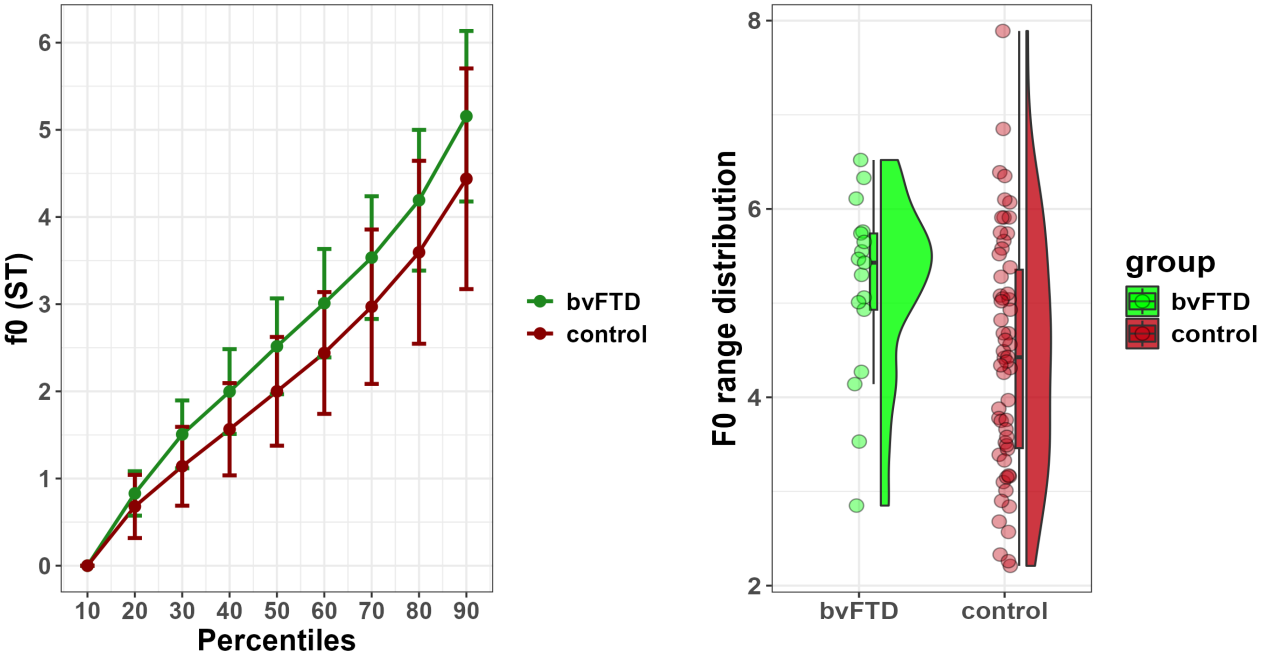
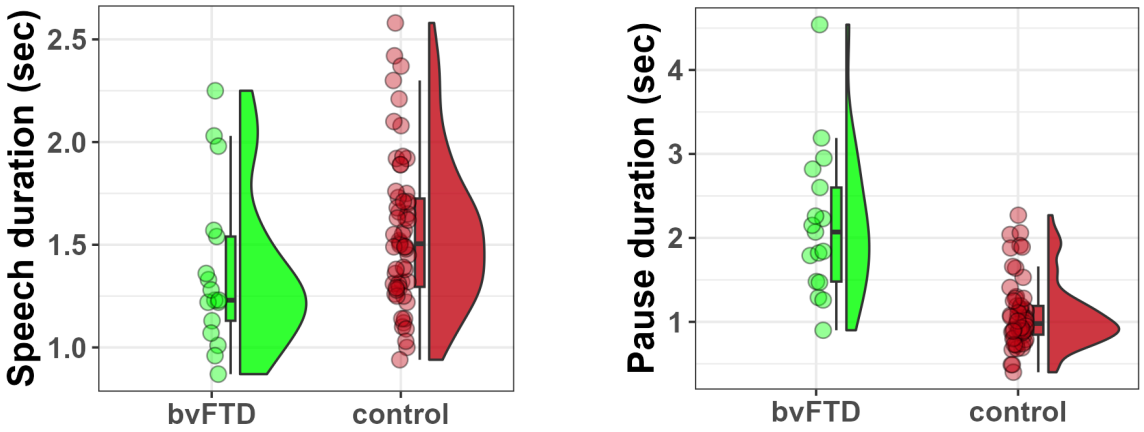


Figure 3. *Mean speech segment duration (left) and mean pause segment duration (right)*

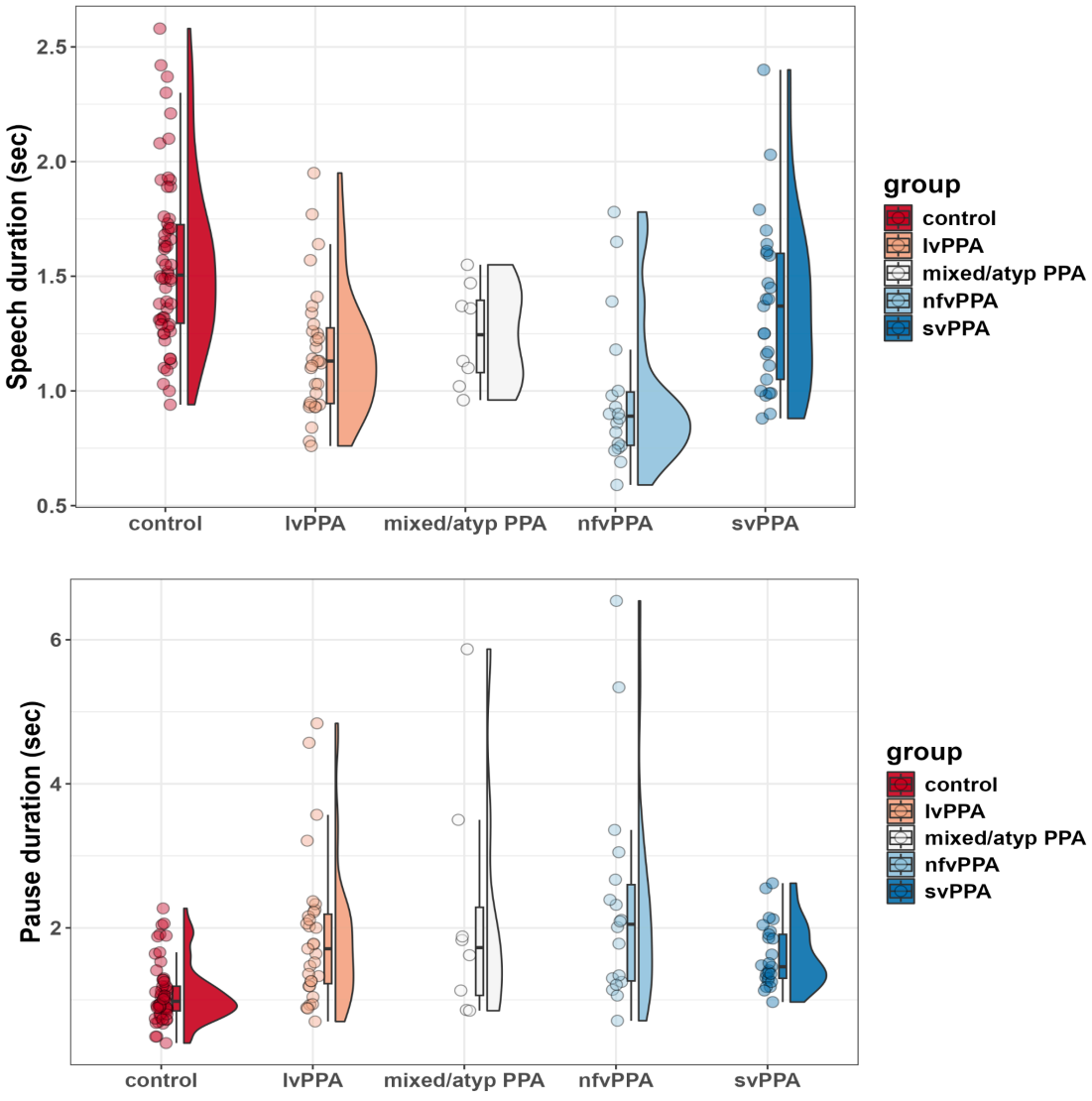


Conversely, mean speech segment duration (Figure 3) was significantly higher for the control group (mean 1.56 ± 0.37 seconds, U = 318, $p < 0.03$) than for the bvFTD group (mean 1.37 ± 0.37 seconds). Mean pause segment duration also differed significantly between the two

groups, the control group exhibiting on average shorter pause segment durations (mean 1.08 ± 0.41 seconds, U = 885, p = 7.259e-07) than the bvFTD group (mean 2.16 ± 0.88 seconds). Lastly, the bvFTD group exhibited a higher pause rate on average (44.3 ± 11.8 ppm, U = 645, p = 0.055) than the control group (38.6 ± 9.53 ppm), but this difference only approached significance slightly.

