The Effect of Levodopa and Fatigue on Dysarthric Speech of Slovenian Parkinson's Disease Patients

Teja Rebernik

university of groningen / faculty of arts

The Effect of Levodopa and Fatigue on Dysarthric Speech of Slovenian Parkinson's Disease Patients

Teja Rebernik

Master thesis
for the requirements of the programme
Research Master Language and Cognition

Word count: 32,526
Student number: S2957191
Supervisor: Prof. Dr. Martijn Wieling
Second reader: Dr. Dicky Gilbers

June 2019

Abstract

Parkinson's disease (PD) is the second most common neurodegenerative disease in the world, affecting millions of people. It is primarily a motor disorder, presenting with motor symptoms such as tremor and problems with gait, however PD patients also often face speech problems, also known as hypokinetic dysarthria. Symptoms of the latter include monopitch, monoloudness, imprecise articulation of vowels and consonants, diminished prosody, and a breathy, harsh voice. While there is no cure for PD, the drug that is used for treating and relieving motor symptoms is levodopa. However, it is currently unclear how exactly levodopa affects speech, as previous studies have shown both detrimental and beneficial effects. The aim of the present study was to determine the effect of levodopa and fatigue on several speech parameters, including two measures of vowel articulation (vowel space area and vowel articulation index) and two measures of voice quality (fundamental frequency and cepstral peak prominence smoothed). 10 native speakers of Slovenian, namely 6 patients with Parkinson's disease and 4 healthy controls, recorded their speech with a headset microphone on twenty occasions, performing four tasks each time.

The results indicate that, compared to healthy control speakers, PD patients show no significant differences in the measured acoustic parameters, although there is a trend towards a smaller triangular vowel space area and lower vowel articulation index, indicating reduced vowels' articulation, as well as lower measures of cepstral peak prominence and lower fundamental frequency, indicating pathological voice quality. Group does not significantly affect speech ($p = 0.54$, $d = -0.5$), however gender does ($p < 0.001$, $d = -8.4$). When analysing the effect of levodopa on speech of PD patients, there were no significant differences between OFF and ON states ($p = 0.6$, $d = 0.04$) nor between OFF, 1 hour and 2 hours after intake ($p = 0.57$, $d = 0.04$). No other variables, including fatigue, time of day or task, significantly affected the measured acoustic parameters.

There is large individual variability across subjects for all measures. The results of the present study have implications for future studies, as they highlight the importance of studying the speech of PD patients at more than one moment in time, choosing a homogenous sample of PD patients and finding age- and gender-matched healthy control speakers.

Keywords: Parkinson's disease; vowel articulation; CPPS; fundamental frequency; Slovenian language; levodopa.

Acknowledgments

The first draft of this section included only the words “Thank you”, in large block letters. The spirit remains the same, the execution a bit lengthier.

First, I would like to thank Trepetlika, the Parkinson’s Disease Society of Slovenia. Without their help, I never would have been able to recruit those suffering from Parkinson’s disease, and this thesis would not exist. A big thank you also to my participants, who recorded hours upon hours of their speech.

Thank you, Martijn, for including me in (what has now become) Speech Lab Groningen and for your (incredibly fast) comments on this thesis, which have made it so much better. As you say: you realize this means you have to work with me for the next five years, right?

Thank you, Dicky, for being my introduction into linguistic (and musical) research when I began my studies. I greatly enjoy our discussions and I’m glad we will work on our topic for at least a while longer.

Thank you, Jidde, for being so welcoming last winter. If you let a shark eat you in Australia, I will never forgive you. Thank you, Lisanne, for being the best lab partner I could have asked for.

Thank you, Ciara and Marjolein*, for being my girls.

Thank you, Casper, for being there. With me and for me. No takebacks.

And most importantly, thank you to my family. *Hvala za podporo, več kot tisoč kilometrov stran. Pogrešam vas in rada vas imam.*

To conclude, I want to say this: This thesis was long in coming. Winter turned into spring turned into summer, and finally here we are. In the immortal words of Douglas Adams:

*“I love deadlines. I like the whooshing sound they make as they fly by.***”*

*The order of names is (not entirely) alphabetical. Ciara is first because of Toeter.

**But I promise I won’t be this late again.

Table of contents

List of tables

List of figures

List of appendices

1 Introduction

Parkinson's disease is the second most common neurodegenerative disease in the world, affecting nearly 10 million people across the globe (Parkinson's Foundation, 2019) – a number that is expected to rise substantially (Dorsey et al., 2007). While Parkinson's disease (PD) is not limited to a certain age group and younger individuals can suffer from it, it does predominantly affect the elderly, with 1 in 100 people above the age of 60 receiving the PD diagnosis (Tysnes and Storstein, 2017). The cause of PD is largely unknown (de Lau and Breteler, 2006), i.e. it is an idiopathic condition, but it is characterized by common motor and non-motor symptoms. Motor symptoms of PD among others include rest tremor, slowness of movement, rigidity and gait imbalance, while most common non-motor symptoms include sleep disorders, reduced ability of smell, and depression (Marsden, Parkes and Quinn, 1981; Jankovic, 2007; Tysnes and Storstein, 2017; Tropea and Chen-Plotkin, 2018).

Another important symptom of Parkinson's disease is deteriorating speech. Although the exact percentage of patients whose voice is affected by the disease differs depending on the study, it is known that approximately three quarters of patients experience some sort of voice problems at some stage in the disease (Ho et al., 1998; Logemann, Fisher, Bosher and Blonsky, 1978; Hartelius and Svensson, 1994). Speech problems occurring in PD are known under the umbrella term hypokinetic dysarthria, which is characterized by diminished prosody, imprecise articulation, harsh and breathy voice, and decreased intensity (Darley, Aronson and Brown, 1969; Walsh and Smith, 2012; Ho et al., 1998; Harel et al., 2004; Brabenec, Mekyska, Galaz and Rektorova, 2017). Speech problems can appear both in early and late stage PD, but voice quality and speech performance have been found to deteriorate across the disease stage (Holmes, Oates, Phyland and Hughes, 2000; Skodda, Grönheit, Mancinelli and Schlegel, 2013).

Speech in PD has been frequently studied, but many unknowns remain. One of the main questions is how PD medication affects speech. Even though PD has no cure, most patients take the drug levodopa to treat and relieve their symptoms. Since its introduction in the 1960s, levodopa has been the gold standard in PD (Jankovic, 2007), but its effect on speech remains unclear. While some studies have shown an improvement of speech symptoms after levodopa intake (Sanabria et al., 2001; Ho et al., 2008; Wolfe et al., 1975), others have shown either no change (Plowman-Prine et al., 2009; Goberman, Coelho and Robb, 2002; de Letter, Santens, Bodt, Boon and Borsel, 2006) or even a worsening of symptoms (Louis, Winfield, Fahn and

Ford, 2001). Interpretation of study results is further complicated because other factors, for example fatigue, influence PD patients' speech. Furthermore, different studies use different tasks (e.g. read vs. spontaneous speech), measure different parameters (e.g. voice quality vs. articulation), and test at differently defined periods.

The present thesis will investigate the effect of levodopa and fatigue on dysarthric speech of Slovenian patients with Parkinson's disease. To answer our research question, we will recruit Slovenian participants (PD patients and their partners) with help of Parkinson's Disease Association of Slovenia. They will record their speech on four separate days, five sessions each day, the exact time of which will be defined depending on the patients' levodopa intake. During each session, a fatigue survey will be filled out, followed by four speaking tasks that will be recorded with headset microphones. Several acoustic parameters, including vowel space area and voice quality, will be measured, taking both levodopa and fatigue into account. The study is a replication of a pilot study conducted with Dutch PD patients within the Speech Lab Groningen and is innovative for several reasons. First, it thoroughly investigates Slovenian parkinsonian speech from an acoustic perspective, which has not been done before. Second, it measures speech on twenty different occasions, allowing for large amounts of speech samples to be collected and compared in the same sample of patients. Finally, the recordings are made by the participants themselves in the comfort of their own homes, without the experimenter present, which reduces stress levels and improves ecological validity.

The following sections will first provide the theoretical underpinnings behind the thesis topic (Chapter 2), starting by examining the pathophysiology, symptoms, causes and treatments for idiopathic Parkinson's disease. This will be followed by an overview of speech problems in Parkinson's disease, including what is currently known about the effect of levodopa on articulation, prosody, and voice quality. As the language studied in the thesis is Slovenian, the third part of Chapter 2 will introduce the general and phonological characteristics of the Slovenian language. The chapter will be concluded with a section on the importance of doing cross-linguistic research in the field of clinical linguistics and speech motor control problems. The subsequent chapter (Chapter 3) will present research questions and hypotheses, and their justifications. This will be followed by a presentation of the method (Chapter 4), including the recruitment of participants, the equipment and experimental set-up, stimuli design, and experimental tasks.

Chapter 5 will present the analysis performed on the obtained speech samples, including pre-processing in Adobe Audition and PRAAT, speech analysis in MATLAB, and statistical analysis in RStudio. Following is Chapter 6, in which the results are presented, including the differences between PD patients and healthy control speakers as well as the effect of levodopa on PD patients' speech. These results are subsequently discussed in Chapter 7 in terms of our research questions and hypotheses. The thesis concludes with a reflection on what the study has shown, to what extent researchers should consider the levodopa intake when investigating parkinsonian speech, what the potential limitations of our study were, and how to proceed with further research on speech in Parkinson's disease.

2 Theoretical background

The present section will first discuss Parkinson's disease (PD), in terms of causes and risk factors, prevalence, symptoms and treatment. This will be followed by a section on general characteristics of PD speech, seminal studies on PD speech that this research study leans on, and an overview of studies that have so far investigated the effect of levodopa on speech. Third, Slovenian speech characteristics, including general characteristics and dialects, phonology and previous studies on Slovenian parkinsonian speech will be discussed. Finally, the chapter is concluded with a short section highlighting on the importance of conducting cross-linguistic research in a time when so many studies are carried out on native English speakers.

2.1 Idiopathic Parkinson's Disease

2.1.1 What is Parkinson's disease?

Parkinson's disease is the second most common neurodegenerative neurological disorder, meaning a disorder characterized by a loss of neuronal (i.e. brain) cells. Specifically, PD is characterized by a loss of dopamine neurons in the substantia nigra in the basal ganglia (Factor and Weiner, 2008), a brain region involved in planning and executing movements (Albin, Young and Penney, 1989), which also includes the movements necessary for producing fluent speech. It is a multisystem disorder, spreading from the gut and the nose to the central nervous system (Klingelhoefer and Reichmann, 2017), leading to a wide range of symptoms spanning from motor to non-motor.

Parkinson's disease was first described by James Parkinson in his influential text from 1817 entitled *An Essay on the Shaking Palsy*. Since then, the high prevalence and fast growth of Parkinson's disease (see Section 2.1.2 below) has not only gained the interest of researchers but also the attention of the public. This is also partially due to celebrities spreading awareness of the disease, most famously Michael J. Fox from the blockbuster film *Back to the Future*, who established his own foundation dedicated to finding the cure for Parkinson's disease (The Michael J. Fox Foundation, 2019).

Despite this public and scientific interest, however, Parkinson's disease remains difficult to diagnose and treat. Already in his 1817 essay, James Parkinson recognized that diagnosing Parkinson's disease is no easy task. He stated that

> "so slight and nearly imperceptible are the first inroads of this malady, and so extremely slow its progress, that it rarely happens, that the patient can form any recollection of the precise period of its commencement."
>
> (Parkinson, 1817, p. 224)

This remains true today, as there are still no diagnostic tests or markers that would clearly, easily and reliably diagnose PD (De Lau and Breteler, 2006). Scientists have thus been led on a search for biological markers that could serve as disease indicators, from biochemical processes (Tropea and Chen-Plotkin, 2018) to speech (Harel et al., 2004). For now, however, post-mortem confirmation is still required for definitive PD diagnosis (De Lau and Breteler, 2006).

Even more, the underlying causes remain largely unknown (De Lau and Breteler, 2006), i.e. it is an idiopathic condition (wherefrom the name *idiopathic* Parkinson's disease). While there are some *genetic* risk factors, they are identified only in a small number of PD patients (Tysnes and Storstein, 2017; De Lau and Breteler, 2006). Instead, researchers have identified several *environmental* risk factors, including dietary particularities, head injury, and exposure to pesticides, herbicides and heavy metals (Connolly and Lang, 2014; De Lau and Breteler, 2006). Interestingly, there are also several protective factors, such as cigarette smoking[1] and high coffee consumption (Tysnes and Storstein, 2017; De Lau and Breteler, 2006; Connolly and Lang, 2014), and several pre-symptomatic risk factors that can act as a predictor of whether an individual will develop Parkinson's disease. Most commonly these include REM-sleep disorders, constipation and reduced ability of smell (Tysnes and Storstein, 2017; Klingelhoefer and Reichmann, 2017; Connolly and Lang, 2014). Finally, PD seems to be more prevalent in men than women (e.g. Connolly and Lang, 2014; Tysnes and Storstein, 2017), which has led to the belief that oestrogen has a neuroprotective effect (Miller and Cronin-Golomb, 2010).

2.1.2 Prevalence of Parkinson's disease

PD is not only the second most common neurodegenerative disorder in the world behind Alzheimer's disease (Schapira, 2009), but also the *fastest growing* one (Dorsey and Bloem, 2017). It is not limited to a single region of the world nor to a certain age group, although it does seem to have a lower prevalence in Asia compared to Europe and the likelihood of having PD increases in the elderly population (Pringsheim et al., 2014). It affects between 6 and 10

[1] The beneficial effect of nicotine has been frequently studied, as smokers are 50% less likely to develop Parkinson's disease (Dorsey and Bloem, 2017). However, this benefit of nicotine remains under discussion, as the results could also stem from higher morbidity rates of smokers (De Lau and Breteler, 2006), i.e. smokers do not live long enough to develop PD.

million people worldwide, depending on the source (Dorsey et al., 2018; Parkinson's Foundation, 2019), but this number is rising substantially every year (Dorsey et al., 2007). PD affects 0.1-0.2% of individuals in a population at any given time and 1% in a population above 60 years of age (Tysnes and Storstein, 2017). 0.3% of the total population suffers from PD (Connolly and Lang, 2014). This prevalence, however, can rise to up to 4% in the highest age groups (ibid.).[2]

However, the reported numbers about PD prevalence differ and are not entirely reliable, the reasons for which are manifold. First, the definitions of Parkinson's disease (also in comparison to other parkinsonisms[3]) has changed over time (Feigin et al., 2017). Second, PD prevalence remains underreported (Dorsey and Bloem, 2017). Finally, due to difficulties in making a reliable diagnosis, individuals are often not diagnosed properly (Dorsey et al., 2007).

In Slovenia, the number of Parkinson patients is estimated differently depending on the source, however the current consensus is that there are currently around 4000 (Campolounghi-Pegan, 2008) or 5000 (Trošt, 2008) individuals suffering from PD and 7000 suffering from PD *including* other parkinsonisms (Trepetlika, 2019).[4] At 2 million inhabitants, this means that 2.5 inhabitants per 1000 in Slovenia suffer from idiopathic Parkinson's disease, which is in line with other studies reporting global PD prevalence.

2.1.3 Symptoms of Parkinson's Disease

Parkinson's disease is characterized by both motor and non-motor symptoms. The cardinal, primary, motor symptoms of PD are described by the acronym *TRAP*: *T*remor at rest, *R*igidity, *A*kinesia (meaning impairment in voluntary movement), and *P*ostural instability (Jankovic, 2007; Tysnes and Storstein, 2017). Depending on the main cardinal symptom of the patient, he or she can fall into one of the three main groups: tremor-dominant (8% of patients), akinetic-rigid (26%) and mixed (66%) (Connolly and Lang, 2014). Besides these primary motor symptoms, there are secondary motor symptoms, including hypomimia (reduced facial

[2] It is not entirely certain how age and PD interact: while some studies claim that the older the person, the more likely they are to have PD, others claim that PD incidence hits its peak in the age group of 70-79 and declines afterwards (Pringsheim et al., 2014).
[3] Parkinsonisms are disorders that cause similar symptoms to those of Parkinson's disease, such as progressive supranuclear palsy (PSP) and multiple system atrophy (MSA) (Parkinson's Foundation, 2018).
[4] This number, however, comes from 2008 or even earlier. Considering the rise in PD prevalence, it is possible that a decade later, there are more PD patients.

expression), hypokinetic dysarthria (speech problems, described below), shuffling gait and others (Jankovic, 2007).

Non-motor symptoms of PD include behavioural dysfunction, such as depression, anxiety, dementia and psychosis; autonomic dysfunction, such as dysphagia (swallowing difficulties), gastric dysfunction, constipation, and impaired sexual function; various sleep-related dysfunctions, such as insomnia and sleep apnea; and sensory dysfunctions, such as hyposmia (reduced ability to smell) (Pfeiffer and Bodis-Wollner, 2005; Tysnes and Storstein, 2017; Jankovic, 2007; Pavšič and Pirtošek, 2015).

Parkinson's disease is not only characterized by its symptoms, but also diagnosed by them. The criteria for diagnosing PD includes the presence of at least two cardinal symptoms (De Lau and Breteler, 2006). Furthermore, a good measure of PD is asymmetric symptom onset (i.e. the symptoms first appear on one side of the body), a good response to the drug levodopa (described below in Section 2.1.4), and olfactory loss (Tysnes and Storstein, 2017). A patient cannot be diagnosed with PD if, amongst others, he or she shows any cerebellar abnormalities, early gait impairment, bilateral symptoms, and absence of common non-motor symptoms (Tysnes and Storstein, 2017).