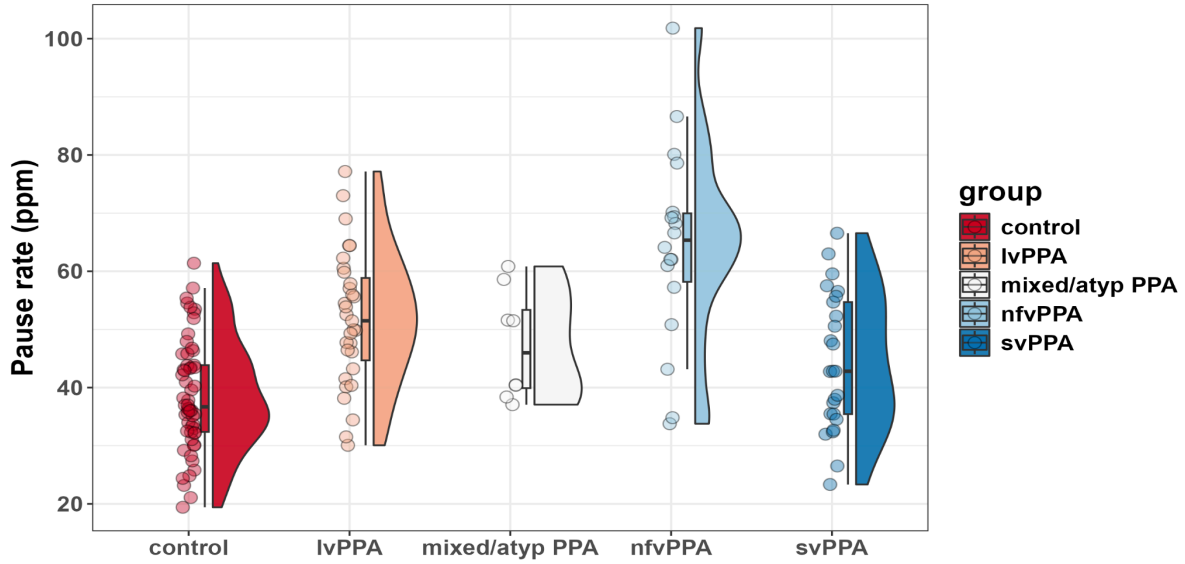
RQ2: F0 range did not differ significantly for any of the PPA variants (lvPPA, mixed/atypical PPA, nfvPPA, svPPA) compared to the control group (p = 0.076). Kruskal-Wallis tests for mean speech segment duration, mean pause segment duration, and pause rate, on the other hand, returned highly significant p-values (p < 0.001 per contrast). Post hoc analyses showed that on average, speech segments of the control group were longer (mean 1.56 ± 0.37 seconds) than speech segments of all of the PPA variants (Figure 4), but that this difference was only significant for the control group compared to the lvPPA group (mean 1.17 ± 0.28 seconds, U = 1466, p < 0.001) and the nfvPPA group (0.98 ± 0.37 seconds, U = 928, p < 0.001).

Figure 4. *Mean speech segment (top) and mean pause segment (bottom) durations per group*



Pause segment duration (Figure 4) was shortest on average for the control group (mean 1.08 ± 0.41 seconds), differing significantly from the mean pause segment durations of the lvPPA group (mean 1.89 ± 1 seconds, U = 352, $p < 0.001$), the nfvPPA group (mean 2.32 ± 1.52 seconds, U = 152, $p < 0.001$), and the svPPA group (mean 1.61 ± 0.44 seconds, U = 226, $p < 0.001$), but not for the mixed/atypical PPA group.

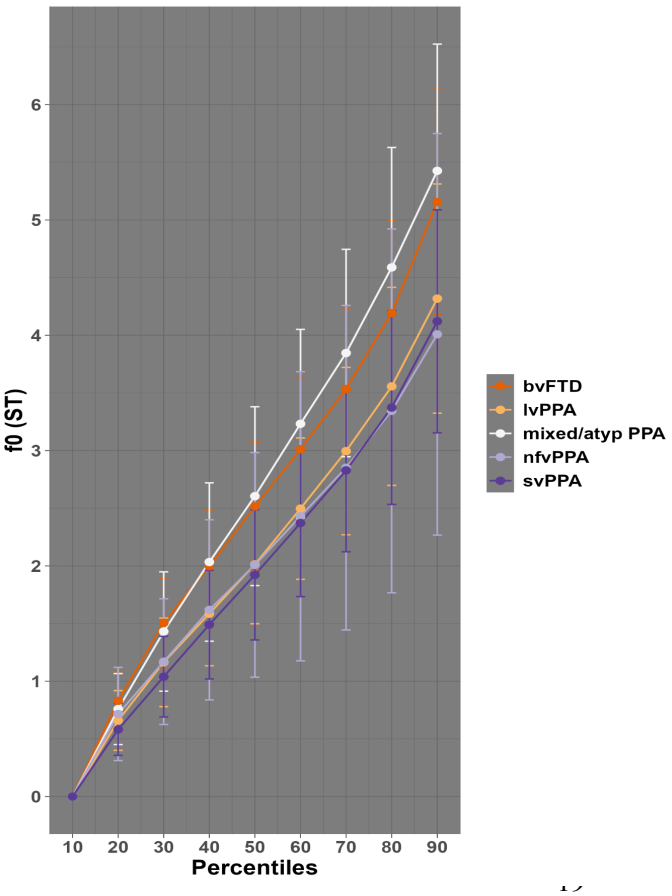
Figure 5. *Pause rate per group*



The control group also demonstrated a lower pause rate (mean 38.6 ± 9.53 ppm) compared to all PPA variant groups (Figure 5), and this difference was significant compared to the lvPPA group (mean 51.8 ± 11.5 ppm, U = 346, $p < 0.001$) and the nfvPPA group (mean 64.4 ± 17.1 ppm, U = 106, $p < 0.001$) but not for the svPPA group or the mixed/atypical PPA group.

RQ3: Kruskal-Wallis tests for group differences between bvFTD, lvPPA, mixed/atypical PPA, nfvPPA, and svPPA groups, returned significant p-values for f0 range (Figure 6), mean speech segment duration, and pause rate ($p < 0.003$, $p <$

Figure 6. *F0 percentiles per group*



0.001, and $p < 0.001$, respectively), but not for mean pause segment duration ($p = 0.31$). A post hoc analysis on the difference in f0 range between groups showed that the f0 range was shallower for lvPPA (mean 4.32 ± 0.99 ST, U = 392, $p = 0.05$) and svPPA (mean 4.12 ± 0.97 ST, U = 337, $p = 0.01$) as compared to bvFTD (mean 5.16 ± 0.98 ST). The nfvPPA group exhibited the shortest mean speech segment duration (mean 0.98 ± 0.33 seconds) compared to the other variant groups (Figure 7), and this difference was significant compared to bvFTD (mean 1.37 ± 0.39 seconds, U = 253, $p = 0.006$), lvPPA (mean 1.17 ± 0.28 seconds, U = 417, $p < 0.05$), and svPPA (mean 1.37 ± 0.37 seconds, U = 78.5, $p = 0.003$). Differences were not statistically significant between groups in terms of mean pause segment duration (Figure 8).

Figure 7. *Mean speech segment duration per group*

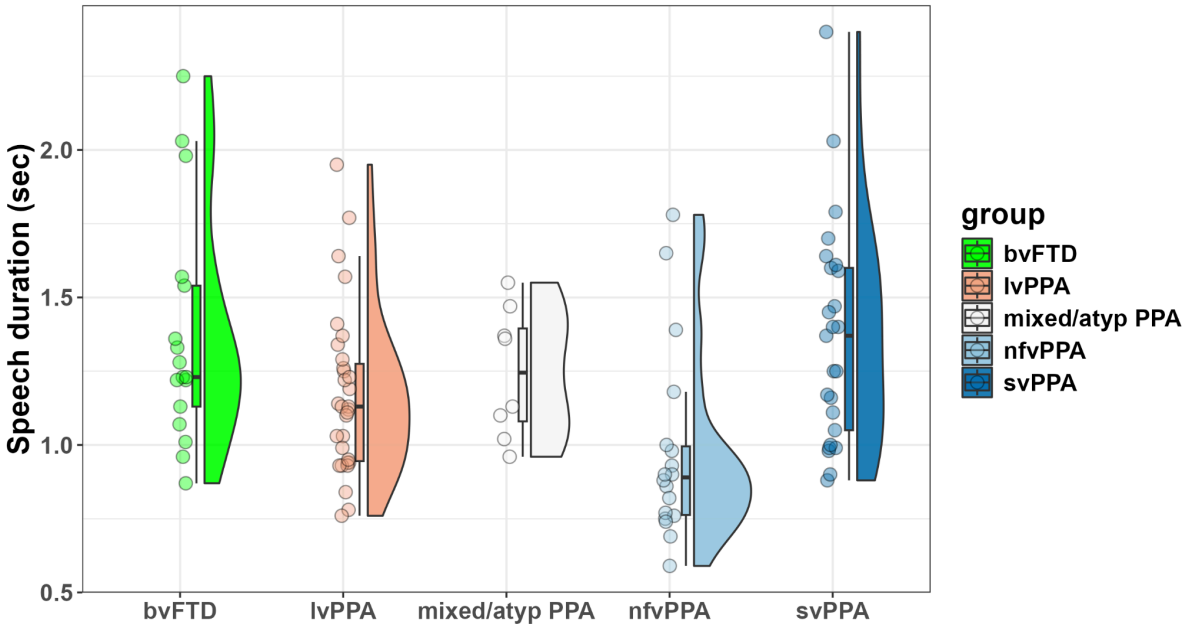
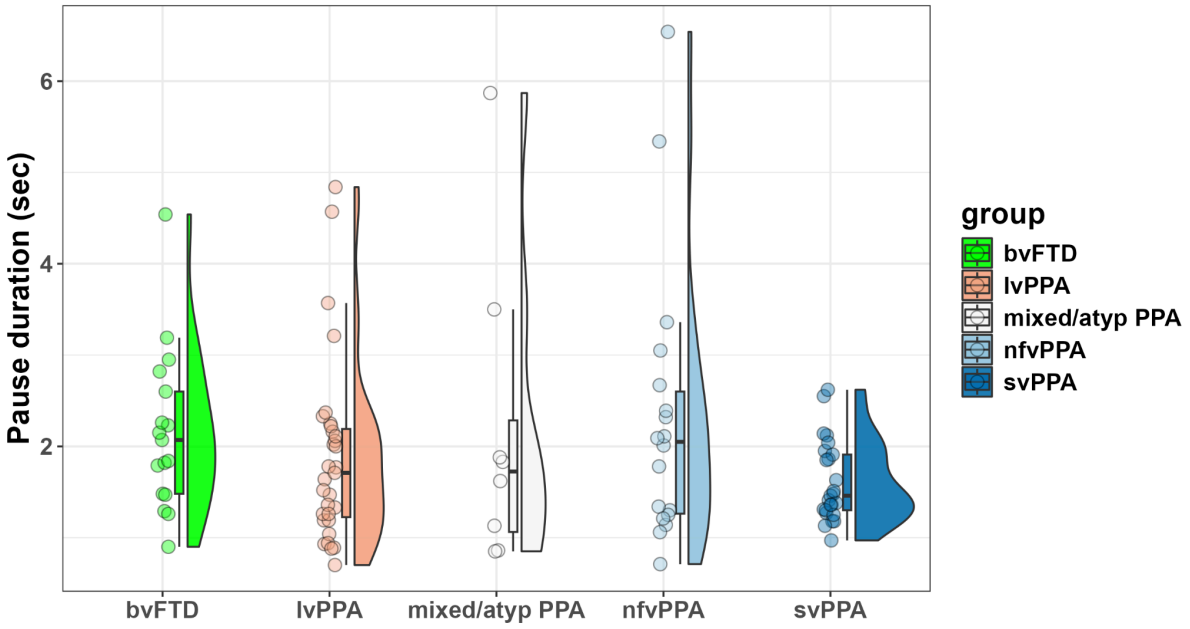
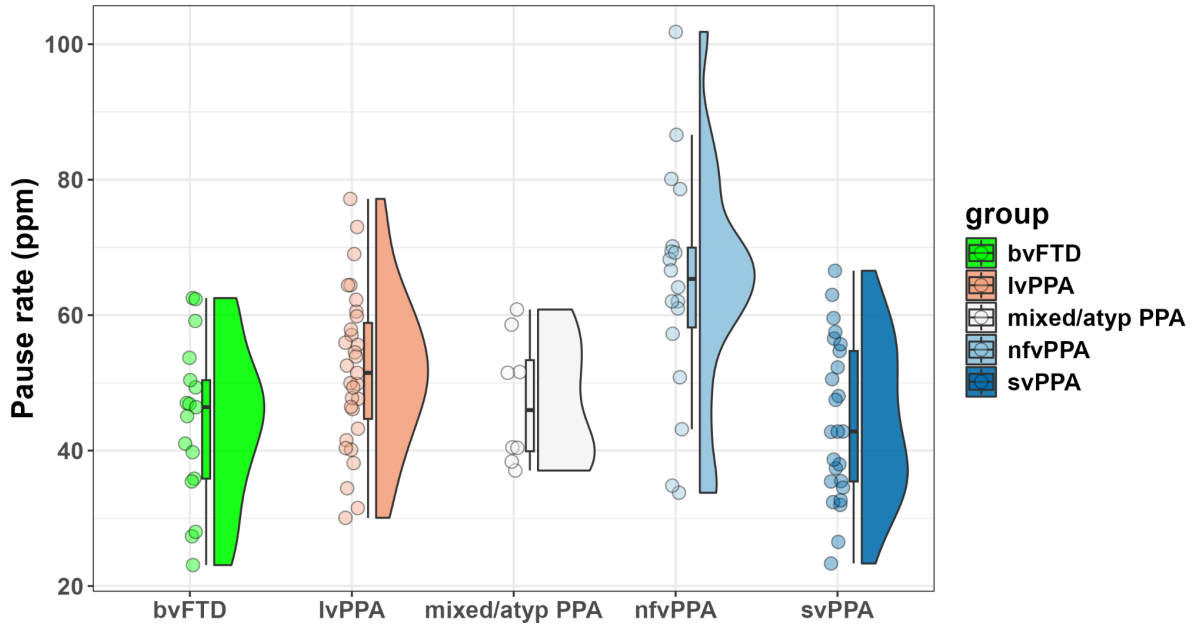


Figure 8. *Mean pause segment duration per group*



Pause rate differed significantly between groups as well. The nfvPPA group exhibited the highest pause rate on average (mean 64.4 ± 17.1 ppm), and this difference was significant compared to the bvFTD (mean 44.3 ± 11.8 ppm, U = 51, p = 0.005), lvPPA (mean 51.8 ± 11.5, U = 141, p < 0.04), and svPPA (mean 44.3 ± 11.8, U = 381, p < 0.001) groups.

Figure 9. *Pause rate per group*



RQ4 & RQ5: Lastly, a Kruskall-Wallis test exploring the differences between controls compared to the *C9orf72*, *GRN*, and *MAPT* groups on the one hand and differences between the *C9orf72*, *GRN*, and *MAPT* groups on the other, did not yield a statistically significant outcome on the basis of the acoustic measures (f0 range: p = 0.19, mean speech segment duration: p = 0.88, mean pause segment duration: p = 0.08, and pause rate: p = 0.82).

6. Discussion

The present study hypothesised abnormal prosody in various Dutch FTD populations compared to healthy controls and adopted a validated approach, the automated speech analysis protocol (Nevler et al., 2017), in order to investigate differences in acoustic measures of f0 range, mean speech segment duration, mean pause segment duration, and pause rate, between (and within) groups of patients with bvFTD, patients with variants of PPA, presymptomatic mutation carriers, and healthy controls. In keeping with studies by Nevler et al. (2017; 2019; 2024), several research questions and predictions were formulated regarding the extent to which prosody would be abnormal in these populations compared to healthy controls and to one another.

The first research question investigated prosodic differences between healthy controls and patients with bvFTD: based on findings by Nevler et al. (2017) and language patterns observed in bvFTD, a reduced f0 range in Dutch patients with bvFTD as compared to healthy controls was expected. Yet, in our sample, patients with bvFTD produced a wider average f0 range than healthy controls. On measures of speech and pause duration, our findings were more consistent with the results of Nevler et al. (2017), as patients with bvFTD produced shorter mean speech segments and longer mean pause segments as compared to healthy controls. For pause rate, it was not possible to identify a significant difference, although patients with bvFTD did exhibit an elevated average pause rate compared to healthy controls. As three out of four acoustic measures differed significantly between patients with bvFTD and healthy controls, we could conclude here that prosody is abnormal in Dutch patients with bvFTD to a large extent. However, the problem we face is that the direction of the difference in f0 range is not supported by previously listed studies and linguistic profiles within the disorder, which describe a reduction in bvFTD f0 range. How could we account for the increased f0 range in this population?

A potential explanation could be that, although subjects' f0 percentiles were converted to the semitone scale, and although each subject's 10th percentile bin was taken as the reference for the semitone scale in order to control for individual differences in pitch, variables like age, height, or sex, may have still had some influence on the f0 range outcome. Another explanation could concern heterogeneity within the bvFTD group in terms of disease progression, disease severity, and affected brain regions in each subject. Any of these explanations would need to be viewed in light of the small sample size (n = 17) and an uneven distribution. Alternatively, compensatory mechanisms may also account for the production of a wider f0 range: increased vocal effort or changes in speech production strategies might be adopted in order to compensate for cognitive or neurological decline. It could also be the case that other factors were interfering with f0 range, which could perhaps be related (though not exclusively) to other changes in behavioural characteristics of bvFTD: one subject chose to use a differently pitched voice during the picture description task in order to act out the reaction of one of the characters; another talked in an amused manner in parts of their recording. However, no other aspects stood out in the recordings of the other patients with bvFTD. It is also important to consider differences in expressive prosody between (here) Dutch and American English in Nevler et al. (2017): although we could not find a study directly comparing the average f0 ranges of these two languages, a number of studies do describe British English having a larger standard pitch range than Dutch (Willems,

2010; Chen, Chen, Kager & Wong, 2014; Marcoux & Ernestus, 2019), and this difference in mean f0 ranges appears to be reflected in Nevler et al. (2017) and the present study as well, with Dutch controls having a narrower mean f0 range compared to American English controls (4.4 ST and 5.8-6.2 ST, respectively). Could the wider f0 range in Dutch patients with bvFTD therefore also be attributed to language-specific effects on manifestations of neurodegeneration, and if so, what could be the cause of higher pitch variability? All of the above listed possibilities warrant further investigation.