As the disease progresses, the symptoms get worse (Jankovic, 2007) and it takes years for the disease to reach its full extent (Braak et al., 2004). While primary motor symptoms such as tremor and bradykinesia appear at diagnosis, other symptoms (e.g. swallowing problems or psychiatric disturbances) only appear 5 – 10 years after symptom onset (Connolly and Lang, 2014). At the beginning, the disease causes fast deterioration in bradykinesia, rigidity and activities of daily living (Maetzler, Liepelt and Berg, 2009), but the exact rate of deterioration depends on the patient and tends to be more rapid in patients who show postural instability and gait difficulties (Jankovic, 2007).

PD symptoms are commonly assessed with two scales, namely the Hoehn and Yahr scale, which is used to provide an assessment of disease progression (Müller et al., 2000), and the United Parkinson's Disease Rating Scale (UPDRS), which assesses disability and impairment (Martinez-Martin et al., 1994), although the latter has been contested to be culturally- and gender-biased (Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, 2003).

Assessing symptoms is additionally complex due to a large symptomatic variation (Marsden, Parkes and Quinn, 1984). Sex, age at which PD begins, and stage of disease have all been shown to influence the presentation of symptoms. Women, additionally, show slower disease progression (Pringsheim et al., 2014), but the presentation of symptoms in women unfortunately remains understudied (Miller and Cronin-Golomb, 2010). Many other variables also influence the severity and appearance of symptoms, including the patients' current emotional state (Jankovic, 2007) as well as changes in environment and drug therapy (Marsden, Parkes and Quinn, 1984).

2.1.4 Treatment

There is no cure for Parkinson's disease, however there are symptomatic therapies that can improve the patients' quality of life (Connolly and Lang, 2014). For the motor symptoms of PD, doctors prescribe levodopa and dopamine agonists[5]. Levodopa has been the gold standard for relieving the symptoms of PD since its development in the 1960s (Jankovic, 2007), but while the primary motor symptoms (e.g. tremor) respond well to it, axial motor symptoms (e.g. falls, postural instability) are treatment-resistant (Connolly and Lang, 2014). Non-motor symptoms, likewise, are treatment resistant, so individual therapies have been developed, for example the Lee Silverman Voice Therapy for improving dysarthric speech.

Due to the treatment-resistance of axial motor symptoms and due to the dyskinesias (i.e. involuntary rhythmic movements) caused by levodopa, dopamine agonists are often prescribed at the beginning stages of the disease, as they are equally efficacious and less likely to cause motor complications and side effects (Connolly and Lang, 2014). Consequently, if the patient is younger than 60, the first prescribed treatment are dopamine agonists, followed by levodopa (ibid.). This is especially the case because younger patients are more likely to develop levodopa-induced dyskinesias than older patients (Jankovic, 2007). In the later stages of the disease, dopamine agonists are sometimes prescribed to reduce the OFF time between two levodopa treatments (ibid.).

The peak level of levodopa in the blood occurs at about 1 hour after the intake of the drug (Kempster et al., 1989) but the responses are not always stable, as they are dependent on several factors. First, individuals can vary in their drug absorption rate (Olanow, Gauger and

[5] Dopamine agonists *enhance* already existing dopamine in the brain, while levodopa replaces the dopamine that is not there anymore (Schulz and Grant, 2000).

Cedarbaum, 1991). Second, fluctuating responses to levodopa can be attributed to disease stage and drug therapy (Goberman and Coelho, 2002; De Letter et al., 2010). Finally, environmental factors, such as time of day, anxiety level and fatigue, play an important role in how the patient will respond to the medication (Marsden, Parkes and Quinn, 1981).

An additional factor when taking levodopa and studying the patients' responses is the appearance of so-called ON/OFF symptoms. The ON state refers to periods where the patient experiences a good response to medication, while the OFF state refers to periods when the medication wears off and symptoms re-emerge (Connolly and Lang, 2014). This can happen suddenly, without any changes to the treatment schedule (Djaldetti and Melamed, 1998) and has been compared to someone flicking a light switch ON and OFF.

Despite the common misconception of patients, levodopa does *not* lose efficacy after 5 years, although there is a greater risk for dyskinesias, dystonia, and ON/OFF effects (Connolly and Lang, 2014; De Letter, Santens and van Borsel, 2005). Instead, it remains relatively stable during that period, however patients do start experiencing motor performance fluctuations (Goberman, Coelho and Robb, 2005), which can lead to an impression of decreased efficacy. When levodopa does lose efficacy, tremor-dominant patients undergo deep brain stimulation (DBS) surgery to relieve their symptoms.

The European Parkinson's Disease Association emphasizes that while Parkinson's disease is not life-*threatening*, it is life-*altering* (EPDA, 2019). It doesn't only affect the PD patients' quality of life, but also puts stress and economic burden on those taking care of them (Chen, 2010).

2.2 Speech problems in Parkinson's Disease

2.2.1 Hypokinetic dysarthria

Patients with Parkinson's disease don't only face motor and non-motor symptoms, but also problems with their speech. Speech in PD patients, known under the umbrella term *hypokinetic dysarthria*, is characterized by deficient respiratory control, abnormal voice quality, monoloudness, articulatory imprecision, poor control of voice onset and offset, abnormal and irregular speech rate, hypophonia (soft speech), and defective prosody, namely monotonous and reduced pitch (De Letter, Santens and Borsel, 2005).

Already in their seminal 1969 study, Darley, Aronson and Brown defined 35 speech parameters to help distinguish between different types of dysarthria, including categories relating to voice quality (9 dimensions, e.g. harshness, breathiness, voice stoppages, hypernasality), respiration (3 dimensions, e.g. audible inspiration), prosody (10 dimensions, e.g. rate, inappropriate silences, excess stress), articulation (5 dimensions, e.g. prolonged phonemes, distorted vowels), and an additional "bizarreness" category (2 dimensions, including intelligibility) (Darley, Aronson and Brown, 1969). After describing the speech of 32 PD patients, they defined PD speech in terms monopitch, monoloudness, reduced stress, imprecise consonants, inappropriate silences, short rushes of speech, breathiness and harshness. The study remained unquestioned until nearly four decades later, when Plowman-Prine et al. (2009) replicated the study and confirmed the original results.

Several speech production subsystems – categorized as *respiration*, *phonation*, *articulation*, *resonance* and *prosody* (Borden, Harris and Raphael, 1994) – are thus impaired in speakers with hypokinetic dysarthria (Swigert, 1997). *Articulation* difficulties are demonstrated in imprecise consonants and short rushes of speech, *rate* and *prosody* difficulties manifest as reduced stress, pitch monotony, inappropriate silences and variable rate, *phonatory* difficulties are evident from harsh voice quality, breathy voice and low pitch, and *respiratory* difficulties are shown in reduced loudness and monotony of loudness (Pinto et al., 2004). This is due both to the neuronal changes of Parkinson's patients as well as to the changes in their anatomy caused by the disease.

Approximately three quarters of PD patients face speech impairments after the onset of the disease (Hartelius and Svensson, 1994; Defazio et al., 2016; Ho et al., 1998; Logemann, Fisher, Bosher and Blonsky, 1978). They can appear both in early and late stage PD, but certain parameters of voice and speech performance have been found to deteriorate across the disease stage (Holmes, Oates, Phyland and Hughes, 2000; Skodda, Grönheit, Mancinelli and Schlegel, 2013). Voice quality features such as jitter and harshness of voice remain fairly stable across disease progression, but monoloudness and monopitch get worse (Holmes et al., 2000). Importantly, while the cardinal symptoms can still be improved by levodopa, even in the later stages of the disease, dysarthria steadily worsens, no matter the treatment (Pinto et al., 2004). Speech problems can eventually become severe enough to disturb communication and reduce the patients' quality of life (Dykstra, Hakel and Adams, 2007).

As the focus of the present thesis is on vowels articulation and voice quality, the following are some of the seminal studies regarding the differences between PD and healthy speakers.

2.2.2 Vowel articulation in PD

PD patients show a smaller (less dispersed) vowel space compared to healthy controls (Watson and Munson, 2008; Tjaden and Wilding, 2004). In a study by Skodda, Visser and Schlegel (2011), they tested 68 PD patients (34 male) and 32 age-matched controls who had to perform a reading task. The study extracted vowels from target words in the text and measured both the triangular vowel space area (tVSA) and the vowel articulation index (VAI), both of which are also used in the present study. They found that VAI values were smaller in PD speakers than the healthy controls, while tVSA values were smaller than healthy controls only in male PD speakers.

Rusz et al. (2013) tested 20 early-stage PD patients and 15 healthy controls, who sustained vowels /a/, /i/ and /u/ separately, performed a sentence repetition task, read a passage, and did a monologue. Both VAI and VSA scores were lower in PD patients compared to healthy speakers, with significant differences between tasks. More specifically, VAI scores were highest in sentence repetition and lowest in sustained phonation, while VSA scores were also highest in sentence repetition but lowest in reading passage.

2.2.3 Voice quality in PD

Compared to healthy controls, PD patients have lower intensity, a harsher voice and show more variable fundamental frequency and intensity. As PD severity worsens, the fundamental frequency (*f0*) of patients increases (Metter and Hanson, 1986), most likely because the laryngeal muscles get more rigid, leading to increased vocal folds and higher *f0* (Goberman, Coelho and Robb, 2002). Doyle et al (1995) showed that the mean *f0* was significantly higher for PD patients when measured during sustained vocalizations. PD patients show abnormal jitter and shimmer measures (Ramig et al., 2004) compared to healthy controls.

Goberman, Coelho and Robb (2005) tested 9 PD patients and 8 healthy control speakers. The participants performed sustained vowel vocalization, read a passage and produced a monologue. The researchers extracted the mean *f0* of prolonged vowels and from the reading/monologue task. They showed that PD patients produce higher *f0* values during the

reading/monologue task and an increased standard deviation of fundamental frequency compared to healthy controls.

Another way to measure voice quality is with *cepstral peak prominence (smoothed)* (CPPS), which relates to several voice quality parameters, including pitch, and to voice perturbation measures, such as jitter (Fraille and Godino-Llorente, 2014). Burk and Watts (2018) tested 32 PD speakers and 10 healthy controls who performed sustained vowels and connected speech. Results showed that patients with the non-tremor type of PD have a significantly lower CPP value compared to healthy controls. Further, Heman-Ackah et al. (2014) evaluated 835 patients with hypophonia (soft speech that can appear as a symptom of PD speech but also other diseases) and set the cut-off point for normal CPPS values at 4.0 or higher.

2.2.4 The effect of levodopa on PD speech

Although levodopa has been used since the 1960s to treat PD symptoms, it is currently still unclear how it affects speech and which *parts* of speech are affected (the discussion here especially concerns the difference between levodopa's effect on articulation versus other variables such as speech quality). While some studies have shown an improvement of speech symptoms after levodopa intake (Sanabria et al., 2001; Ho et al., 2008; Wolfe et al., 1975), others have shown either no change (Plowman-Prine et al., 2009; Goberman, Coelho and Robb, 2002; de Letter, Santens, Bodt, Boon and Borsel, 2006) or even a worsening of symptoms (Louis, Winfield, Fahn and Ford, 2001). Interpretation of study results is further complicated because different studies use different tasks (e.g. read vs. spontaneous speech), measure different parameters (e.g. voice quality vs. articulation), and test at differently defined periods. Especially early studies, from the 1970s and 1980s, often used perceptual measures, making them difficult to compare with later studies, utilizing acoustic and kinematic measures.

Additionally, studies differ in how they define the ON/OFF state of patients, measuring speech either: before and after morning medication; depending on the patients' symptom perception; in set intervals across a drug cycle; or in relative intervals across the drug cycle (Goberman and Coelho, 2002). Especially in advanced PD, the entire drug cycle needs to be taken into account in order to properly assess speech variation (De Letter et al., 2010).

See Table 1 on page 23 below for an overview of studies on the effect of levodopa on speech. Where possible, tasks and number of patients were included. Studies especially relevant to the

present research, namely those using acoustic methods to investigate the effect of levodopa on vowels pronunciation and voice quality[6], are further discussed below. Other studies, e.g. those on intelligibility (such as De Letter, Santens and van Borsel, 2005) or using perceptual measures (such as Plowman-Prine et al., 2009), are not discussed further.

Sanabria et al. (2001) tested 20 patients before and after levodopa, who vocalized a sustained /a/ vowel for at least two seconds. They performed acoustic analysis of voice quality (including tremor, noise, frequency and amplitude parameters) and found that the fundamental frequency significantly increased after medication (by at least 20 Hz) while other voice quality measures, such as jitter and tremor, significantly decreased[7] after medication. They speculated that the "improvement in fundamental frequency and other vocal parameters may be a result of decrease in laryngeal hypokinesia and rigidity" (Sanabria et al., 2001, p. 99).

Ho, Bradshaw and Iansek (2008) tested 9 PD patients once before their first medication (OFF state) and then once an hour for three consecutive hours (several ON states). The patients vocalized a sustained vowel, counted numbers and carried out conversation in three conditions (without volume instruction, as quietly as possible without whispering, and as loudly as possible without shouting). The study measured the intensity and duration, finding that levodopa consistently increased the loudness and rate of the patients' speech, while pitch and articulation remained unchanged.

A study by de Letter, Santens, Bodt, Boon and Borsel (2006) evaluated the effects of levodopa on speech rate of 25 PD patients who had to read a standardised passage (The North Wind and the Sun) while ON and OFF medication. While they did not find any significant improvement of speech rate, they did find an increase in *variability* during the ON state, perhaps caused by levodopa-induced dyskinesias (De Letter et al., 2006). On the other hand, Fabbri et al. (2017) tested 24 PD late-stage patients who had to produce a sustained vocalization of vowel /a/, repeat a declarative sentence, and read out 5 words and 5 sentences. There, levodopa did not significantly change any of the voice and speech variables. The authors point out that their results might be incompatible with those of De Letter et al. (2007) because the latter used PD patients in earlier stages and a better levodopa response.

[6] It is important to note that most studies, even when measuring vowels, predominantly focused on the fundamental frequency of the vowels and on other voice quality measures, such as jitter and shimmer, as opposed to focusing on vowel articulation measures, such as tVSA or VAI.
[7] Higher jitter and shimmer measures indicate pathological voice quality.

A study by Skodda, Grönheit and Schlegel (2011) tested 138 PD patients and 50 age-matched controls on a reading task consisting of four complex sentences. They analysed the variability of fundamental frequency and found that the fundamental frequency is generally reduced in PD patients compared to healthy controls, however this reduction is improved by levodopa treatment.[8]

Goberman, Coelho and Robb (2002) studied the phonatory characteristics of 9 PD patients who performed sustained vocalizations in two different conditions (as quietly as possible without whispering and as loudly as possible without shouting), performed a monologue, and read the first paragraph of a standardized passage. They tested them before morning medication, 1 hour after medication and 2 hours after medication, however each recording session took place on a different day to avoid fatigue. The results indicated that the only change occurred in the SD of *f0*, which increased in the OFF state. However, although as a group, PD patients showed no significant improvement in speech, individual patients did show improvements.

De Letter et al. (2010) tested 7 advanced-stage PD patients at 9 time segments, namely 15 minutes before their first medication (OFF state) and then at 15-minute intervals (ON states) for a total of 135 minutes. The patients produced a sustained vocalization of vowel /a/, DDK task (repeating the syllable "pa"), repeated vocalic transition /i-u/ and a reading task. De Letter et al. (2010) measured various parameters related to voice frequency, however only the standard deviation (SD) of diadochokinetic rate was shown to be significantly improved in the ON state.

A study on prosodic characteristics of Parkinsonian speech by Goberman, Coelho and Robb (2005) tested 9 PD before morning medication (OFF state) and one and two hours after morning medication (ON states). The patients performed paragraph reading and monologue tasks. While some differences were found between controls and patients (see above), none of the measures (including articulation rate, percent pause time, *f0* SD, and speech rate) were significantly affected by levodopa intake, although there were improvements seen in some individuals.

Finally, to our knowledge, only one previous study exists that explicitly evaluated the effect of levodopa on patients' vowel space area. Okada, Murata and Toda (2015) studied 21 PD patients who performed sustained vocalization of five Japanese vowels. The patients had to vocalize

[8] Both Skodda, Grönheit and Schlegel (2011) and Sanabria et al. (2001) indicate that PD patients generally have a lower *f0* than healthy controls. However, this does not negate the results of Goberman, Coelho and Robb (2002), which indicate that the *f0* increases as the disease progresses.

vowels 20 times before their levodopa intake and 1 hour after. The calculated pentagonal vowel space areas were discovered to be significantly larger after levodopa intake. However, as most of the previous studies, also those comparing speech of PD and healthy speakers, were conducted on Germanic languages, it is uncertain to what extent levodopa affects speech differently depending on the language (see Section 2.3.5 for a discussion on the importance of cross-linguistic research).