For the second research question, a reduced f0 range was expected to be found only in our nfvPPA sample compared to that of healthy controls, and although healthy controls did have a wider average f0 range than patients with nfvPPA, a significant effect was not found. Findings concerning f0 range for our svPPA and lvPPA samples were consistent with Nevler et al. (2019), in that the mean f0 ranges of patients with svPPA and lvPPA were similar to the mean f0 range of healthy controls. In terms of mean speech segment duration, a significant reduction was expected for patients with nfvPPA, lvPPA, and svPPA compared to healthy controls. This reduction turned out to be robust for patients with nfvPPA and lvPPA, but it was not possible to find a significant difference between healthy controls and patients with svPPA, although they did produce shorter mean speech segment durations on average than healthy controls. In our sample, mean pause segment durations were significantly longer in patients with nfvPPA, lvPPA, and svPPA in comparison to healthy controls, whereas this effect was not reported in Nevler et al. (2019). Last, the measure of pause rate was expected to be significantly higher in patients with nfvPPA, svPPA, and lvPPA, but this difference was only borne out for patients with nfvPPA and lvPPA compared to healthy controls. Please note that nearly every contrast involving patients with mixed/atypical PPA compared to healthy controls returned a significant p-value initially, but received a p-value above 0.05 after correction (Bonferroni). Besides the likelihood of being a false positive, other possible explanations for this adjustment could be linked to small sample size (n = 8) and high variability within the sample. Due to these confounds, no further conclusions will be drawn about abnormalities in the prosodic profile of mixed/atypical PPA at this time.

In determining to what extent prosody is abnormal in patients with variants of PPA compared to healthy controls, we find that performance on the measure of f0 range cannot be considered abnormal in patients with PPA variants based on the analysis outcome. An effect of f0 range was expected in patients with nfvPPA in accordance with Nevler et al. (2019), yet when we consider once more that pitch variability appears to be less wide in Dutch as compared to English, it could be the case that this might not be the most challenging aspect of

prosody for Dutch patients with nfvPPA, although there could be other confounds involved. The remaining three measures of fluency, on the other hand, were all found to be abnormal in lvPPA and nfvPPA as compared to healthy controls, which is consistent with the pattern of language breakdown found in these populations, but not entirely in line with the study by Nevler et al. (2019), as their subjects did not exhibit an increase in pause duration. Finally, prosody was abnormal to a limited extent in patients with svPPA, as it was only possible to identify a significant effect of mean pause segment duration within this population. Here too, Nevler et al. (2019) demonstrated different results than we found in the present study, as they found a significant effect of pause rate and mean speech segment duration for English patients with svPPA. However, according to the literature, fluency in svPPA remains relatively intact, which is why it is not alarming that we did not find the same effects. These differences in outcomes between Dutch and English patients could be ascribed to various causes, such as language-specific differences, heterogeneity within and between groups, or interaction with other linguistic features in each language.

Research question three explored which acoustic measures could possibly function as indicators of differences between patients with bvFTD, nfvPPA, lvPPA, svPPA, and mixed/atypical PPA. In terms of f0 range, we expected to find similar f0 ranges for patients with nfvPPA and bvFTD (narrower) on the one hand, and for patients with lvPPA and svPPA on the other (wider). We did find a difference of f0 range in bvFTD compared to lvPPA and svPPA, but the direction of this difference was opposite to what was hypothesised, as patients with bvFTD exhibited a wider f0 range than patients with lvPPA and svPPA. Additionally, it was not possible to identify a significant difference in f0 range between patients with nfvPPA on the one hand, and lvPPA and svPPA on the other. Next, contrary to our expectations for the acoustic measure of mean speech segment duration, it was actually possible to identify differences in our sample: patients with nfvPPA produced shorter speech segments than patients with bvFTD, lvPPA, and svPPA. It was not possible, on the other hand, to identify significant differences in mean pause segment duration between groups. Last, we expected to find differences between all PPA variants in terms of pause rate, but the results indicated that the differences were only found to be significant for patients with nfvPPA, who exhibited the highest average pause rate compared to patients with bvFTD, lvPPA, and svPPA. For patients with atypical/mixed PPA, we encountered the same type of results as we found for research question two, and we will thus refrain from drawing conclusions.

In determining to which extent it is possible to differentiate between populations with variants of FTD on the basis of one or more acoustic measures, we conclude that it is only

possible to a limited extent: nfvPPA appears to be the only variant of FTD that can be distinguished from all other variants (excluding mixed/atypical PPA here) on the basis of two out of four acoustic measures, namely mean speech segment duration, and pause rate. As both are fluency measures, the result is in line with the literature on the profile of language breakdown in nfvPPA compared to bvFTD, lvPPA, and svPPA, yet we would have expected to find more differences, also when considering findings in Nevler et al. (2017; 2019). The possible explanations that have been discussed under research questions one and two are equally relevant to research question three, both with regard to differences between Dutch FTD variants and English FTD variants on the one hand, and among Dutch FTD variants on the other.

The final two research questions investigated whether it was possible to find differences in prosody in presymptomatic carriers of mutations or gene expansions in *MAPT*, *GRN*, and *C9orf72* as compared to noncarriers, as well as prosodic differences between carriers of mutations/expansions in *MAPT*, *GRN*, and *C9orf72*. Unfortunately, it was not possible to identify any differences between mutation carriers and non-carriers, nor was it possible to identify differences between mutation carriers of mutations/expansions in *MAPT*, *GRN*, and *C9orf72*. Like Nevler et al. (2024), we cannot rule out that there are differences, but that these may be too subtle considering the lack of statistical power within the present study, due to a small sample size and uneven distribution of subjects within groups.

From the discussion of the five research questions, it becomes clear that there are a number of limitations within the present study that need to be addressed, the most important of which can be summarised as follows: the current study design may have oversimplified the number of, and the interaction between variables that might be involved in the characterisation of the established variants of FTD. It may not have sufficed to control for individual differences through converting to a semitone scale and using subjects' own 10th percentile bins, without explicitly covarying for demographic variables such as sex, height, age, and level of education, as well as other variables such as disease progression and duration. Due to this simplification, it becomes more challenging to provide a singular explanation for unexpected results, such as a wide f0 range in bvFTD. Another limitation concerns the small sample sizes and great variation within groups, although this is a common occurrence in research on FTD and other neurodegenerative disorders: partly due to the sensitivity surrounding including patients in research and partly due to the diverse nature of the disorder. It also explains why the patient groups, the presymptomatic groups, and the control group were not as well-matched in demographic variables. Another potential

methodological confound concerns the selection of recordings: the most recently available recording was used, as opposed to the most early available recording, which could have caused some subjects to be familiarised to the task, and it would be difficult to find out in which sense this would have manifested in the results.

Furthermore, although the Speech Activity Detector is not language dependent, and although healthy controls were used as the baseline for the analysis outcome, and although the trajectory of neurodegeneration in FTD is not necessarily considered to be language-specific, the present study could have benefitted from investigating language-specific characteristics of Dutch prosody, in comparison to (American) English prosody so as to formulate more precise predictions on the expected outcome of the study. What also ties into this, is the fact that prosody interacts with different linguistic levels such as phonology and syntax, which is also language-specific, and such interaction could be responsible for more pauses and shorter or longer speech segments. Finally, one more limitation needs to be addressed briefly with regard to the automated speech analysis protocol: there are no instructions within the protocol on how to treat speech segments that consist of the enactment of how characters in the picture would speak. The decision of a subject to do so may be of interest for research on FTD, especially for patients with bvFTD.

Considering the ambition of the present study, of identifying sensitive clinical biomarkers using an automated speech analysis protocol that could be used as an ancillary tool for the clinical diagnosis of FTD, there are strengths to the present study too: it is the first of its kind that applies the automated speech analysis protocol to another language than English, and its findings show promise; prosody in Dutch patients with FTD is found to be abnormal compared to healthy controls, which has successfully and objectively been demonstrated through the automated speech analysis protocol. In terms of which acoustic measures could aid in the diagnosis of FTD in Dutch speakers, the present study supports the possibility that mean pause segment duration could function as a sensitive clinical biomarker, and that mean speech segment duration and pause rate could be of interest as well. At present, f0 range does not show the same promise yet, although further research is needed to establish these initial findings. Keeping above caveats in mind, the present study supports the feasibility of the automated speech analysis protocol as an ancillary tool for the diagnosis of FTD. Future studies would benefit from exploring the following directions: (1) an examination of the interaction between variables of group, acoustic measures, demographic and disease variables, and inclusion of and interaction with other linguistic levels, within the same cohort of subjects, including the earliest available recording of every subject; (2) a

cross-validation of the results gathered in Nevler et al., (2017; 2019; 2024) and the present study; (3) an investigation of language-specific effects on language breakdown in variants of FTD, and finally; (4) longitudinal follow-up analyses of the presymptomatic mutation carrier group using the automated speech analysis protocol.

Acknowledgements

To Dirk Vet, for setting up and running the Speech Activity Detector, and for writing the Praat script that enabled the retrieval of pitch values that were not defined by the Speech Activity Detector. To Marloes Roosingh, for their help identifying and correcting an issue in the R script used for the calculation of summarised acoustic measure outcomes. And to Thomas Tienkamp, for introducing me to the Tidyverse.

Bibliography

Ahmed, R. M., Hodges, J. R., & Piguet, O. (2021). Behavioural Variant Frontotemporal Dementia: Recent Advances in the Diagnosis and Understanding of the Disorder. *Frontotemporal Dementias*, *1281*, 1–15.

Bäckström, T., Räsänen, O., Zewoudie, A., Pérez Zarazaga, P., Koivusalo, L., Das, S., Mellado, E. G., Mansali, M. B., Ramos, D., Kadiri, S., & Alku, P. (2022). Introduction to speech processing (2nd ed.).