Study (year)	Sample	Task	Conditions	Results
Audelman et al. (1970)	N = 25	/*	OFF / ON	Improved intelligibility (*perceptual* ratings)
Azevedo et al. (2003)	N = 8	Reading task	OFF / ON	Fundamental frequency and speech intensity: improved in the ON state (higher fundamental frequency and higher intensity)
Azevedo et al. (2013)	N = 10	3 sentences	OFF / ON	Lower duration in the ON state
Critchley (1981)	/	/	/	Improved articulation, loudness and persistence of phonation* (provides no experimental evidence for his claims)
Daniels et al. (1996)	/	/	OFF / ON	No effect of medication
De Letter et al. (2006)	N = 25	Reading (The North Wind and the Sun)	OFF / ON	Rate: no effect of medication Variability: increase in the ON state
De Letter et al. (2005)	N = 10	Word intelligibility	OFF / ON	Improved intelligibility on single words in the ON state (perceptual ratings)
De Letter et al. (2007a)	N = 25	Word test	OFF / ON	Improved speech intelligibility and comprehensibility in the ON state
De Letter et al. (2007b)	N = 10	Reading task	OFF / ON	Improved speech intelligibility and comprehensibility in the ON state Improved pitch and loudness variability (both perceptual ratings)
De Letter et al. (2010)	N = 7	Sustained /a/ DDK ("pa") Repeated /i-u/ Reading task	OFF ON (15-minute intervals, 8x)	Improved SD of DDK period
Fabbri et al. (2017)	N = 24	Sustained /a/ Declarative sentence 5 words, 5 sentences	OFF / ON	No change in vowel duration, average *f0*, pitch break time, jitter, sentence *f0* or speech rate in the ON state
Frota et al. (2018)	N = 83	Speech	OFF / ON	Meta-analysis; modifications in *f0* and jitter in the ON state; vocal intensity not affected by levodopa
Gentil et al. (1999)	N = 1	Clinical evaluation	OFF / ON	No significant improvement of oral function in the ON state
Goberman et al. (2002)	N = 9	Sustained /a/ Monologue Reading task	OFF ON (1 h after) ON (2 h after)	Increased SD of fundamental frequency in OFF state (opposite to Sanabria et al., 2001). No effect of medication on *f0*
Goberman et al. (2005)	N = 9	Reading task Monologue	OFF ON (1h after) ON (2h after)	No effect of levodopa on *f0* or pause time

Ho et al. (2008)	N = 9	Sustained /a/ Counting Conversation	OFF ON – 3 times, once every hour	Loudness: increase in the ON state Rate: increase in the ON state Pitch and articulation: no effect of medication
Jian et al. (1999)	N = 15	Sustained /i/	OFF / ON	Shimmer: decrease in the ON state Tremor in intensity contours: decrease in the ON state Intensity: increase in the ON state No effect of levodopa on *f0*, jitter or mean flow rates.
Larson et al. (1994)	N = 2	/	OFF / ON	No consistent effect of medication on mean *f0*, intensity, jitter and shimmer
Louis et al. (2001)	N = 2	Spontaneous speech	OFF / ON	Worsening of speech
Maillet et al. (2012)	N = 12	Speech sequence	OFF / ON	No changes in brain activation in the ON state
Mawdsley and Gamsu (1971)	N = 20	/	/	Increased speech intelligibility in the ON state
Nakano et al. (1973)	N = 18	General speech	OFF / ON	Overall speech improvement and increased speech intelligibility in the ON state
Okada	N = 21	Sustained vowels	OFF / ON	Vowel space area expanded in the ON state
Pinto et al. (2005)	N = 4	General speech	OFF / ON	Case studies; variant effects of levodopa, depending on the individual
Plowman-Prine et al. (2009)	N = 16	Reading task	OFF / ON	No significant differences between ON / OFF state (perceptual ratings)
Poluha et al. (1998)	N = 10	Vowels	OFF / ON	No effect of medication on general speech performance
Quaglieri and Celesia (1977)	N = 30	Global speech score	/	Little difference in global speech score in ON versus OFF state
Sanabria et al. (2001)	N = 20	Sustained /a/	OFF / ON	Fundamental frequency: significant increase in the ON state Jitter: significant decrease in the ON state Voice tremor: significant decrease in the ON state Shimmer: no effect of medication
Skodda et al. (2011)	N = 138	Reading task	OFF / ON	SD of fundamental frequency: increase in the ON state
Skodda et al. (2010)	N = 22	DDK	OFF / ON	No changes in syllable repetition in the ON state
Solomon and Hixon (1993)	N = 14	Reading task	Medication cycle	Perceptual; more speech characteristics defective at the end of the medication cycle than at the beginning; few respiratory variables showed change
Wolfe et al. (1975)	N = 17	Speech passage	Before and after treatment	Improved articulation, loudness and persistence of phonation No effect of levodopa on speech rate

Table 1: An overview of studies investigating the effect of levodopa on speech

*In the table, / marks missing information. Especially in older studies, the exact tasks and conditions used are not explicitly discussed.

2.2.5 Other factors influencing PD speech

When investigating the effect of levodopa on speech and perusing previous studies, it is important to consider possible confounding factors that have led to differences in reported results. First, pathological speakers with hypokinetic dysarthria show large individual variation in their speech (Metter and Hanson, 1986). As Ho et al. (2008) point out, there is a high level of variability between patients as well as between measures, which could account for differences in findings. This also refers to the time of testing, as some patients perform better in the morning, while others do better in the evenings (Goberman and Coelho, 2002).

Second, it is important to ensure the lowest possible anxiety and fatigue levels, as anxiety can worsen PD symptoms, causes sudden OFF states and makes dyskinesias worse (Goberman and Coelho, 2002). Anxiety and depression are known to affect speech production (Schulz and Grant, 2000), and fatigue, which PD patients are prone to, exacerbates ON/OFF effects (Marsden et al., 1981). This has led Goberman, Coelho and Robb (2002) to suggest that the results of some studies of levodopa-related fluctuations on speech were influenced by the fact that patients provided multiple samples on a single day. However, the solution that the afore-mentioned authors suggest (i.e. recording each state on a different day) is not suitable either, as it does not take into account that speech is variable across days.

Third, age and gender are significant factors when studying speech production in PD, as elderly speakers' voices are perceived as hoarse and unsteady (Gorham-Rowan and Laures-Gore, 2006) and elderly (female) speakers generally have a lower fundamental frequency (Eichhorn, Kent, Austin and Vorperian, 2018). Time of day additionally needs to be taken into account due to the so-called "vocal warm-up effect", where the fundamental frequency is higher in the evening than in the morning (Garrett and Healey, 1987). Finally, the choice of task affects differently PD patients and healthy controls (Goberman, Coelho and Robb, 2002), so the type of task used to assess the PD patient's speech matters significantly (Ho et al., 1998). They might speak slower when they generate their own speech and faster when they are reading, for example (Schulz, Greer and Friedman, 2000).

The present study accounts for levodopa intake, general speech variability and fatigue levels. Furthermore, it follows the recommendation of Goberman and Coelho (2002) that the patients' home is the optimal environment for data collection, causing least anxiety. Importantly, by also

including healthy control speakers, it looks at general variability of speech measures across different sessions and days, ensuring that any found differences are indeed due to the effect of medication.

2.3 Slovenian language

2.3.1 General characteristics and dialectal variation

Slovenian (also: Slovene) is a South Slavic language with approximately 2 million native speakers, who live predominantly in Slovenia. It has an SVO (Subject-Verb-Object) word order with pro-drop, does not have any definite or indefinite articles, and is highly inflectional, with 6 cases. Each noun is marked for gender and number. A particularity of Slovenian, which it shares with few other languages in the world, is that it does not only have singular and plural grammatical number but also dual, which affects the endings of verbs, nouns and adjectives (see examples below).

Pametn**a** znanstvenic**a** gre na sprehod.
A smart female scientist has gone for a walk.

Pametn**i** znanstvenic**i** gre**sta** na sprehod.
Two smart female scientists have gone for a walk.

Pametn**e** znanstvenic**e** gre**do** na sprehod.
Several smart female scientists have gone for a walk.

What makes studying Slovenian language more difficult is its great dialectal diversity. There is even an old Slovenian saying: “Every village has its own voice”, which is beneficial for national diversity, but not that good for researchers trying to draw general conclusions about Slovenian speech. Dialects vary greatly depending on the region and speakers strongly identify with their regional dialect. The differences can be so big that two people from different sides of Slovenia find it difficult to understand each other unless they both use the Standard Slovenian dialect. Standard Slovenian is thus the variant that people speak and write in (semi-)official contexts, while dialectal Slovenian is spoken in everyday life. Speakers also greatly identify with their regional dialect. Figure 1 is a map of the 7 main dialect groups and altogether 50 dialects (Škofic et al., 2016). Due to the dialectal diversity, the present research has focused on a single region, namely the region surrounding Ljubljana (marked in dark violet) on the map below.

Figure 1: Map of the 7 main Slovenian dialect groups (the group of interest in dark violet)

2.3.2 Slovenian phonology: stress and phonemicity

Slovenian does not have a predictable stress place (Tivadar, 2004), as it arbitrarily positions stress on any syllable. Depending on the regional dialect, the place of stress changes (ibid.). It is clear, however, that most Slovenian words have only one stressed syllable (ibid.) and that stress *can* help distinguish between words (Toporišič, 2006): for example, p**à**rtija denotes a political party, while part**î**ja refers to a game of chess. Final stress tends to be eliminated (Greenberg, 2003).

Slovenian has a pitch accent, i.e. it is tonemic[9], at least in some dialects (Greenberg, 2003). However, tonemicity in Slovenian does not remain undisputed: nowadays, speakers seem to have a relatively low-level awareness of Slovenian tonemicity (Šuštaršič and Tivadar, 2005) and are undergoing a process of tone loss (Woznicki, 2006). Pitch distinctions do not carry a functional load anymore (Greenberg, 2003).

Slovenian is a syllable-timed language when described in terms of its internal rhythm (Komar, 2007). While the syllable-timed versus stress-timed distinction has become a popular way of characterizing languages, it is notable that few linguists mention it for Slovenian. On the other

[9] Tonemicity refers to a phenomenon where the meaning of the word depends on the intonation of the stressed syllable.

hand, research on other South Slavic languages, for example Croatian (Josipović, 1994) or Bulgarian (Dimitrova, 1997), has shown that it might be better to consider these languages on a scale between stress- and syllable-timed rather than in absolute terms. It is also likely that internal rhythm of Slovenian differs depending on the dialect, but while this has often been described in non-scientific ways (e.g. the dialect in the Western part of Slovenia, bordering on Italy, is described as "song-like"), it has not yet been empirically shown.

2.3.3 Slovenian phonology: sounds

Slovenian has 21 consonant phonemes and allows (or, rather, loves) consonant clusters, including words containing just consonants, such as for example *čmrlj* ("bumblebee") or *grm* ("bush"). Sonorants are always closer to vowels than non-sonorants and the formula for a syllable in Slovenian is NSVSN (Srebot-Rejec, 1992), where N stands for non-sonorant (e.g. /p, t, k, f, s/), S stands for sonorant (e.g. /v, m, n, l, r, j/) and V stands for vowel. Pronunciation of words with consonant clusters can differ depending on common colloquial pronunciation, as different dialects have different ways of realizing consonant clusters. Consonant clusters can appear word-initially as well as word-finally (Srebot Rejec, 1992).

While it was long assumed that Slovenian has 8 vowel sounds, Jurgec (2011) argues that it actually has two low vowels, both /ʌ/ and /a/. Further evidence for this additional vowel is that in an unstressed position, the pair {ʌ, a} undergoes the same neutralization process as the pairs {e, ɛ} and {o, ɔ}. From all Slovenian vowels, /e/ and /o/ can appear only in stressed positions, while others can be both stressed and unstressed (Srebot Rejec, 1998). Slovenian also has vowel sequences, which usually appear in the middle of the word and, in colloquial speech, see the insertion of the glide [j] (Jurgec, 2004).

Slovenian was previously described as having both short and long vowels (Toporišič, 2006), even though Srebot Rejec (1998) has empirically shown that the phonetic difference between short and long vowels does not exist anymore (i.e. there are no minimal pairs distinguishing the two types). Similarly, Tivadar (2004) found that there are no differences in vowel duration, although unstressed vowels were significantly shorter than stressed ones, with a ratio of at least 1:2 in most speakers. Slovenian thus tends to turn vowel quantity contrasts (distinction between short and long vowels) into vowel-quality contrasts (distinction between high and low vowels) (Greenberg, 2003).

Since the early 2000s, Slovenian vowels have begun to be objectively measured. Jurgec (2005) analysed vowels in tonal and non-tonal dialects and discovered that generally, the dispersion of the phoneme /u/ is far greater than expected. It is also generally the case that /e/ and /o/ are mid high vowels and higher than /ɛ/ and /ɔ/, and that the biggest differences can be found if they are in a stressed position (Jurgec, 2006), although the actual vowel placement in the vowel space plot does not change. Figure 2 illustrates the vowel space of stressed vowels in standard Slovenian.

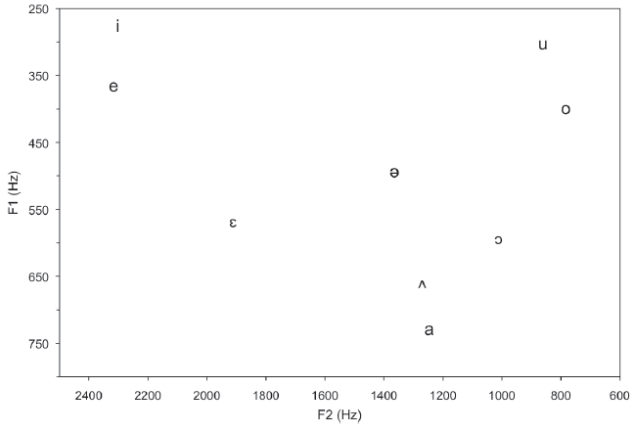


Figure 2: Vowel space of Slovenian speakers - vowels in stressed position (Jurgec, 2011)

While formant analysis shows both /e/ and /ɛ/ as well as /ɔ/ and /o/ to be four distinct phonemes, a Slovenian speaker will sense the two pairs as connected. One of the reasons for this is that the four sounds are sometimes in free variation, especially in colloquial speech (Srebot Rejec, 1988). One well-known example is the pronunciation of the word *pes* ("dog"), which can be realized as /*pes*/, /*pɛs*/ or /*pəs*/, depending on the region. Because this dialectal distribution of various vowel sounds is different from standard Slovenian, speakers sometimes find it difficult to translate their dialectal pronunciation into standard Slovenian (Jurgec, 2005).

Table 2 below shows the average formant values found in previous studies[10] by Srebot-Rejec (1988), Tivadar (2004) and Jurgec (2005).

[10] Although formant frequencies of vowels /ɔ/, /a/, and /ɛ/ have been shown to be influenced by whether the speaker's dialect is tonal or not (Jurgec, 2005), studies have not yet charted dialectal differences in vowel space of Slovenian speakers. For the purposes of present research, this does not matter much, as all 10 speakers came from the same general (tonemic) region. However, a previous short study by author of the thesis (Rebernik, 2018, unpublished; term paper for course *New Sounds* supervised by Dr. Dicky Gilbers and Dr. Wolfgang Kehrein) has shown that vowels /u/ and /o/ have a higher F2, i.e. they are more fronted, than previous studies indicated, which emphasizes the need for more empirical acoustic studies on Slovenian.

F1	/i/	/e/	/ɛ/	/a/	/ə/	/ɔ/	/o/	/u/
Srebot-Rejec (1988)	382	451	585	726	539	565	430	393
Šuštaršič et al. (1996)	301	387	328	735	456	577	414	317
Tivadar (2004) – women	385	429	606	774	/	655	453	423
Tivadar (2004) – men	351	397	526	603	/	551	397	385
Jurgec (2005)	280	386	586	717	498	587	423	321
F2	/i/	/e/	/ɛ/	/a/	/ə/	/ɔ/	/o/	/u/
Srebot-Rejec (1988)	2116	1973	1849	1332	1376	993	823	747
Šuštaršič et al. (1996)	2250	1916	1730	1362	1370	973	733	621
Tivadar (2004) – women	2318	2169	1943	1578	/	1125	879	770
Tivadar (2004) – men	2219	2169	1821	1324	/	1000	872	754
Jurgec (2005)	2309	2257	1860	1256	1368	993	814	852

Table 2: Average formant values for Slovenian vowels

The above-mentioned studies used predominantly speakers from the area of Ljubljana and its surroundings.

2.3.4 Slovenian PD speech and cross-linguistic studies

To our knowledge, only two studies so far discussed PD speech of Slovenian native speakers, leaving a gap to be filled. In her review article *Logopedic Treatment of Patients with Parkinson's Disease* Maja Ogrin, a clinical speech therapist, discusses the treatment of parkinsonian speech disorders in Slovenia. Amongst other things, she points out that Slovenia currently does not have any dysarthria tests. The Slovenian Institute for Rehabilitation instead uses a "Speech and Language Ability Overview" test, which includes a test of motor movements of speech organs, a test of articulation and speech intelligibility, a reading and memory test, a test of writing, and a description of voice disorders and swallowing difficulties (Ogrin, 2000).

The only other study was a bachelor's thesis by Širca (2012) who analysed the speech of 4 male Parkinson's disease patients with dementia. She does not use acoustic methods, but rather counts lexical, syntactic, phonemic and pronunciation mistakes (Širca, 2012). She characterizes the voice quality and speech of patients in line with previous studies on PD speech (e.g. hoarse voice) but fails to take into account that dementia could be at fault for some other mistakes (e.g. difficulties finding words).

2.3.5 Conclusion

Unfortunately, a lack of (English-written) studies on PD speech in languages other than English seems to be a rule rather than an exception. In a list of 269 articles on PD speech taken from 19 journals[11], 174 articles (65%) describe English patients, followed far behind by French (17 articles, 6.3% of total articles) and Dutch (16 articles; 5.9% of total results). Significantly, nearly half of the articles do not explicitly mention the language of the patients (most often, judging by the article authors' affiliations, the language is not mentioned if the patients in question are native English speakers).

Speech research can contribute towards the diagnosis of PD as well as towards the development of new speech therapies that help patients maintain their linguistic and communicative abilities. Considering the differences in languages across the world, it seems crucial to adopt a cross-linguistic perspective to studying PD speech. After all,

> "although speech motor control is a universally shared human ability, the evolution and impact of speech disorders may depend on the linguistic and cultural environment of the patients."
>
> (Pinto et al., 2017, p. 157)

Languages have vastly different characteristics, making it crucial to know which ones are universal and which are language-specific. However, while there have been strides and encouragements made towards cross-linguistic research on speech disorders (see e.g. Miller and Lowit's 2011 book on a cross-language perspective on motor speech disorders), it remains clear that the prevalent language of patients studied is English, and English seemingly remains the driving force behind the development of new research approaches. Studies such as ours, looking at Slovenian patients and, in the future, allowing for direct comparison with data of Dutch patients, are crucial if we are to make improvements in PD speech research.

[11] The following information comes from a review paper on languages of PD speech research, written by thesis author for purposes of LOT school in summer 2018 (Rebernik, 2018, unpublished; paper entitled *Speech Studies in Parkinson's Disease Research: Linguistic Bias and Diversity*).

3 Research questions and hypotheses

Considering the theoretical underpinnings (see Chapter 2), the present research aims to answer several research questions relating to the dysarthric speech of Slovenian Parkinson's disease patients. The questions will be answered using acoustic methods (see Chapter 4), namely by investigating changes in several acoustic parameters, including vowels' pronunciation and voice quality (see Table 3 for an overview of the measured parameters). In all measures, we expect a large degree of individual variation (between different PD speakers as well as within different session of the same PD speaker), as that is characteristic of pathological speech (e.g. Metter and Hanson, 1986).

The research questions we aim to answer are the following:

- How does Parkinson's patients' speech differ from healthy speech?
- What is the effect of levodopa on parkinsonian speech?
- Does fatigue affect the measured acoustic parameters?
- Does choice of task affect vowel articulation measures in PD patients and healthy control speakers?

Acoustic parameter	**Measure**
Vowels' pronunciation	Vowel Articulation Index (VAI) Triangular Vowel Space Area (tVSA)
Voice quality	Cepstral Peak Prominence Smoothed (CPPS) Fundamental frequency (*f0*)

Table 3: Acoustic parameters and their measures (for more details see Chapter 5)

The first main research question concerns **the differences between Parkinson's disease patients' speech compared to healthy speech.** More specifically we aim to determine whether PD patients' speech differs from healthy speech. Our null hypothesis is that participants in the PD group will not show significantly deteriorated acoustic parameters compared to participants in the healthy control group. Our alternative hypothesis is that the PD group will show significantly deteriorated acoustic parameters. The latter is in line with results from previous studies, which have shown that compared to healthy controls, PD patients have:

- a smaller triangular vowel space area (tVSA) than healthy controls (following the results of Skodda, Visser and Schlegel, 2010);
- a lower vowel articulation index (VAI) than healthy controls (following the results of Skodda, Visser and Schlegel, 2010);

- smaller cepstral peak prominence smoothed (CPPS) compared to healthy controls (following the results of Burk and Watts, 2018);
- lower fundamental frequency (*f0*) compared to healthy controls (following the results of Goberman, Coelho and Robb, 2002).

The second main research question concerns **the effect of levodopa on dysarthric speech of Slovenian patients with Parkinson's disease.** More specifically, we aim to determine to what extent Slovenian parkinsonian speech changes as a result of medication, i.e. the levodopa intake. Due to varied reports on the effects of levodopa on parkinsonian speech, we have set the null hypothesis that levodopa will not significantly affect the defined acoustic parameters, and the alternative hypothesis that levodopa will significantly affect the acoustic parameters.

We will investigate the effect of overnight withdrawal from levodopa (i.e. the so-called OFF state) on the acoustic parameters described in Table 1. It is hypothesized that when comparing the ON and OFF states:

- tVSA measures will not significantly differ in the ON and OFF states;
- VAI measures will not significantly differ in the ON and OFF states;
- the CPPS will not significantly differ in the ON and OFF states;
- the *f0* will not significantly change in the ON and OFF states.

Two additional subquestions have been set. The first subquestion concerns the effect of fatigue on speech. More specifically, we aim to determine **whether fatigue has an effect on healthy and dysarthric speech**, and whether it plays a role when studying the effect of levodopa.

The second subquestion concerns the difference between read versus (semi-)spontaneous speech, namely we will investigate **whether task plays a role in vowel articulation of PD patients and healthy control speakers.** As different studies use different tasks, it is important to determine whether there is a task-specific effect on speech that could also account for differences in results of previous studies. More specifically, it is hypothesized that:

- the two measures of vowels' pronunciation, tVSA and VAI, will be bigger in semi-spontaneous (elicited) speech compared to read speech (following the results of Rusz, Cmejla and Tykalova, 2013).

By answering the research questions and confirming/rejecting our hypotheses (outlined above), we will contribute towards the study of speech in Parkinson's disease patients in several ways.

First, by collecting data not only in five different sessions and different stages of the medication cycle but also across different days, we will be able to determine whether there is indeed a (consistent) effect of levodopa on speech. Unfortunately, since many studies only measure ON/OFF moments of one medication cycle, the results they obtain might not be representative of the general effect of medication, as there are too many influencing factors. Second, by also collecting fatigue data and taking into account the time of day when the speech was recorded, we eliminate another potential variable that could influence speech. Third, by having the participants record their speech at home, with their partners, we know that we are recording their speech without needlessly causing them anxiety, thus improving ecological validity. Finally, by using Slovenian patients, we will be able to study a previously untested population. In the future, such studies could help us determine whether certain results are valid cross-linguistically, highlighting the importance of doing cross-linguistic research in speech motor control disorders.

4 Method

4.1 Participants: recruitment and sample

To recruit Slovenian participants with Parkinson's disease, we established contact with Trepetlika, the Parkinson's Disease Society of Slovenia. It is the main (and only) support association for PD patients in Slovenia, which has been active for more than 27 years. Trepetlika is a humanitarian and volunteer organization that connects nearly 1000 members and has branches in several cities across Slovenia (Trepetlika, 2019). It connects people with Parkinson's disease and other parkinsonisms as well as their family and friends, with the goal of spreading awareness of the disease, organizing sports and cultural activities, and collaborating with health experts.

After contacting them and meeting with the association's president, Ms. Cvetka Pavlina Likar, and main coordinator, Ms. Mirjam Martini Gnezda in May 2018, they agreed to collaborate with us and help with recruitment of patients. We agreed that they would put our recruitment letter (see Appendix 1) on their website, www.trepetlika.si, and inform patients about our study during their organized activities. Afterwards, we would receive a list of names and phone numbers of potential suitable and willing participants. From this list, we could choose which participants to contact. Without the collaboration of Trepetlika, whom the patients already trust, it would have been difficult to recruit and test a suitable number of patients.

In the recruitment letter given to Trepetlika, there were several inclusion criteria for participants with Parkinson's disease. Namely, eligible participants were those who:

- didn't have any brain injury and had not suffered from any strokes in the past;
- didn't suffer from any other speech disorders (e.g. stuttering);
- weren't diagnosed with depression (or, alternatively, had a milder form of depression for which they did not need to take anti-depressants);
- hadn't undergone deep brain stimulation (DBS) surgery;
- took levodopa in the form of pills;
- had Slovenian as their first, native, language.

The inclusion criteria that we used are standard for speech studies with PD patients. Brain injury and stroke are exclusionary, also because symptoms arising from physical injury cannot be reliably distinguished from symptoms arising from the disease and the neural damage it causes.

Speech disorders are listed as an exclusion criterion for a similar reason, namely if an individual suffers from speech problems in general, we cannot know to what extent these problems are exacerbated or, alternatively, caused by PD.

Deep brain stimulation (DBS) and depression were, likewise, excluded due to their effect on speech. DBS affects speech by improving some motor components but decreasing intelligibility (Pinto et al., 2004; Santens et al., 2003). Depression can affect speech production, and is frequently listed as an exclusion criterion, also in non-PD studies (Schulz and Grant, 2000). Finally, the participants had to take levodopa as opposed to the Duodopa pump. The latter administers a continuous stream of a levodopa-carbidopa intestinal gel, meaning that patients do not experience ON/OFF states, which would be contrary to our study's aim.

Additionally, the goal was to recruit participants from the same region of Slovenia, as dialectal differences between different regions are substantial. This was not included in the recruitment letter, as we could not know in advance how many individuals would be interested, and we wished to ensure that we would get a sufficient number of participants. The recruitment letter stated that we were looking for 5-10 PD participants in total. The number of participants was set relatively low due to constraints in time (i.e. a total of 5 weeks was available for testing) and limited equipment (a total of 3 sets were available). However, it is necessary to note that such low numbers are not unusual in studies on speech of PD patients, as recruitment is oftentimes difficult, especially considering the strict inclusion criteria that need to be observed (see above).[12]

Trepetlika provided a list of 10 potential participants, whom we contacted for collaboration via a telephone call. From all the potential participants, 5 were considered unsuitable: 2 were considered unsuitable due to a strong regional accent; 1 was unsuitable because he had suffered from a parkinsonism, not Parkinson's disease, and did not feel an effect of levodopa; and 2 (who were supposed to record as a pair and had already received instructions and equipment) dropped out of the study due to time demands. The study originally intended to recruit 5 (or more) participant pairs, meaning 5 PD patients and 5 healthy controls. However, due to a rare opportunity, one of the participant pairs included *2* PD patients – a pair of monozygotic twins, both diagnosed with PD but in different stages of the disease.

[12] The number of PD patients was comparable in the study on Dutch native language speakers that used the same design and on which the present study is based.

The information on participants (both PD patients and their recording partners) can be found in Table 4. All scores and duration information refer to the time of testing. All participants were born in the Central Slovenian region (marked in violet on the map of Slovenia in Chapter 2.3) and have also lived there their entire lives. This ensured a fairly homogeneous accent across the participants and allows comparability with other acoustic studies on Slovenian speech, which mostly recruit participants from the capital city.

Subject code	Group	Sex	Age
01	PD	M	57
02	PD	M	57
03	PD	M	58
04	PD	M	62
05	PD	M	61
06	PD	F	71
07	HC	M	71
08	HC	F	49
09	HC	F	49
10	HC	F	62

Table 4: Participant information

Information on PD participants' medication and disease specifics can be found in Table 5. All scores and duration information refer to the time of testing.

Subject code	Disease duration	MMSE	L-dopa duration	L-dopa daily dose	Other medication (daily dose)	Speech problems	ON/OFF state
01	6	29	10 days	75 mg	18.75 mg carbidopa 1 mg rasagiline	No.	No.
02	7	28	3 years	450 mg	112.5 mg carbidopa 5.4 mg biperiden 120 mg propranolol	Tight muscles, talks faster in ON state.	Yes.
03	9	25	3-3.5 years	600 mg	150 mg carbidopa 1000 mg entacapone 1 mg rasagiline	Quiet speech, unintelligible.	Yes.
04	5	21	2 years	700 mg	100 mg benserazide 75 mg carbidopa 800 mg entacapone 3.15 mg pramipexole 1 mg rasagiline 500 mg aspirin	No.	Yes.
05	12	20	8 years	850 mg	212.5 mg carbidopa 1130 mg entacapone	Runs out of words.	Yes.
06 (f)	2	24	2 years	400 mg	100 mg carbidopa 6 mg ropinirole	Blocked speech, runs out of words.	No.

Table 5: PD participants' disease information

With PD patients, the average disease duration was 6.8 years (range: 2–12). The amount of time for which the PD patients have been taking levodopa ranged between 10 days and 8 years, with the majority of patients taking it between 2 and 3 years. The levodopa equivalent daily dose (LEDD) ranged between 75 mg and 850 mg (mean = 512.5 mg), combined with carbidopa, which helps levodopa pass into the bloodstream (range between 18.75 mg and 212.5 mg, mean = 111.5 mg). The patients' levodopa intake was supplemented with other medication, most commonly rasagiline (1 mg), which is designed to treat non-motor symptoms such as fatigue. Four out of six patients (all excepting patient 01 and 04) faced some sort of speech problems[13] due to the disease. Four out of six patients (all excepting patient 01 and 06) reported that they experience fluctuations to the medication (i.e. ON/OFF states).

The study received a Letter of No Objection from the Research Ethics Committee (CETO) of the Faculty of Arts, University of Groningen, establishing that the research protocol follows internationally recognized standards to protect the research participants (see Appendix 2).

4.2 Equipment and experimental set-up

Each PD patient was asked to recruit their (life) partner – this ensured not only that they had someone to record with, but also that we simultaneously obtained speech from patients and healthy controls (in all pairs but one, see above). The first meeting with each pair of participants was scheduled over the phone. The testing took place at the participants' homes, as that was also where all the acoustic recordings would be made. During the first visit, the participants first received an explanation of the study[14], read the information letter (see Appendix 3) and signed the consent form (Appendix 4). They were told in advance that the visit would take around 2 hours, and they were free to ask as many questions as they wished. Furthermore, they were also asked to ensure that their (recording) partner was present.

Second, after reading the information letter, the participants received instructions for the tasks they would need to perform. The tasks are described below (Section 4.4); the instruction sheet (including information on using the equipment) can be found in Appendix 5. The participants received four envelopes, one for each day of testing. The envelopes contained the

[13] Due to time constraints, we were unable to recruit patients who had been officially diagnosed with hypokinetic dysarthria. The term "dysarthric" is thus used as an adjective denoting characteristics of hypokinetic dysarthria (quiet speech, mumbling, reduced articulation).

[14] They were informed that their speech was measured because we wished to study the effect of levodopa and fatigue. They were also informed of the purpose behind using several different tasks. However, they were not informed of the purpose behind target words, as we did not wish for their pronunciation to change or for them to focus on the target words.

spot-the-differences task and fatigue survey for every session of that day. The participants were also asked to choose a colour before recording for the first time (each partner chose either green or yellow), as it was important to keep the fatigue surveys and spot-the-differences sheets consistent. Specifically, we needed to ensure that we knew which fatigue survey was filled out by which participant, and which sheet was being described by which participant in any given session. Colours were chosen instead of names in order to preserve anonymity.

Third, the participants received instructions for using recording equipment. Each pair of participants received two Shure WH20 XLR headset microphones, an iRig Pro Duo audio interface, and a Motorola C Plus smartphone and charger. The two headset microphones were connected to the iRig, which digitized and transferred the speech recordings to the phone. Recordings were automatically uploaded to the Google Drive cloud server, which enabled us to monitor the study remotely and provide assistance if needed. Permission to upload the recordings using third-party apps was obtained from the participants beforehand in order to comply with the new European GDPR regulations. The equipment (including internet access of the phone) was then tested to ensure that the participants understood the tasks and how to use the equipment. We had altogether 3 phones and 6 headset microphones, meaning 3 participant pairs could record simultaneously.

Finally, during the home visit, we administered the Mini Mental State Exam (MMSE) to the PD patients to test their cognitive abilities. The MMSE (original by Folstein, Folstein and McHugh, 1975; Slovenian version by Granda, Mlakar and Vodušek, 2003) was originally designed as a way to estimate the level of dementia in Alzheimer patients, but has since also been frequently used to discover potential cognitive deficits in non-demented patients. Questions refer to time and date (e.g. “What date are we today” and “Where do you live”), memory recall (e.g. “Repeat these three words after me”), subtraction or spelling backwards, repetition, object naming, task instructions, and drawing. The Slovenian version of the MMSE can be found in Appendix 6.

Additionally, we also prepared a short questionnaire for the participants (see Appendix 7). It included demographic information (e.g. gender, educational level, place of birth and place of living), language information (which other languages the participants speak, if any), questions concerning PD diagnosis (when the disease was diagnosed, whether the disease affects their speech and how), and questions concerning medication (which medication they are taking, how

often they take levodopa, whether they experience ON/OFF states, and for how many years they have been taking levodopa).

Session	Time
1 – OFF	15 minutes before first levodopa intake
2 – ON morning 1	60 minutes after levodopa
3 – ON morning 2	120 minutes after levodopa
4 – ON afternoon 1	60 minutes after afternoon/evening levodopa
5 – ON afternoon 2	120 minutes after afternoon/evening levodopa

Table 6: Recording times in terms of levodopa intake

Based on the times of day that the participants took levodopa, we calculated when the recording sessions should take place (see Table 6) and each participant received a list of exact times based on his or her own levodopa schedule. The participants held on to the equipment for two weeks. During these two weeks, they had to record themselves on four separate days, five times on each of the days. We did not specify which days the participants should record themselves on (they could choose four consecutive days, if they so wished).[15]

Participants were further instructed to call us if they had any additional questions, and we set the approximate date for when we would pick up the equipment. During the two weeks, we were in contact with all participant pairs at least once, to give them additional instructions and clarifications. In the second (and last) visit, when the equipment was picked up, the participants received two University of Groningen mugs as a thank-you gift for their participation.

4.3 Stimuli design

To be able to measure vowel and consonant articulation (which are said to be impaired in PD patients, see discussion in Section 2.2 above), we needed to choose target words in regular phonetic contexts. The target vowels (V_t) were corner vowels: front close vowel /i/, close back vowel /u/ and central open vowel /a/. The reason is twofold. First, if the vowel space area of PD patients is indeed affected, then the reduced vowel space area would be most apparent from the corner vowels. Second, the pronunciation of other vowels in Slovenian is strongly marked and shows a lot of variation (see Section 2.3 on Slovenian). Choosing corner vowels eliminated potential individual differences in the pronunciation of target words. The target consonants (C_t)

[15] The exact order of days did not matter as the only argument against recording on consecutive days is fatigue (e.g. due to the strain of recording 5 times a day, 4 days in a row). However, since a fatigue survey was included before every session, this potential effect was accounted for.

were plosives (i.e. stop consonants): bilabial plosives[16] /p/ and /b/, dental plosives /t/ and /d/, and dorsal plosives /k/ and /g/.

Due to phonotactic constraints in Slovenian and its strong use of cases, it was crucial to ensure that the stressed syllable would remain the same in all cases, i.e. that the combination $V_t + C_t$ would be pronounced not only in the first task (reading carrier phrases) but also in subsequent tasks (semi-spontaneous elicited speech). The words needed to be well-known and, furthermore, for the purposes of picture tasks, it was necessary for at least some of the target words to be easily depictable. Table 7 below is an overview of the target words.