Boersma, P. & Weenink, D. (2024). *Praat: doing phonetics by computer* [Computer program], version [6.3.18], http://www.praat.org/

Breteler, A., & Schrijvers, E. M. C. (2009). Epidemiologie. *Handboek dementie*, 13–22. Bohn Stafleu van Loghum.

Chen, A., Chen, A., Kager, R., & Wong, P. (2014). Rises and falls in Dutch and Mandarin Chinese. In 4th International Symposium on Tonal Aspects of Language (pp. 83-86).

Duits, A., & Kessels, R. (2014). Schatten van het premorbide functioneren. *Neuropsychologische diagnostiek: De klinische praktijk*, 176-178.

Geraudie, Battista, P., García, A. M., Allen, I. E., Miller, Z. A., Gorno-Tempini, M. L., & Montembeault, M. (2021). Speech and language impairments in behavioral variant frontotemporal dementia: A systematic review. *Neuroscience and Biobehavioral Reviews*, *131*, 1076–1095.

Gorno-Tempini, Hillis, A. E., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M.,

Grossman, M., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., & Boeve, B. F. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014.

Hirst, & De Looze, C. (2021). Measuring Speech. Fundamental frequency and pitch. In *Cambridge Handbook of Phonetics* (Issue 1, pp. 336–361). Cambridge University.

Hogan, Jetté, N., Fiest, K. M., Roberts, J. I., Pearson, D., Smith, E. E., Roach, P., Kirk, A., Pringsheim, T., & Maxwell, C. J. (2016). The Prevalence and Incidence of Frontotemporal Dementia: a Systematic Review. *Canadian Journal of Neurological Sciences*, *43*(S1), S96–S109.

Jiskoot, L. C., Dopper, E. G. P., Heijer, T. den, Timman, R., van Minkelen, R., van Swieten, J. C., & Papma, J. M. (2016). Presymptomatic cognitive decline in familial frontotemporal dementia: A longitudinal study. *Neurology*, *87*(4), 384–391.

Jiskoot, L. C., Panman, J. L., van Asseldonk, L., Franzen, S., Meeter, L. H. H., Donker Kaat, L., van der Ende, E. L., Dopper, E. G. P., Timman, R., van Minkelen, R., van Swieten, J. C., van den Berg, E., & Papma, J. M. (2018). Longitudinal cognitive biomarkers predicting symptom onset in presymptomatic frontotemporal dementia. *Journal of Neurology*, *265*(6), 1381–1392.

Josephs, Duffy, J. R., Strand, E. A., Whitwell, J. L., Layton, K. F., Parisi, J. E., Hauser, M. F., Witte, R. J., Boeve, B. F., Knopman, D. S., Dickson, D. W., Jack, C. R., & Petersen, R. C. (2006). Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. *Brain (London, England : 1878)*, *129*(6), 1385–1398.

Kassambara, A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr

Kassambara, A. (2022). rstatix: Pipe-friendly Framework for Basic Statistical Tests. https://CRAN.R-project.org/package=rstatix

de Koning, I. & Schmand, B. (2009). Neuropsychologisch onderzoek. *Handboek dementie* (33–42). Bohn Stafleu van Loghum.

Lev-Ari, S., Amsalem, S., Geva, S., Shalvi, D., & Feiglin, T. (2024). ggrain: Grain size distribution and sedimentary process analysis. https://CRAN.R-project.org/package=ggrain

Linguistic Data Consortium. (n.d.). LDC Broad Phonetic Class Speech Activity Detector (ldc-bpcsad) command line interface (CLI) documentation. Retrieved from https://linguistic-data-consortium.github.io/ldc-bpcsad/cli.html

Litvan. (2005). What is an Atypical Parkinsonian Disorder? *Atypical Parkinsonian Disorders* (pp. 1–9). Humana Press.

McKee, A. C., Stein, T. D., Kiernan, P. T., & Alvarez, V. E. (2015). The Neuropathology of Chronic Traumatic Encephalopathy. *Brain Pathology (Zurich, Switzerland)*, *25*(3), 350–364.

Marcoux, K., & Ernestus, M. (2019, August). Pitch in native and non-native Lombard speech. Paper presented at the 19th International Congress of Phonetic Sciences (ICPhS 2019), Melbourne, Australia (pp. 2605-2609). Australasian Speech Science and Technology Association Inc.

Mesulam, Coventry, C., Bigio, E. H., Geula, C., Thompson, C., Bonakdarpour, B., Gefen, T., Rogalski, E. J., & Weintraub, S. (2021). Nosology of Primary Progressive Aphasia and the Neuropathology of Language. In *Frontotemporal Dementias* (Vol. 1281, pp. 33–49). Springer International Publishing.

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes. https://CRAN.R-project.org/package=RColorBrewer

Nevler, Ash, S., Irwin, D. J., Liberman, M., & Grossman, M. (2019). Validated automatic speech biomarkers in primary progressive aphasia. *Annals of Clinical and Translational Neurology*, *6*(1), 4–14.

Nevler, Ash, S., Jester, C., Irwin, D. J., Liberman, M., & Grossman, M. (2017). Automatic measurement of prosody in behavioral variant FTD. *Neurology*, *89*(7), 650–656.

Nevler, N., Cho, S., Cousins, K. A. Q., Ash, S., Olm, C. A., Shellikeri, S., Agmon, G., Gonzalez-Recober, C., Xie, S. X., Barker, M. S., Manoochehri, M., Mcmillan, C. T., Irwin, D. J., Massimo, L., Dratch, L., Cheran, G., Huey, E. D., Cosentino, S. A., Van Deerlin, V. M., … Grossman, M. (2024). Changes in Digital Speech Measures in Asymptomatic Carriers of Pathogenic Variants Associated With Frontotemporal Degeneration. *Neurology*, *102*(2).

Nevler, N., Cho, S., Cousins, K. A. Q., Ash, S., Olm, C. A., Shellikeri, S., Agmon, G., Gonzalez-Recober, C., Xie, S. X., Barker, M. S., Manoochehri, M., Mcmillan, C. T., Irwin, D. J., Massimo, L., Dratch, L., Cheran, G., Huey, E. D., Cosentino, S. A., Van Deerlin, V. M., … Grossman, M. (2024). Changes in Digital Speech Measures in Asymptomatic Carriers of Pathogenic Variants Associated With Frontotemporal Degeneration. *Neurology, 102*(2). Supplement: Mutation specific sub-analyses.

Nevler, N. (2022). UPenn-FTDC-Acoustic-pipeline [GitHub repository]. Retrieved from https://github.com/naomi-nevler/UPenn-FTDC-Acoustic-pipeline

Pakhomov, Smith, G. E., Marino, S., Birnbaum, A., Graff-Radford, N., Caselli, R., Boeve, B., & Knopman, D. S. (2010). A computerized technique to assess language use patterns in patients with frontotemporal dementia. *Journal of Neurolinguistics*, *23*(2), 127–144.

Pakhomov, Smith, G. E., Chacon, D., Feliciano, Y., Graff-Radford, N., Caselli, R., & Knopman, D. S. (2010). Computerized analysis of speech and language to identify psycholinguistic correlates of frontotemporal lobar degeneration. Cognitive and Behavioral Neurology, 23(3), 165–177.

Plack, C. J., Oxenham, A. J., & Fay, R. R. (Eds.). (2006). *Pitch: neural coding and perception* (Vol. 24). Springer Science & Business Media.

Praat. (n.d.). Get_pitch_quantiles_duration_5.praat

R Core Team. (2023). R (v.4.3.3). R Foundation for Statistical Computing. https://www.R-project.org/

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., ... & Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, *134*(9), 2456-2477.

Rascovsky, & Grossman, M. (2013). Clinical diagnostic criteria and classification controversies in frontotemporal lobar degeneration. *International Review of Psychiatry (Abingdon, England)*, *25*(2), 145–158.

Riedl, Mackenzie, I. R., Förstl, H., Kurz, A., & Diehl-Schmid, J. (2014). Frontotemporal lobar degeneration: Current perspectives. *Neuropsychiatric Disease and Treatment*, *10*, 297–310.

Rohrer, Sauter, D., Scott, S., Rossor, M. N., & Warren, J. D. (2012). Receptive prosody in nonfluent primary progressive aphasias. *Cortex*, *48*(3), 308–316.

Rohrer, J. D., Warren, J. D., Fox, N. C., & Rossor, M. N. (2013). Presymptomatic studies in genetic frontotemporal dementia.

Rohrer, Nicholas, J. M., Cash, D. M., van Swieten, J., Dopper, E., Jiskoot, L., van Minkelen, R., Rombouts, S. A., Cardoso, M. J., Clegg, S., Espak, M., Mead, S., Thomas, D. L., De Vita, E., Masellis, M., Black, S. E., Freedman, M., Keren, R., MacIntosh, B. J., … Laforce, R. (2015). Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the Genetic Frontotemporal dementia Initiative (GENFI) study: a cross-sectional analysis. *Lancet Neurology*, *14*(3), 253–262.

Rohrer, J.D. & Boxer, A. L. (2021). The Frontotemporal Dementia Prevention Initiative: Linking Together Genetic Frontotemporal Dementia Cohort Studies. *Frontotemporal Dementias*, *1281*, 113–121.

Rosness, Haugen, P. K., Passant, U., & Engedal, K. (2008). Frontotemporal dementia--a clinically complex diagnosis. *International Journal of Geriatric Psychiatry*, *23*(8), 837–842.