	Word (meaning)	**Preceded by**	**$V_t + C_t$**	**Followed by**	**Syllables**
Bilabial voiceless	Kapa (hat)	voiceless plosive /k/	ap	vowel /a/	2
	Papež (the Pope)	voiceless plosive /p/	ap	vowel /ɛ/ + fricative /ʒ/	2
	Pipa (pipe)	voiceless plosive /p/	ip	vowel /a/	2
	Ekipa (team)	/ɛ/ + voiceless plosive /k/	ip	vowel /a/	3
	Lupa (glass)	approximant /l/	up	vowel /a/	2
	Pupa (doll)	voiceless plosive /p/	up	vowel /a/	2
Bilabial voiced	Žaba (frog)	voiced fricative /ʒ/	ab	vowel /a/	2
	Kabel (cable)	voiceless plosive /k/	ab	vowel /ɛ/ + approximant /l/	2
	Riba (fish)	flap /ɾ/	ib	vowel /a/	2
	Šiba (rod)	voiceless fricative /ʃ/	ib	vowel /a/	2
	Rubelj (rouble)	flap /ɾ/	ub	vowel /ɛ/ + approximant /l/	2
	Tuba (tuba)	voiceless plosive /t/	ub	vowel /a/	2
Dental voiceless	Solata (salad)	dental approximant /l/	at	vowel /a/	3
	Vrata (door)	consonant cluster /vɾ/	at	vowel /a/	2
	Kita (braid)	voiceless plosive /k/	it	vowel /a/	2
	Pita (pie)	voiceless plosive /p/	it	vowel /a/	2
	Ruta (bandana)	flap /ɾ/	ut	vowel /a/	2
	Valuta (currency)	dental approximant /l/	ut	vowel /a/	3
Dental voiced	Brada (beard)	consonant cluster /bɾ/	ad	vowel /a/	2
	Čelada (helmet)	dental approximant /l/	ad	vowel /a/	3
	Robida (bramble)	voiceless plosive /b/	id	vowel /a/	3
	Piramida (pyramid)	nasal /m/	id	vowel /a/	4
	Pudelj (poodle)	voiceless plosive /p/	ud	vowel /ɛ/ + approximant /l/	2
	Buda (Budha)	voiceless plosive /b/	ud	vowel /a/	2
Velar	Omaka (sauce)	vowel + nasal /m/	ak	vowel /a/	3
	Mlaka (puddle)	consonant cluster /ml/	ak	vowel /a/	2

[16] The first listed sound is voiceless, the second is voiced.

	Pika (dot)	voiceless plosive /p/	ik	vowel /a/	2
	Slika (painting)	consonant cluster	ik	vowel /a/	2
	Bukev (beech tree)	voiceless plosive /b/	uk	vowel /e/ + approximant /ʊ/	2
	Kljuka (doorhandle)	consonant cluster /klj/	uk	vowel /a/	2
Velar voiced	Žaga (saw)	voiced fricative /ʒ/	ag	vowel /a/	2
	Glagol (verb)	consonant cluster /gl/	ag	vowel /ɔ/ + approximant /l/	2
	Figa (fig)	voiceless fricative /f/	ig	vowel /a/	2
	Knjiga (book)	consonant cluster /knj/	ig	vowel /a/	2
	Vijuga (winding)	dorsal approximant /j/	ug	vowel /a/	3
	Uganka (riddle)	/	ug	vowel /a/ + nasal /n/	3

Table 7: Target words with target sounds and environments

The goal was to have two target words for each combination V_t + C_t. With 3 target vowels and 6 target plosives, there were 18 combinations, resulting in 36 target words in total. We selected the target words by first perusing a list of the most common 2000 Slovenian words, compiled with help of the biggest Slovenian corpus (Jezikovna svetovalnica, 2016) and the online Dictionary of Slovenian Literary Language (IJS ZRC SAZU, 2019).

As seen in Table 7 above, all 36 target words follow the construction **C V_t C_t V**[17] (e.g. /ka.pa/ or /ri.ba/). Most consonants preceding the target vowel are voiceless and most target consonants are followed by corner vowel /a/. Furthermore, most target words have two syllables and stress in all words (no matter the number of syllables) falls on the syllable with the target vowel.

4.4 Experimental tasks

The experimental paradigm included several tasks, which allowed the measurement of elicited read speech (i.e. carrier phrases with embedded target words), semi-spontaneous speech (i.e. the card game "kwartet" and a spot-the-differences game) and general oral motor control (i.e. the diadochokinesis task). The participants played the games in pairs, always following the same order, which is also described below. We decided against randomizing the order of the tasks, as it would have been difficult to ensure consistency remotely (i.e. participants would have been more likely to skip tasks or misremember the instructions). Both participants in a pair, not just the PD patient, carried out all the tasks, ensuring that data of both healthy and PD speech was obtained. Each session took approximately 15 minutes.

[17] Although plosives were not analysed for the purposes of this thesis, they were measured and can still be analysed in the future.

4.4.1 Fatigue survey

First, the participants marked their tiredness levels, following the instructions "Before recording, please mark how tired you feel" (see Figure 3). The fatigue survey (in Slovenian) can be found in Appendix 7.

Figure 3: Fatigue scale; ratings from left to right are: not at all, little, medium, very, exhausted (source of original picture in English: 4.bp.blogspot.com)

4.4.2 Carrier phrases

After marking their fatigue levels, the participants read out a list of carrier phrases, which contained embedded target words, described in Section 4.3. The carrier phrase was the following: "Beseda **pipa** ima več kot en zlog" ("The word **target word** has more than one syllable). It was chosen to ensure that the target word was in the nominative case and that the final sound of the preceding word (the /*a*/ in /*be.'se.da*/) would not change the pronunciation of the initial sound of the target word. Likewise, it ensured that the last sound of the target word (often, but not always, the vowel /a/) would not merge with the first sound of the following word (/*i*/ in /*i.'ma*/).

Beseda	**pipa**	ima	več	kot	en	zlog.
bɛ.'se.da	*'pi.pa*	*i.'ma*	*vɛtʃ*	*kɔt*	*ɛn*	*zlɔg*
The word	***pipe***	*has*	*more*	*than*	*one*	*syllable.*

There were altogether 36 carrier phrases, each printed on a card that was laminated afterwards (see Figure 4). For each session, each participant read half of all carrier phrases, namely one target word for every combination *corner vowel* + *target plosive*. For the full list of carrier phrases, please consult Appendix 9.

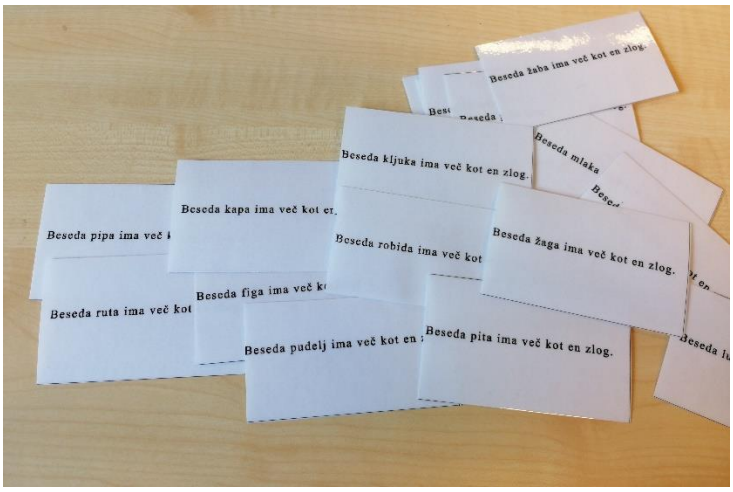


Figure 4: Cards with carrier phrases

Participants received two decks of cards with carrier phrases and were instructed to take a different deck for every session. They had to shuffle the deck and place it in front of them, face down. They took each individual card, read it, and placed it back. This ensured that they took enough time to read each sentence and did not rush through the list (as would happen if they were simply reading the sentences from a single piece of paper).

4.4.3 Spot-the-differences game

After reading out the carrier phrases, the participants proceeded with a spot-the-differences game. They had two sheets of laminated paper with pictures that differed in location, colour and details (see Figure 5). They were instructed to find 10 differences between their picture and the picture of their partner, without looking at their partner's game sheets. They played the game for 5 minutes or less, if they succeeded in finding all 10 differences earlier than in 5 minutes. In the envelope, they also received a list of solutions.

Figure 5: An example of the spot-the-differences game

The items on each sheet represented target words that could be depicted. There were 20 different combinations of the spot-the-differences games (10 sheets with differences), to prevent participants from learning the differences by heart (see Appendix 10 for several example game sheets). Due to time constraints, the game was not analysed for the purposes of the thesis, and the data is thus not included.

4.4.4 Kwartet card game

Following the spot-the-differences game, the participants played the card game kwartet, the goal of which is to collect four cards of the same category. The category was the target word that needed to be elicited. For example, for the target word /*pi.ta*/ ("pie"), there were four cards, featuring the target word at the top and different subtypes of the target word (see Figure 6). As with carrier phrase cards, there were two decks of kwartet cards.

The participants were instructed to take a different deck every session, and to shuffle the cards. There were altogether 18 categories, one for each combination corner vowel + target plosive. In each of the two decks, there were 9 categories, meaning 36 cards altogether. Each deck had 3 instances of every corner vowel and 1-2 instances of the target plosive.

Figure 6: Example of kwartet deck

Participants received the instructions to start with 6 cards and ask their partner about his or her cards in the following way: "From the category [target word], can I get [subtype]?" (e.g. "From the category pie, can I get a quiche?". They played the game for 5 minutes, after which they stopped, no matter the score. While the game of kwartet is not known in Slovenia, this turned out to be an advantage, as participants made sure to use the prescribed phrasing when asking their partner about the cards. Cards from all 18 categories can be found in Appendix 11.

4.4.5 Diadochokinesis task

The final task was a diadochokinesis (DDK) task, measuring the participants' ability to rapidly carry out alternating speech motor movements. They were instructed to repeat the syllables /pa/, /ta/ and /ka/ as well as the nonsense multi-syllabic word /pa.ta.ka/ as many times and as quickly as possible in a single breath. The DDK task was analysed but is not included in the thesis due to time and space constraints.

5 Measures and analysis

During our analysis, aimed at determining the effect of levodopa on parkinsonian speech, we performed several measurements using the data collected during the various tasks (described in Section 4.4 above). The recordings were first manually pre-processed, i.e. prepared in Adobe Audition and segmented in PRAAT (Boersma & Weenink, 2018), and then automatically analysed with help of various software programs, most notably MATLAB (version 9.4.0, r2018a). Finally, statistical analysis was performed in RStudio (R Core Team, 2013).

5.1 Data pre-processing

After the participant pair had informed us that they had finished recording, we downloaded the recordings from Google Drive and made a local backup copy. Upon collecting the equipment, the recordings were uploaded to several locations.[18] The saved raw files contained no identifiable information. The recordings were subsequently removed from the phone and from the Google Drive folder (which had been created specifically for that participant pair). There were approximately 50 hours of raw speech files.[19]

The recordings were then processed in Adobe Audition, an audio editing software. They were first transformed from stereo (two channels) to mono (one channel). During this process, no changes in intensity (i.e. loudness) were made, although some participants spoke noticeably more quietly than others. This is possibly due to positioning of the microphone, even though the participants had been instructed to always put it in the same position, namely approximately 5 centimetres from the mouth. As the speech on the recordings was clearly captured and as we did not plan on measuring speech intensity (all our measures, described below, are intensity-independent), this is not problematic.

Each session recording was split into 4 tasks and named according to a unified convention, where "Pp" denotes the number of the participant pair, "Day" and "Session" denote the day and session, "Task" denotes the task ("List", "Kwartet", "Diapix" and "Diado", respectively), and "Participant" denotes whether the participant was a healthy control (mark HC) or Parkinson's disease patient (mark PD). For example, the wordlist task read by the healthy control from

[18] Recording backups can be found on the researcher's local drive, two external encrypted hard drives and the University of Groningen's Y: drive for employees, to which the researcher and other members of Speech Lab Groningen have access. Processed recordings were additionally added to supplement the raw recordings.
[19] 10 participants x 20 sessions x 15 minutes = 3000 minutes (50 hours); note that the actual recordings were longer, as the sessions sometimes took more than 15 minutes (especially at the beginning).

participant pair 1 on the first session of the first day would be named "Pp1Day1Session1ListHC". This enabled us to maintain the participants' anonymity while still clearly differentiating between the different pairs, individuals and recording sessions.

The prepared recordings were further processed in PRAAT (Boersma & Weenink, 2019), a program for phonetic analysis of speech. The target vowels /a, i, u/ and plosives /p, b, t, d, k, g/ in the target words in the first three tasks were manually annotated and segmented. The target words (see Chapter 4.3 for target words) were marked in two interval tiers, namely the ORT tier (with the orthographic transcription of the word) and the MAU tier (with the target sounds). Exact phonemic transcription was avoided as it was not necessary for further processing.[20] The target vowels and plosives that did not appear in the desired environment (e.g. the first /p/ sound in the target word /*'pu.pa*/) were marked with an *x*, to prevent further processing in MATLAB. See Figures 7 and 8 for an annotated and segmented spectrogram. By preparing the recordings in such a way, the .wav and TextGrid files could be further (automatically) analysed in MATLAB.

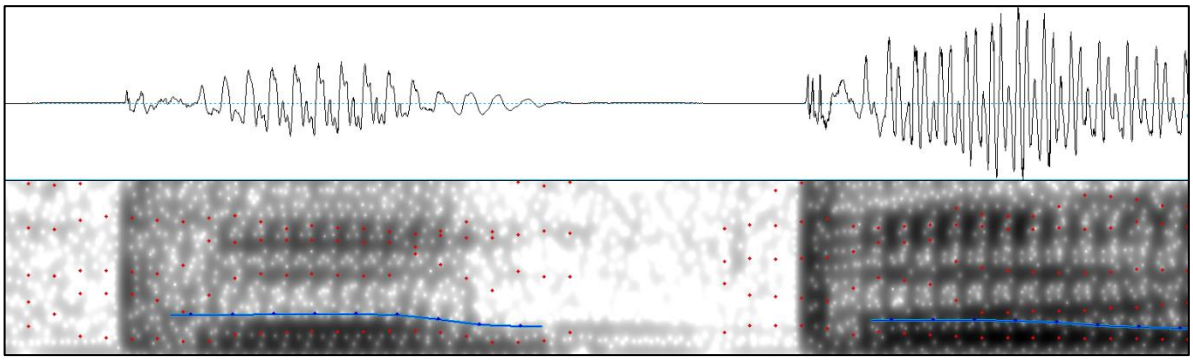


Figure 7: Waveform (above), spectrogram (below) with pitch (blue line) and formants (red dots) of the target word pupa.

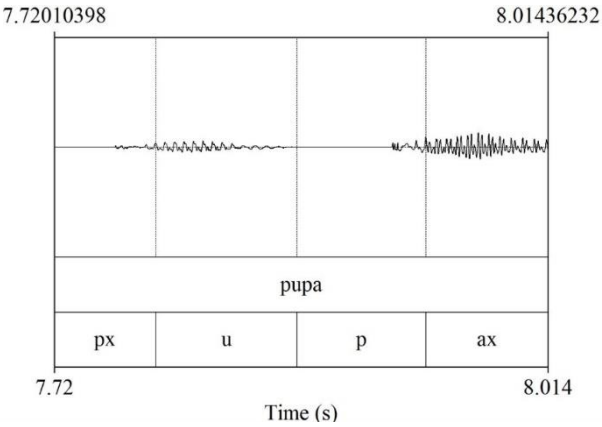


Figure 8: Waveform (above) with orthographic tier and sound segmentation in the target word pupa.

[20] That is also why we refer to a "sound segmentation" tier as opposed to "phoneme segmentation".

5.2 Vowel articulation: tVSA and VAI

To assess vowel articulation, we calculated the triangular Vowel Space Area (tVSA) and the Vowel Articulation Index (VAI), which Skodda, Visser and Schlegel (2011) propose to be more sensitive to dysarthric speech than the tVSA. First, we used scripts for MATLAB, created by Prof. Michael Proctor from Macquarie University, which automatically extracted the first and second formants (F1 and F2) of target vowels in the annotated target words[21]. The input for the scripts were .wav files (speech samples) and .TextGrid files (their accompanying segmentation, i.e. the ORT and MAU tiers described above). The formants for the vowels in the target words of the *carrier phrases* (read speech) and *kwartet game* (semi-spontaneous speech) were saved to .csv files.

The formants were then further analysed in RStudio (R Core Team, 2013). As there were no systematic outliers in the data, we did not remove any vowel values. Further, the tVSA and VAI were calculated according to the following formulae (following Skodda, Visser and Schlegel, 2011), where *F1a*, *F1i* and *F1u* stand for the first formant of vowels /*a*/, /i/, and /*u*/, and *F2a*, *F2i* and *F2u* stand for the second formants.

$$(1)\ tVSA = abs\frac{((F1i*(F2a-F2u)+F1a*(F2u-F2i)+F1u*(F2i-F2a))}{2}$$

$$(2)\ VAI = \frac{F2i+F1a}{F2u+F2a+F1i+F1u}$$

First, the triangular Vowel Space Area (tVSA; equation 1 above) calculates the area filled by the vowel triangle (grey area in Figure 9), which plots the first formant frequency of all three vowels as a function of the second formant frequency (Blomgren, Robb and Chen, 1998). It is not standard procedure to report the actual values obtained through the calculation, however the tVSA can reliably indicate changes in the vowel space area (Turner, Tjaden and Weismer, 1995).

[21] As we did not create the scripts in question, we cannot include them as an appendix.

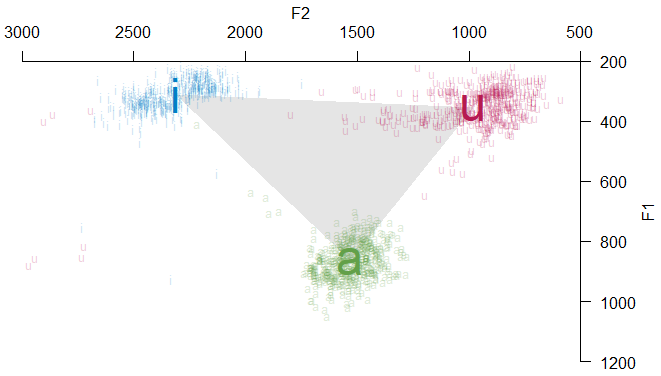


Figure 9: Calculating tVSA (grey area)

Second, the vowel articulation index (VAI; equation 2 above) is a measure of vowel centralization, namely how close vowels are to each other (arrows in Figure 10 indicate distances of each vowel from the center of the vowel space area). The closer the vowels are to the centre of the vowel space area, the lower the elements in the numerator and the higher the elements in the denominator, also leading to a decreased VAI when vowel formants are centralized (Roy, Nissen, Dromey and Sapir, 2009).

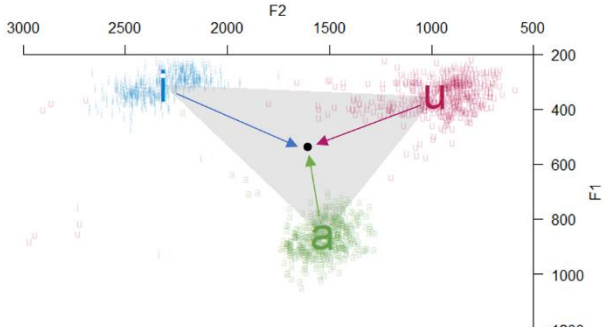


Figure 10: Calculating VAI (distance of each vowel from the center of vowel space)

When studying PD speech, the measures are used with the assumption that PD patients have a reduced vowel space, which will be mirrored in lower tVSA and VAI values. Furthermore, the VAI measure has been proposed as a measure that is more sensitive and suitable than the tVSA when studying PD speech (Skodda, Visser and Schlegel, 2011).

5.3 Voice quality: CPPS and *f0*

Cepstral Peak Prominence (CPP) and *Cepstral Peak Prominence Smoothed* (CPPS) are used to assess voice quality by measuring the degree of harmonic organization (Heman-Ackah, Michael and Goding, 2002). Harmonic organization is, at its basis, a measure of voice breathiness and it assesses the harmonic energy and periodicity of individual peaks in the sound waves (Watts, Awan and Maryn, 2016). CPPS transforms the voice signal from time to

frequency, showing the intensity of each frequency (Balasubramanium et al., 2011). Figure 11 represents the cepstral peak prominence of a healthy speaker.

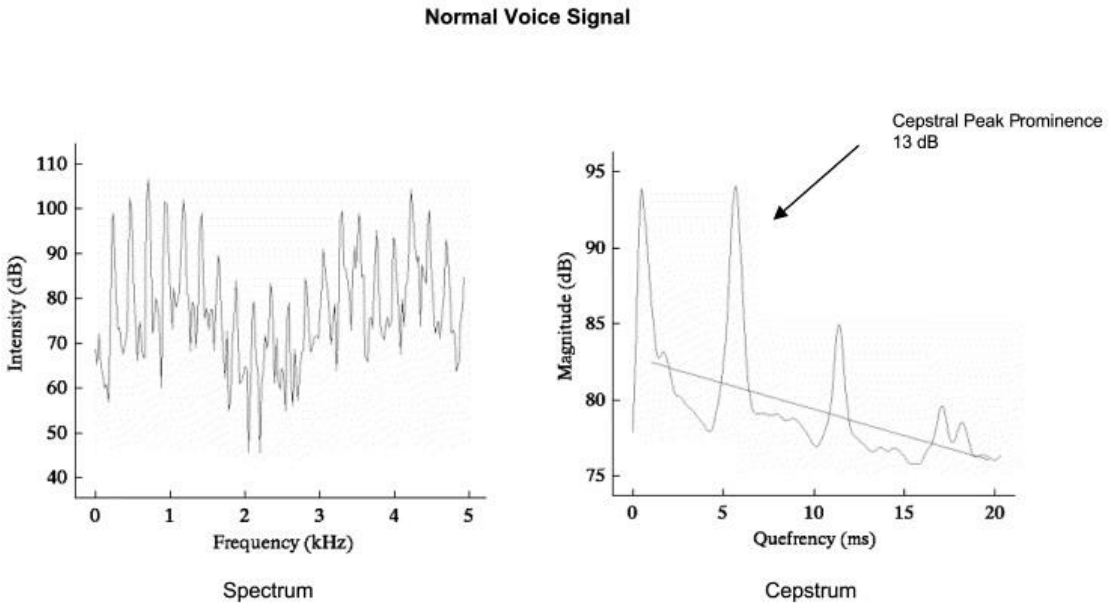


Figure 11: Spectrum and cepstrum (with the cepstral peak prominence) of a healthy speaker (Heman-Ackah, Michael and Goding, 2002).

The measure CPP(S) has been used in several studies on voice disorders (e.g. Maryn et al., 2010) and is considered to be a more reliable measure of pathological voice quality than other measures such as jitter and shimmer (Heman-Ackah et al., 2003). The latter describe variations of fundamental frequency in the voice (jitter) and variations of waveform amplitude (shimmer) (Farrus, Hernando and Ejarque, 2008), but are less robust, as they are more sensitive to accurate extraction of fundamental frequency and accurate measurement of amplitude.[22]

Voice quality analysis in our study was performed on voice samples from each session. The voice samples consisted of approximately 10 seconds of running speech taken from the read-out carrier phrases. We cut three sentences[23] from a single recording and ran them through the freely available software *Speech-Tool* for CPPS analysis (Hillenbrand and Houde, 1996). With it, we obtained CPP and CPPS measures as well as the fundamental frequency (*f0*) of each speech sample. The fundamental frequency, also known as the first harmonic, is the lowest frequency of the voice signal (Figure 12, first sine wave).

[22] Especially the latter is difficult to achieve in non-lab conditions such as ours, as accurate amplitude measurements demand an exact and constant positioning of the microphone throughout all sessions.
[23] We disregarded the first sentence and took the next three consecutive correct sentences. In most cases, these were the second, third and fourth sentence; if any of these phrases contained mispronunciations, the next available phrase was analysed.

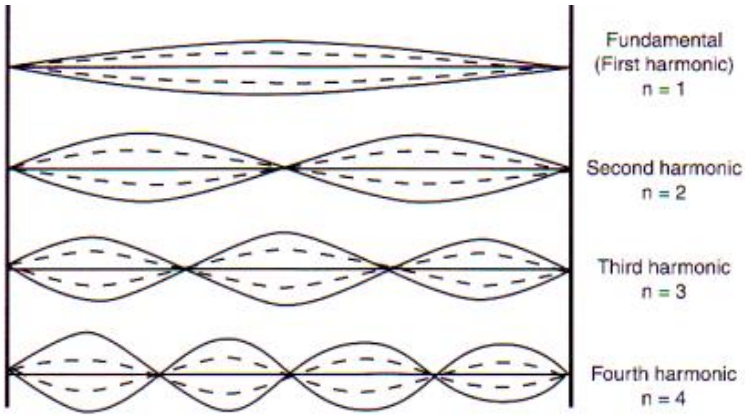


Figure 12: Voice fundamental frequency (source of original picture: https://standingwavesch4.wordpress.com/)

A previous study on voice quality, which used Hillenbrand and Houde's algorithm to study the speech of more than 800 patients with various disorders and 50 healthy volunteers, found that the cut-off value for pathological voice quality shown by CPPS is 4.0 or higher (Heman-Ackah et al., 2014). However, it is unclear whether this value is language-specific or not.

5.4 Statistical analysis

Statistical analysis was performed in RStudio, namely we built multiple linear mixed-effects regression models with help of the *lme4* package (Bates, Mächler, Bolker and Walker, 2015). Mixed models are *mixed* because they include both fixed effects (i.e. the variables that we control experimentally) and random effects (i.e. the variables that are out of our control) (Winter, 2013). Due to our highly heterogenous sample and differing number of vowel instances per speaker, it was crucial to carry out analysis that could reliably account for differences between subjects and perform well in case of missing data or unevenly distributed data points.

For our analysis, we first created three columns: Task, Type, Score. "Task" contained two tasks, namely wordlist and kwartet. All four measures were available for the wordlist data, but only vowel articulation (tVSA and VAI) measures were available for the kwartet data. Consequently, we could only distinguish between tasks for vowel articulation measures, not voice quality. "Type" contained the four different recorded measures (tVSA, VAI, *f0* and CPPS) and "Score" contained the scores from the individual measures. When testing the first hypothesis (comparing PD patients and healthy control speakers), we *z*-transformed the scores by type of

measure and task. When testing the second hypothesis (comparing OFF and ON states of PD patients), we *z*-transformed the scores by type of measure, task, and participant.

For testing the hypotheses, we built a simple model, testing the effect of predictor group (for the first hypothesis) or predictor state (for the second hypothesis) on the *z*-transformed scores. We included participants as random intercepts. After building this simplest model, we proceeded with an exploratory analysis by including different predictors, i.e. fixed effects. Our fixed effects included predictors related to the participant (namely group, sex and fatigue levels), our measures (namely task and type of measure), time of recording (namely time of day, session and day) and medication (namely state – ON/OFF – and effect – OFF/+1 hour/+2 hours). We also tested for interactions between different predictors (e.g. group and sex, task and sex, time of day and fatigue). When testing and adding different predictors, we always changed only one part of the model in order to ensure it would be comparable to the previous model.

For model comparison, we checked significance and measured goodness of fit using the Akaike Information Criterion (AIC; Akaike, 1974). This represented a means for model selection: we chose the more complex model if the AIC dropped by 2 units or more (Wieling, 2018). Participants were included as random intercepts throughout our analysis. We tested additional random effects and tested for correlation parameters in random effects based on significant predictors. For the first hypothesis (comparing PD patients and healthy speakers) this led to the inclusion of the variable "Type" as a correlated random slope.

Furthermore, for every model, we calculated effects sizes (namely Cohen's d), where $d = 0.2$ represents a small effects size, $d = 0.5$ represents a medium effects size and $d = 0.8$ represents a large effects size. For our best-fitting model, we tested the required assumptions, including multicollinearity, autocorrelation, normality and heteroscedasticity. For the results of our analysis see Section 6.3 and Section 6.4 below. See Appendix 12 for our statistical analysis.

6 Results

This section presents the trends in the recorded data and describes the statistical analysis, including our hypothesis test, exploratory analysis and best-fitting linear mixed-effects model. First, we describe the data collection size (Section 6.1). This is followed by descriptive statistics on the four measures (Section 6.2), including two vowel articulation measures (tVSA and VAI) and two voice quality measures (*f0* and CPPS). Each measure subsection describes individual measures and averages for PD patients and healthy control speakers but also averages for men and women.

As men and women differ significantly in their acoustic characteristics, it is necessary to keep the differences in mind when presenting results, especially considering our sample, where the patients were predominantly men, and the healthy controls predominantly women. Finally, each subsection concludes with averages for two different levodopa states (state: ON/OFF and effect: OFF/+1 hour/+2hours).

Next are two sections on our main hypotheses: both sections include the outcomes of the hypothesis test as well as an exploratory analysis to ensure the hypothesis tests are valid (i.e. the mixed-effects model that best describes our data). More specifically, Section 6.3 describes the differences between the two groups (Parkinson's disease patients versus healthy control speakers), while Section 6.4 describes the differences between the Levodopa ON and OFF states for PD patients. The chapter is concluded with a case study (Section 6.5), namely a description of acoustic characteristics of the monozygotic twins who participated in our study.

6.1 Data collection size

We obtained a total of 2067 instances for vowel /a/, 2073 instances for vowel /i/ and 1169 instances for vowel /u/ (see Table 8 and Table 9 below for the distribution across tasks, groups and participants).

		Per task		**Per group**	
Vowel	**Total**	*Kwartet*	*Wordlist*	*PD*	*HC*
/a/	2067	883	1184	1269	789
/i/	2073	893	1180	1254	819
/u/	1169	854	1169	1253	770

Table 8: The number of extracted vowels, in total and per task

Participant ID																				
	01		**02**		**03**		**04**		**05**		**06**		**07**		**08**		**09**		**10**	
V	W	K	W	K	W	K	W	K	W	K	W	K	W	K	W	K	W	K	W	K
/a/	115	67	116	53	119	107	120	108	118	76	120	94	118	116	121	121	119	86	118	55
/i/	110	53	110	77	120	111	121	102	119	96	120	98	120	112	120	99	120	70	120	75
/u/	111	71	110	46	118	91	119	105	119	99	120	69	115	116	119	124	119	70	119	63

Table 9: The number of extracted vowels per participant (W denotes the wordlist task, K denotes the kwartet task)

Due to the experimental design, we obtained approximately six instances of each corner vowel per each session in the wordlist task, but some missing vowels in individual sessions of the kwartet task. Per participant, we obtained approximately 100 instances of each of the 36 target words in the wordlist task, and approximately 150 additional instances of the 18 target words chosen for the kwartet task.

For voice quality data, we obtained the mean CPPS and mean *f0* measures for all sessions, leading to 20 CPPS and 20 *f0* measures per participant.

6.2 Descriptive statistics

6.2.1 Vowel space areas and vowel formants

6.2.1.1 Individual vowel space areas

Based on the obtained data, we first plotted vowel spaces areas for all participants. Figure 13 is a representation of the vowel space areas per participant, depicting all vowel instances across the 20 sessions. Above each vowel space area there is the participant number (Participant 1 – 10) as well as group (PD denotes Parkinson patients and HC denotes healthy control speakers) and sex (F denotes women and M denotes men). For example, the title “Participant 6 (F-PD)” means that the participant in question is a female Parkinson patient with subject code 6.

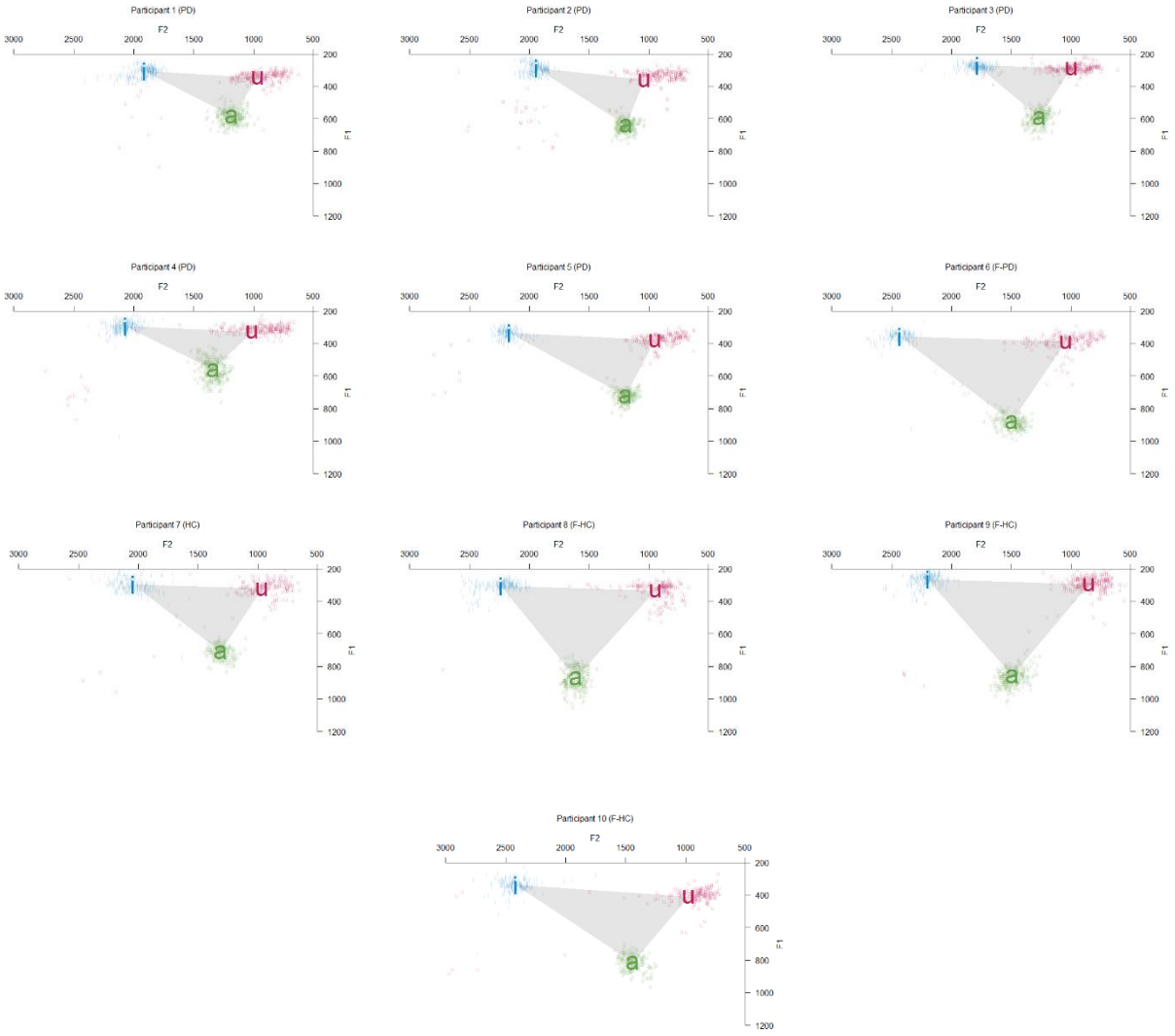


Figure 13: Vowel space areas per participant

6.2.1.2 Average vowel formants for PD and HC

Table 10 presents the average vowel formants for the three measured corner vowels, compared between Parkinson's disease patients (PD group) and healthy controls (HC group). Each table (in this section and sections below) reports the mean values ± 1 standard deviation. Formants for vowels /i/ and /a/ are generally lower or similar in the PD group compared to the HC group, while formants for /u/ are higher in the PD group. Figure 14 is a depiction of the groups' vowel space areas. The vowel space area is smaller for PD patients compared to healthy controls. When looking at the plots of PD patients, we can see a thick group of data points outside the bounds of the vowel triangle, indicating bimodality in the data that we plotted. We can presume that this thick cloud is due to the female PD speaker whose vowel space area is bigger than that of male PD speakers.