RStudio Team. (2023). Rstudio (RStudio-2023.12.1-402). RStudio, PBC. https://www.rstudio.com/.

Ryant, N. (2013). LDC HMM Speech Activity Detector, v. 1.0. 4. *University of Pennsylvania*.

Ryant, N. (2023). Linguistic Data Consortium Broad Phonetic Class Speech Activity Detector (ldc-bpcsad) [GitHub repository]. Retrieved from https://github.com/Linguistic-Data-Consortium/ldc-bpcsad

Seelaar, Pijnenburg, Y. A. L., & van Swieten, J. C. (2009). Frontotemporale dementie (FTD). *Handboek dementie* (pp. 211–218). Bohn Stafleu van Loghum.

Swinburn, K., Porter, G., Howard, D., & Swinburn, Kate. (2014). *CAT - NL : comprehensive aphasia test* (E. G. Visch-Brink, D. Vandenborre, H. J. De Smet, & P. Mariën, Eds.). Pearson.

Varma, Snowden, J. S., Lloyd, J. J., Talbot, P. R., Mann, D. M. A., & Neary, D. (1999). Evaluation of the NINCDS-ADRDA criteria in the differentiation of Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery and Psychiatry*, *66*(2), 184–188.

Verhey, F. & Pijnenburg, Y. (2009). Klinische Diagnostiek. *Handboek dementie* (pp. 25–32). Bohn Stafleu van Loghum.

Vet, D. J. (2024). CalcQuantilesOfUndefined.script [Software]. University of Amsterdam.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., … Yutani, H. (2022). Tidyverse: Easily install and load the 'Tidyverse'. https://CRAN.R-project.org/package=tidyverse

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag. https://ggplot2.tidyverse.org

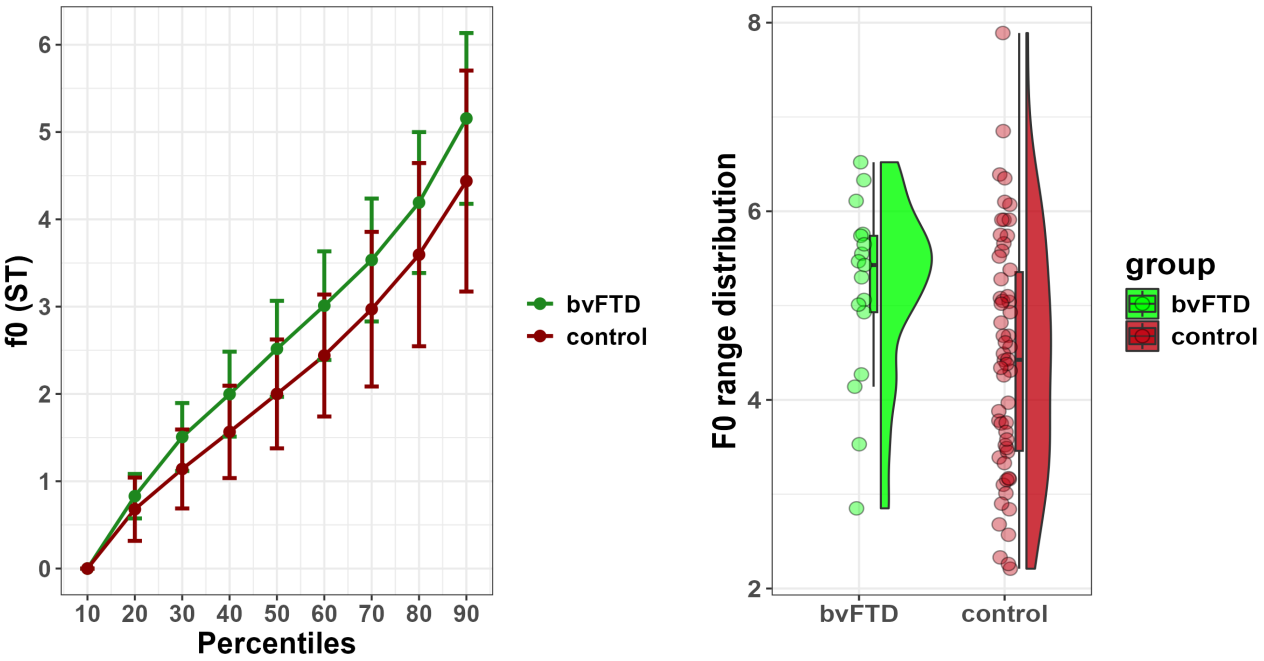
Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation. https://CRAN.R-project.org/package=dplyr

Willems, N. (2010). English intonation from a Dutch point of view. Walter de Gruyter.

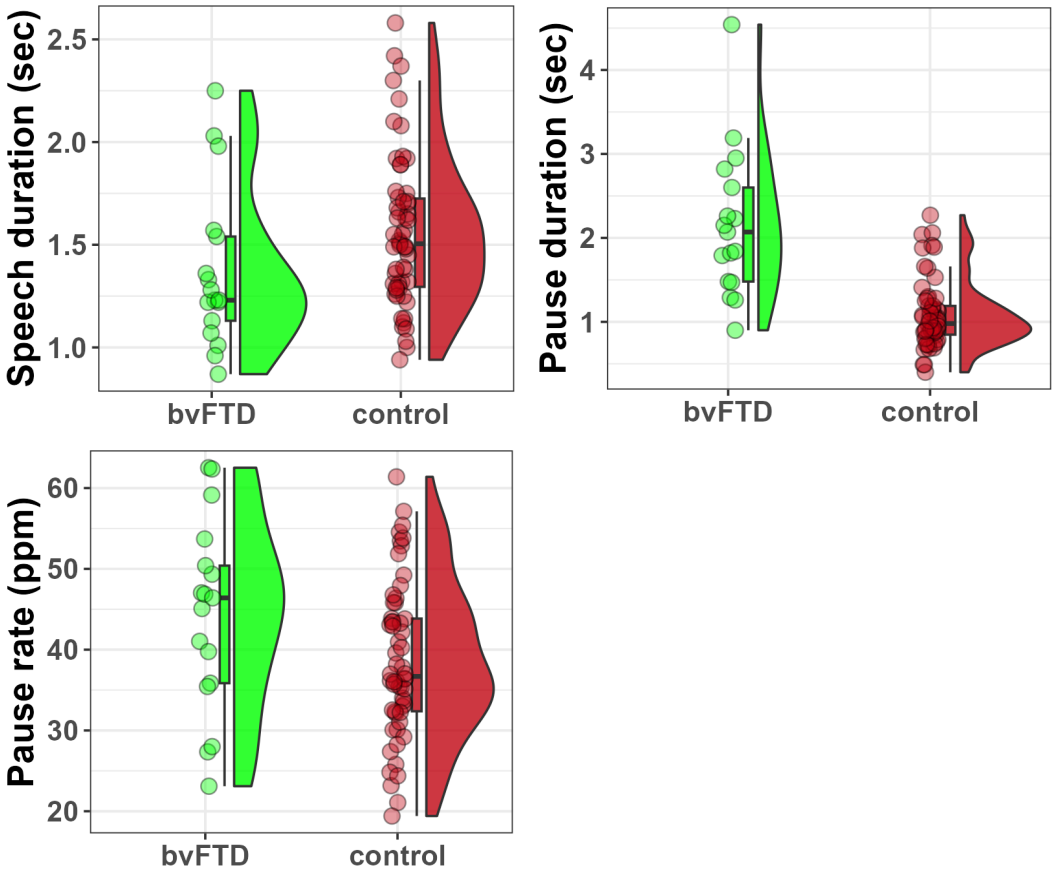
Appendix: List of Figures

Research question 1: Comparison between patients with bvFTD and controls on measures of f0 range, mean speech segment duration, mean pause segment duration, and pause rate:

F0 percentiles (left) and f0 range distribution (right)

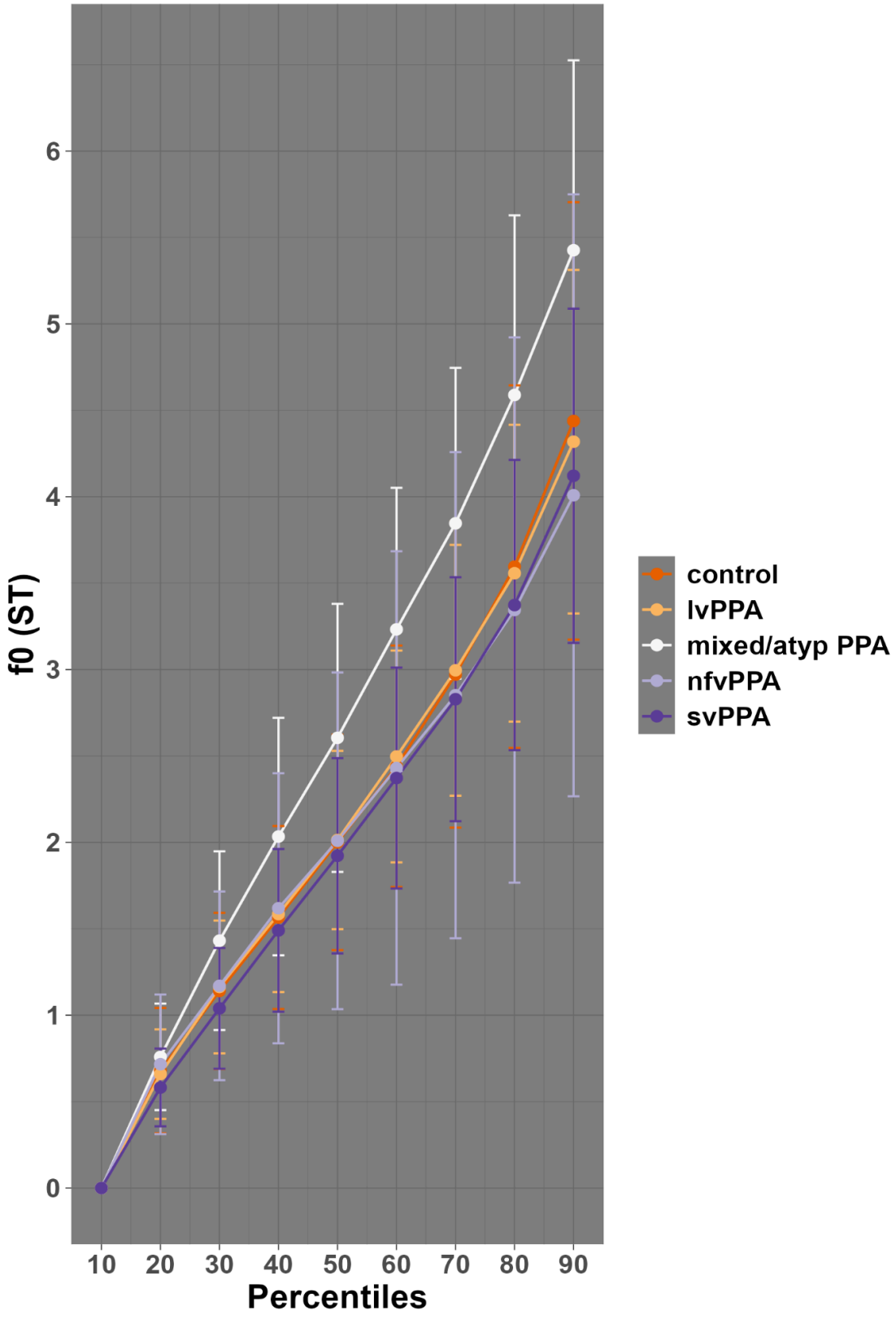


Mean speech segment duration (top left), mean pause segment duration (top right), and pause rate (bottom left)

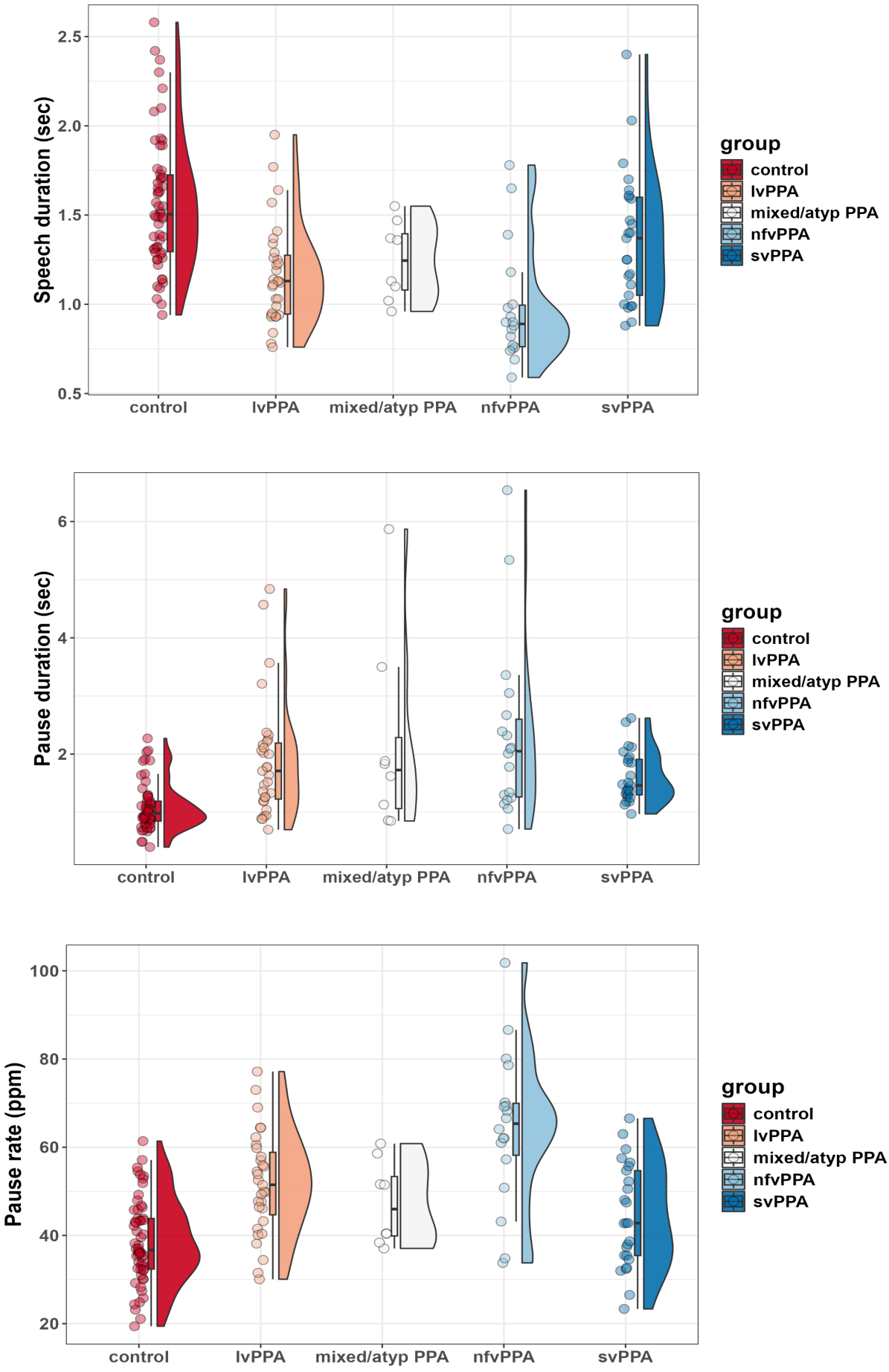


Research question 2: Comparison between patients with lvPPA, svPPA, mixed/atypical PPA, and nfvPPA on the one hand, and controls on the other, on measures of f0 range, mean speech segment duration, mean pause segment duration, and pause rate:

F0 percentiles per group

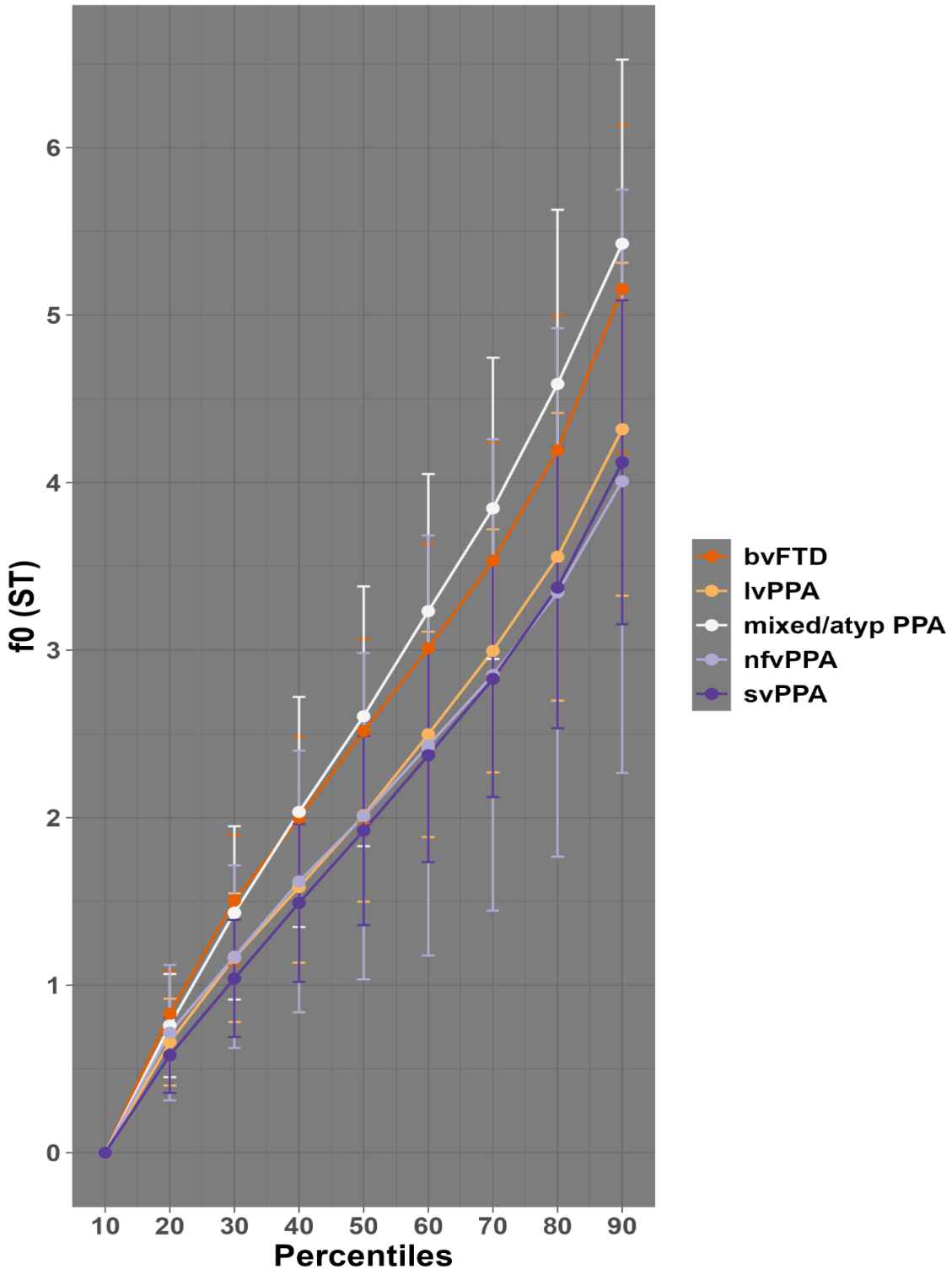


Mean speech segment duration (top), mean pause segment duration (middle), and pause rate (bottom) per group