Vowel →	/i/		/a/		/u/	
Group ↓	f1	f2	f1	f2	f1	f2
PD	305 ± 47	2032 ± 216	656 ± 115	1272 ± 130	346 ± 78	999 ± 313
HC	300 ± 51	2226 ± 174	821 ± 84	1473 ± 135	337 ± 89	925 ± 272

Table 10: Average vowel formants per group

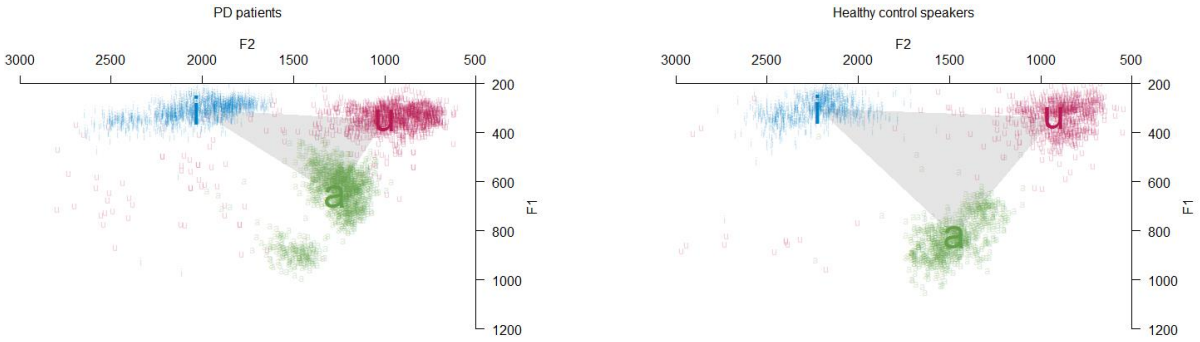


Figure 14: Vowel space areas per group

6.2.1.3 Averages for men and women

Table 11 presents the average vowel formants for the three measured corner vowels, compared between men and women. Formants are generally lower for men than women, excepting the second formant of vowel /u/, which is higher for the men. The vowel space area is smaller for men compared to women (also seen in Figure 13). Furthermore, unlike in the plots above (Figure 15), there are no groups of data points that would lie outside the bounds of the vowel triangle, indicating that there is no bimodality and gender accounts well for the acquired data points (i.e. the vowel space of the female PD patient is closer to the vowel space of female healthy control speakers than to the vowel space of other, male, PD patients).

Vowel →	/i/		/a/		/u/	
Gender ↓	f1	f2	f1	f2	f1	f2
Men	297 ± 42	1982 ± 155	630 ± 76	1245 ± 93	337 ± 80	988 ± 321
Women	312 ± 56	2311 ± 151	858 ± 62	1512 ± 109	352 ± 85	943 ± 263

Table 11: Vowel formants per gender

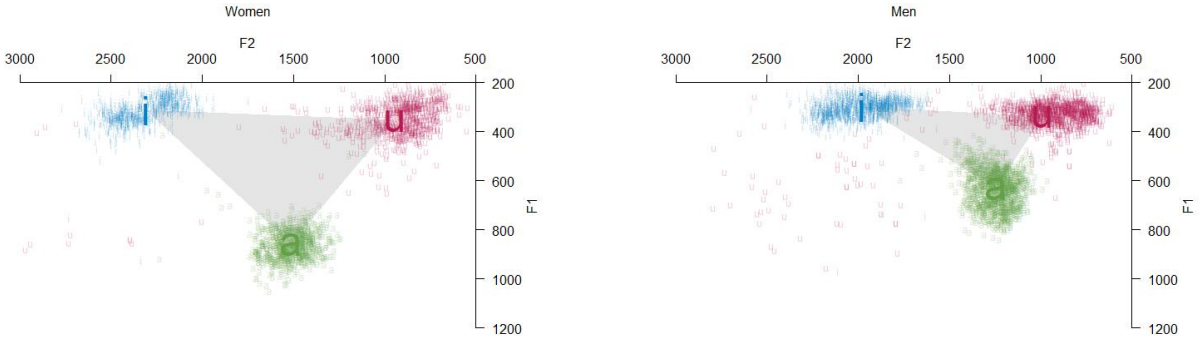


Figure 15: Vowel space areas per gender (women: left; men: right)

As most of our patients were men and most of our healthy controls were women, the patterns for gender comparison are similar to the pattern seen in PD patients versus healthy controls (Section 6.2.1.2 above). Group and sex are therefore highly collinear, and we need to be careful in interpreting our data. We can also see the similarities between group and sex when we plot vowel space areas for each group (PD patients and healthy controls) but separate them by gender (see Figure 16 below).

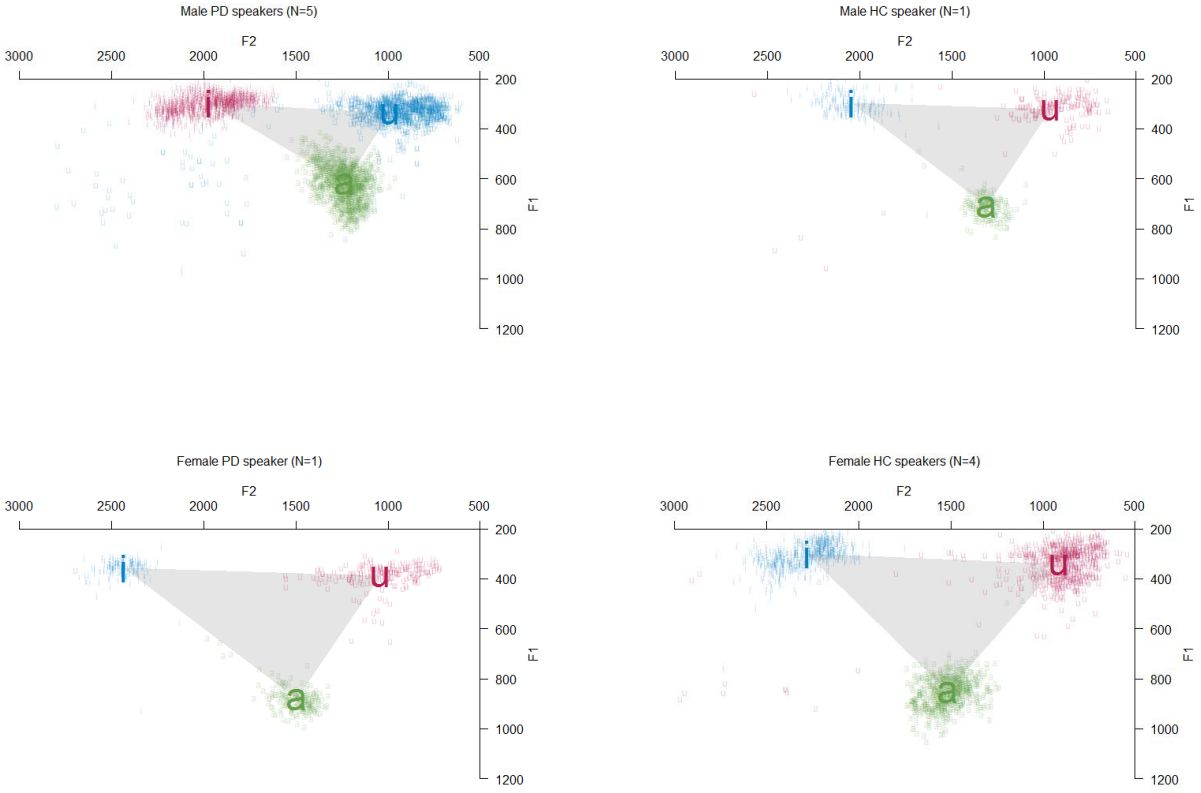


Figure 16: Vowel space areas for men above (PD left; HC right) and women below (PD left; HC right)

6.2.1.4 Average vowel formants across the levodopa cycle

Table 12 presents the average vowel formants for the three measured corner vowels, compared across state (OFF / ON) and levodopa cycle (effect; OFF / +1 hour / +2 hours). Formants for all three vowels are slightly lower or equal in the OFF state compared to the ON state.

This can also be seen in the corresponding vowel space areas (Figure 17). There is greater dispersion of vowels in the ON state compared to the OFF state.

Vowel →	/i/		/a/		/u/	
State ↓	f1	f2	f1	f2	f1	f2
OFF	305 ± 71	2024 ± 222	643 ± 122	1270 ± 122	337 ± 72	979 ± 285
ON	305 ± 39	2034 ± 214	659 ± 113	1272 ± 132	348 ± 79	1003 ± 319
Effect ↓	f1	f2	f1	f2	f1	f2
OFF	305 ± 71	2024 ± 222	643 ± 122	1270 ± 122	337 ± 72	979 ± 285
+1 hour	304 ± 38	2034 ± 213	662 ± 115	1276 ± 132	351 ± 85	1015 ± 346
+ 2 hours	305 ± 40	2034 ± 216	655 ± 112	1268 ± 131	345 ± 73	992 ± 292

Table 12: Average vowel formants per state and effect

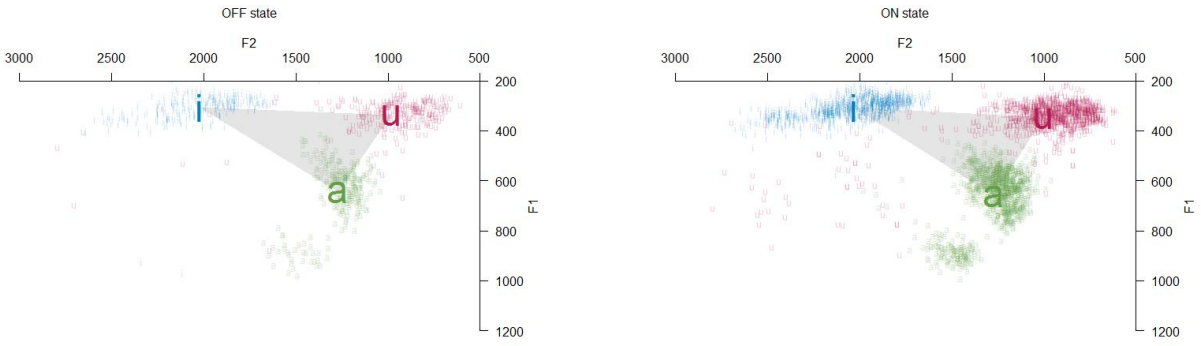


Figure 17: Vowel space areas in OFF state and ON state

Figure 18 is a plot of vowel space area per state, where the dotted line represents the ON state and the full line represents the OFF state.

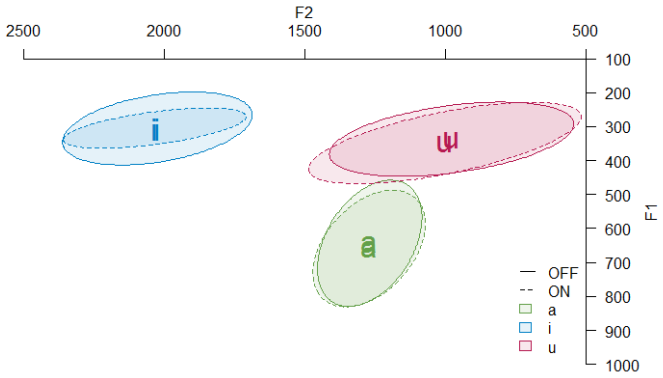


Figure 18: Vowel space areas (ON/OFF state in one plot)

Furthermore, compared to the OFF state, formants are slightly higher 1 hour after the medication has taken effect and decrease again by the second hour (Table 12 above). Figure 19 displays vowel space areas per medication cycle.

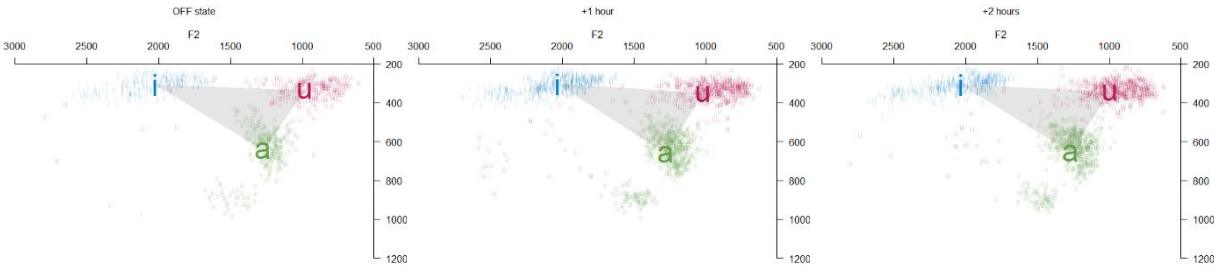


Figure 19: Vowel space areas in OFF state, 1 hour and 2 hours after levodopa intake

Figure 20 is a plot of vowel space area through the levodopa cycle, where the green colour represents the OFF state, blue colour represents the effect 1 hour after levodopa intake, and red colour represents the effect 2 hours after levodopa intake.

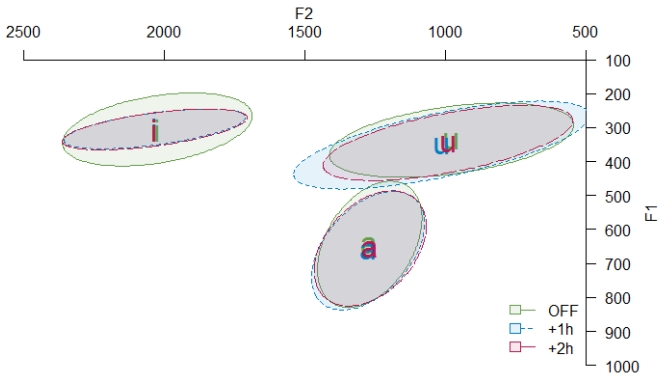


Figure 20: Vowel space areas throughout the levodopa cycle

6.2.2 Vowel articulation: tVSA

We calculated the triangular vowel space area (tVSA) for all participants using the method described in Section 5.2.

6.2.2.1 Individual tVSA measures

Figure 21 displays mean tVSA measures per participant (boxplot, left) and per each of the five sessions per participant (bar chart, right). For figures displaying individual measures, the first 6 participants are PD patients, followed by the 4 healthy controls. Women are marked with (f) in all figures. tVSA measures are higher in healthy control participants than PD patients, and higher in women than men.

Unlike other measures used in this study, tVSA values usually aren't reported on their own (Turner, Tjaden and Weismer, 1995). We chose to nonetheless report the values, which can lie on a scale from 100 000 Hz^2 and above, and the accompanying standard deviation.

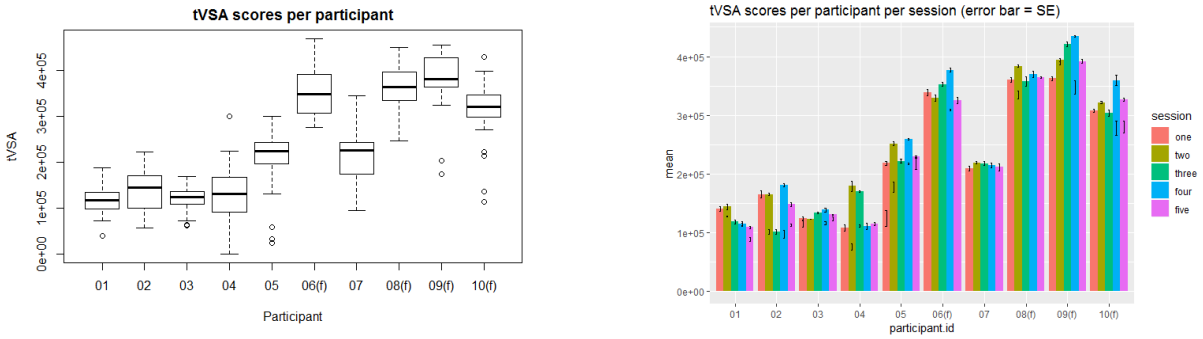


Figure 21: tVSA scores per participant (left) and per session (right)

6.2.2.2 Average tVSA measures for PD and HC

Figure 22 displays tVSA scores per group, namely comparing healthy control speakers and PD patients. Mean tVSA scores are higher for healthy control speakers (*325,115 ± 81,300*) than for PD patients (*169,540 ± 88,277*).

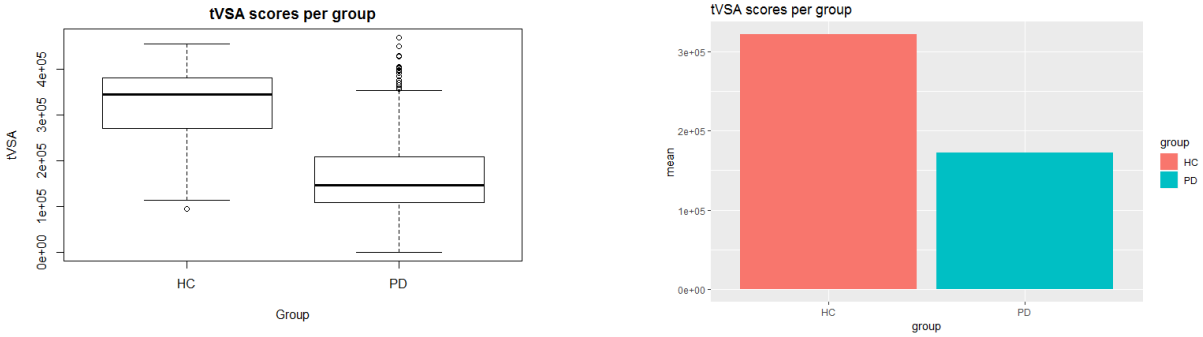


Figure 22: tVSA scores per group (PD patients versus healthy control speakers)

6.2.2.3 Average tVSA measures for men and women

Figure 23 displays tVSA scores per gender, comparing men and women. Mean tVSA scores are higher for women (*352,800 ± 60,436*) than for men (*152,451 ± 59,507*). This mirrors the pattern seen when comparing PD patients and healthy speakers (Section 6.2.2.2).