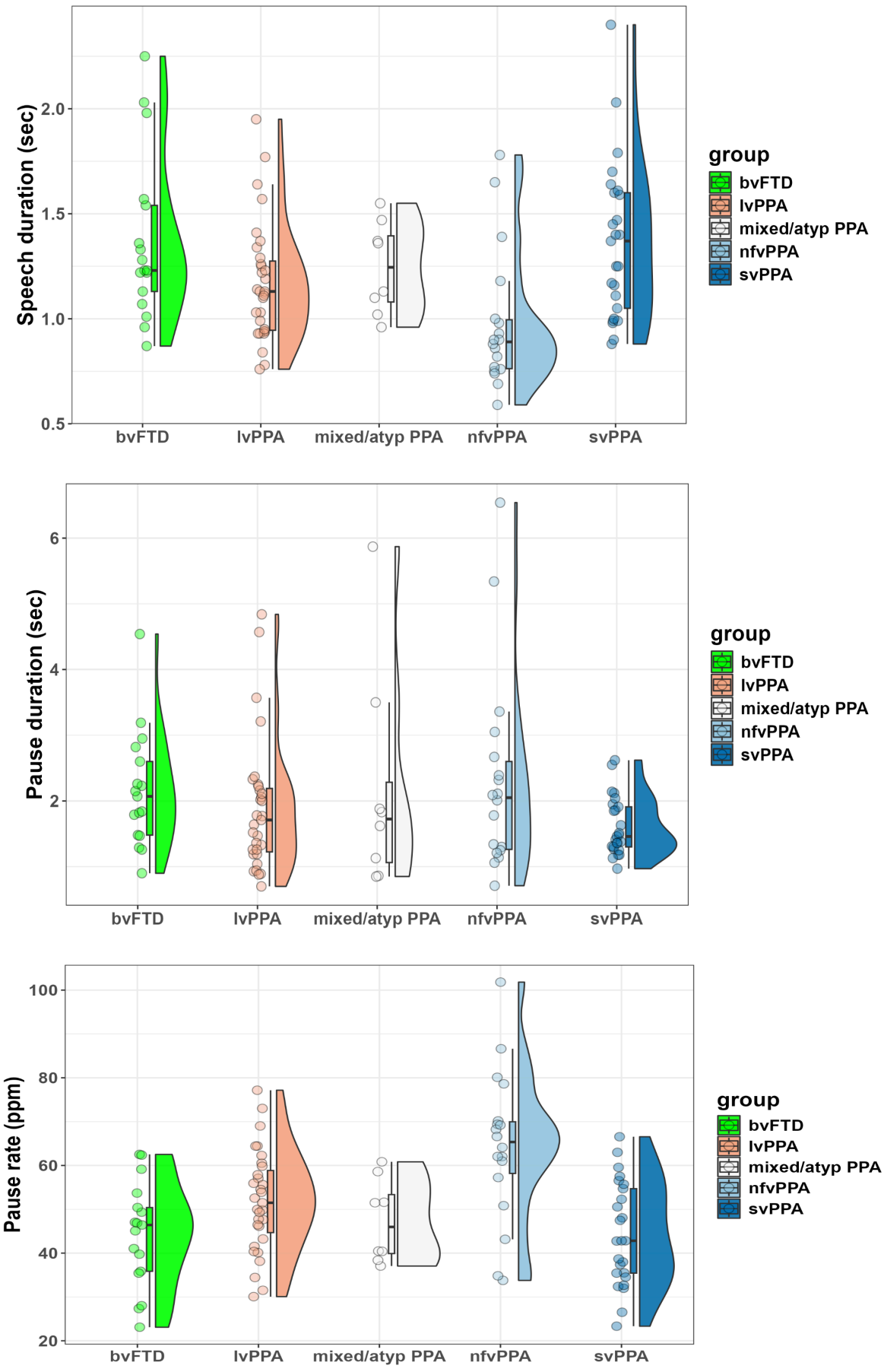


Research question 3: Comparison between patients with lvPPA, svPPA, mixed/atypical PPA, nfvPPA, and bvFTD, on measures of f0 range, mean speech segment duration, mean pause segment duration, and pause rate:

F0 percentiles per group

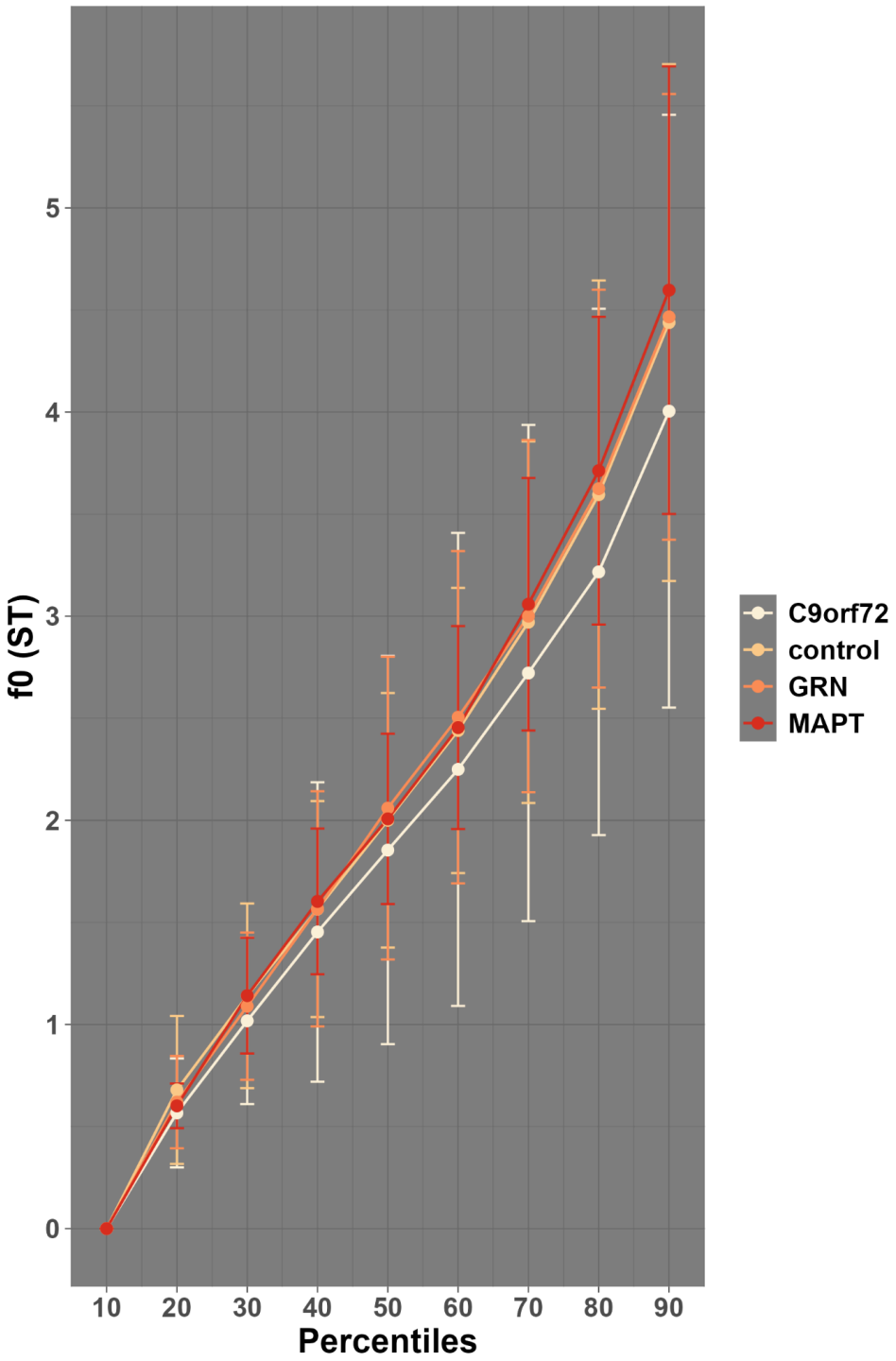


Mean speech segment duration (top), mean pause segment duration (middle), and pause rate (bottom) per group



Research question 4+5: Comparison between presymptomatic mutation carriers (*MAPT, GRN, and C9orf72*) and controls, and between subtypes of presymptomatic mutations, on measures of f0 range, mean speech segment duration, mean pause segment duration, and pause rate:

F0 percentiles per group



Mean segment duration (top), mean pause segment duration (middle), and pause rate (bottom) per group