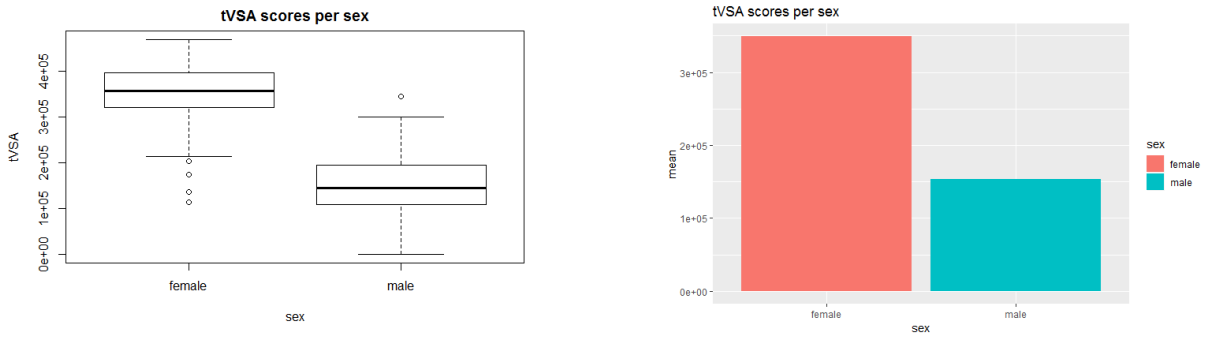


Figure 23: tVSA scores per gender

6.2.2.4 Average tVSA measures across the levodopa cycle

Figure 24 displays the effect of state (ON / OFF) on the mean tVSA score of PD patients as a group (boxplot, left) and on the mean tVSA score of each individual PD patient (bar chart, right). The mean tVSA score for PD patients as a group is *170,102 ± 85,942* in the OFF state and *169,407 ± 88,834* in the ON state.

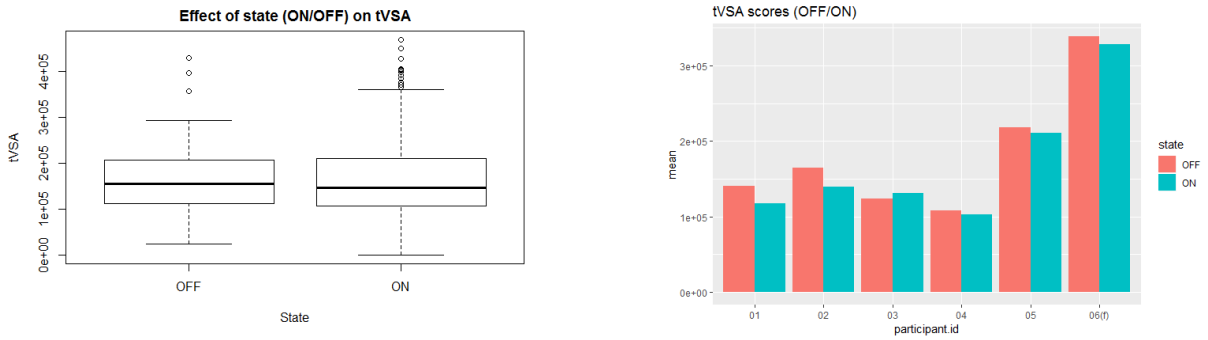


Figure 24: Effect of state on tVSA

Figure 25 displays the effect of levodopa cycle (OFF / + 1h / +2h) on the mean tVSA score of PD patients as a group (boxplot, left) and on the mean tVSA score of each individual PD patient (bar chart, right). The mean tVSA score for PD patients as a group is *170,102 ± 85,942* in the OFF state, *169,345 ± 89,108* one hour after levodopa intake, and *169,471 ± 88,582* two hours after the intake.

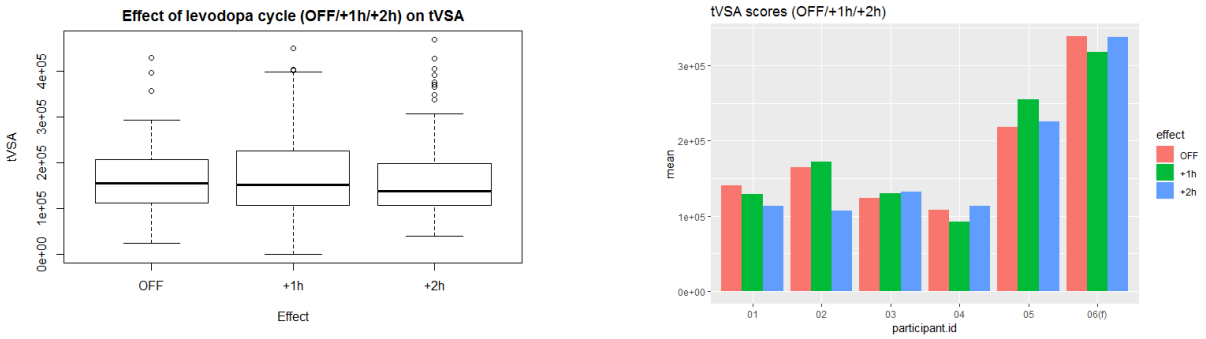


Figure 25: Effect of levodopa on tVSA

6.2.3 Vowel articulation: VAI

We calculated the vowel articulation index (VAI) for all participants using the method described in Section 5.2. The two vowel articulation measures (tVSA and VAI) are highly correlated (r = 0.836).

6.2.3.1 Individual VAI scores

Figure 26 displays mean VAI score per participant (boxplot, left) and per each of the five sessions per participant (bar chart, right).

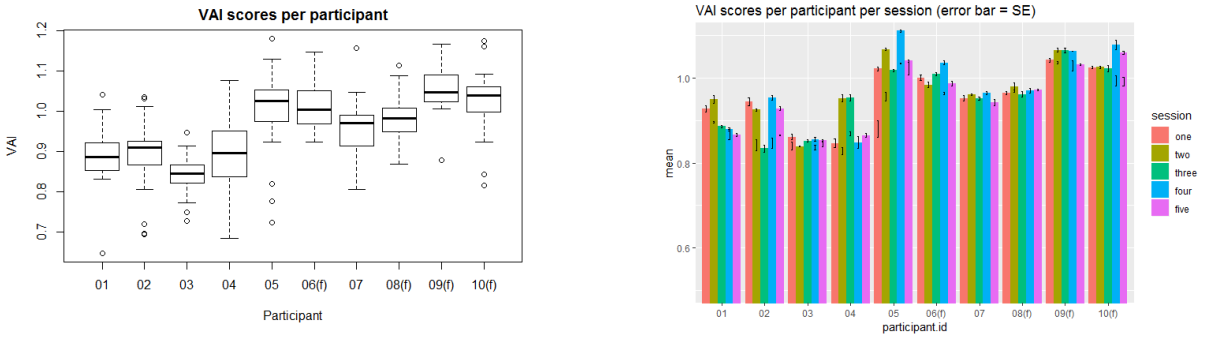


Figure 26: Mean VAI scores per participant (left) and per session (right)

6.2.3.2 Average VAI scores for PD and HC

Figure 27 displays the average VAI scores per group, comparing PD patients and healthy control speakers. Mean VAI scores are higher for healthy control speakers (*1.00 ± 0.068*) than for PD patients (*0.92 ± 0.093*).

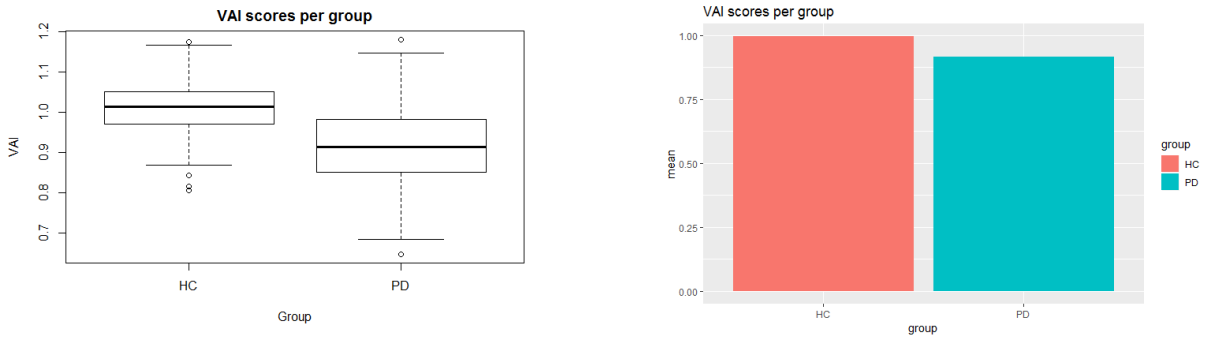


Figure 27: VAI scores per group (PD patients versus healthy control speakers)

6.2.3.3 Average VAI scores for men and women

Figure 28 displays VAI scores per gender. Mean VAI scores are higher for women (*1.01 ± 0.063*) than men (*0.91 ± 0.088*).

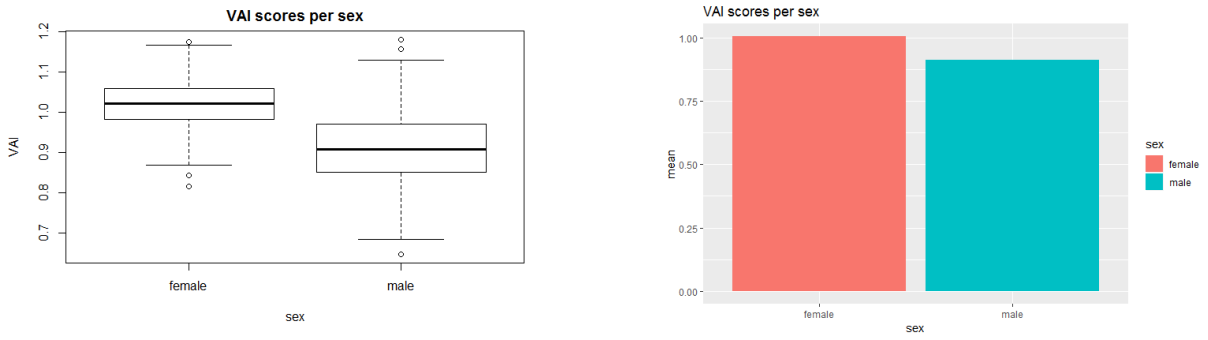


Figure 28: VAI scores per gender

6.2.3.4 Average VAI scores across the levodopa cycle

Figure 29 displays the effect of state (ON / OFF) on the mean VAI score of PD patients as a group (boxplot, left) and on the mean VAI score of each individual PD patient (bar chart, right). The mean VAI score for PD patients as a group is *0.93 ± 0.093* in the OFF state and *0.92 ± 0.093* in the ON state.

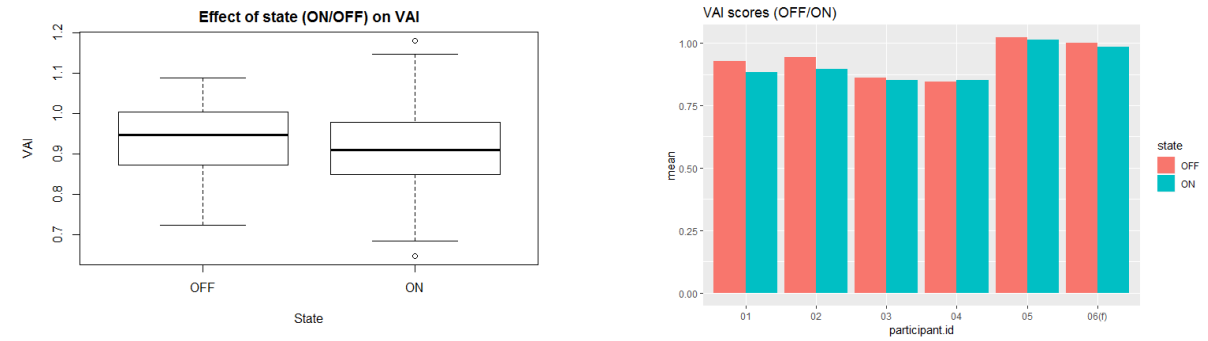


Figure 29: Effect of state on VAI

Figure 30 displays the effect of levodopa cycle (OFF / + 1h / +2h) on the mean VAI score of PD patients as a group (boxplot, left) and on the mean VAI score of each individual PD patient (bar chart, right). The mean VAI score for PD patients as a group is *0.93 ± 0.093* in the OFF state, *0.92 ± 0.1* one hour after levodopa intake, and *0.92 ± 0.084* two hours after the intake.

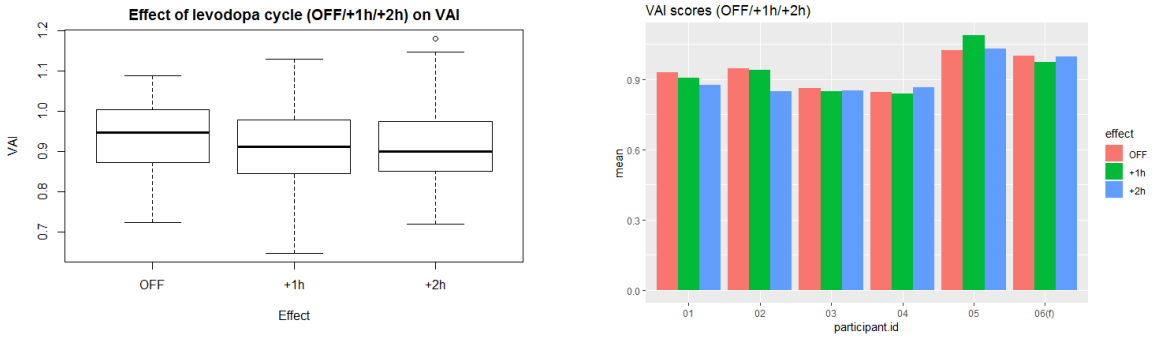


Figure 30: Effect of levodopa cycle on VAI

6.2.4 Voice quality: *f0*

We obtained the mean fundamental frequency (*f0*) using the method described in Section 5.3.

6.2.4.1 Individual f0

Figure 31 displays mean *f0* per participant (boxplot, left) and per each of the five sessions per participant (bar chart, right). Unfortunately, as a measure, fundamental frequency is not very informative on its own with a small number of speakers (as it is only a report of a person's pitch). Instead, it would have been more informative to look at the changes in fundamental frequency or at its dispersion.

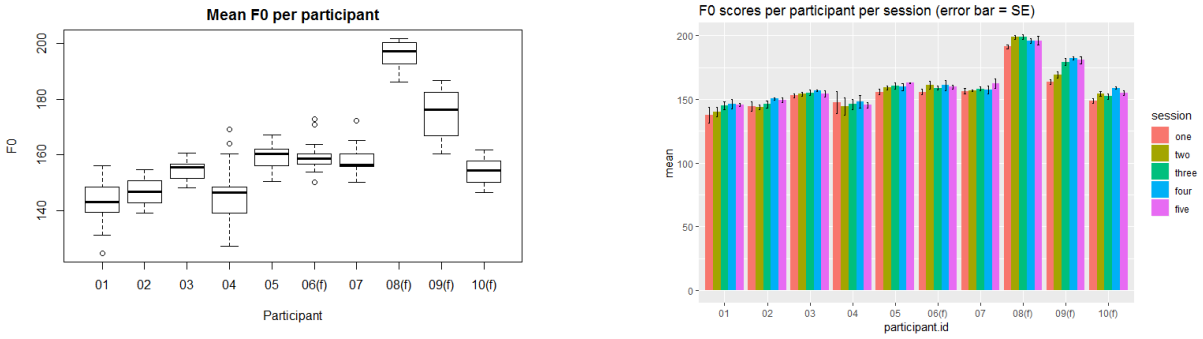


Figure 31: Mean f0 per participant (left) and per session (right)

6.2.4.2 Average f0 for PD and HC

Figure 32 displays the average *f0* per group, comparing PD patients and healthy control speakers. Mean *f0* is higher for healthy control speakers (*170.8 ± 17.7*) than for PD patients (*151.5 ± 9.14*).

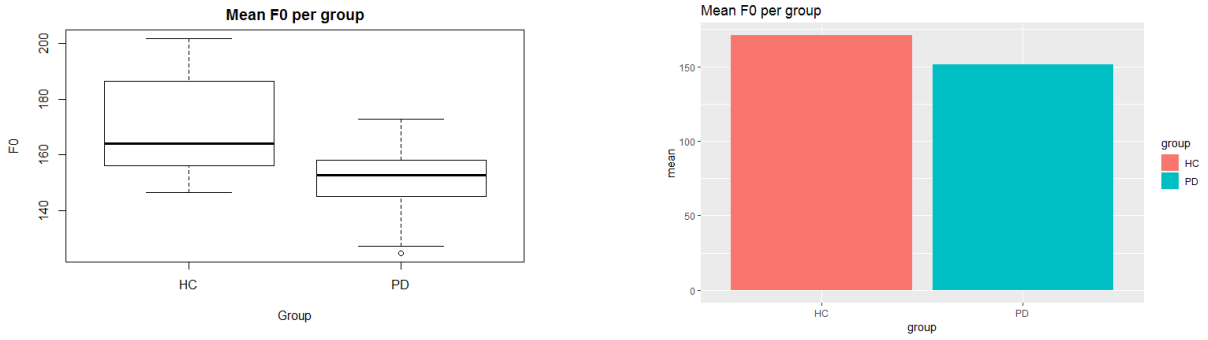


Figure 32: f0 per group (PD patients versus healthy control speakers)

6.2.4.3 Average f0 for men and women

Figure 33 displays mean *f0* per gender. Mean *f0* is higher for women (*171.0 ± 17.5*) than men (*151.3 ± 9.00*).

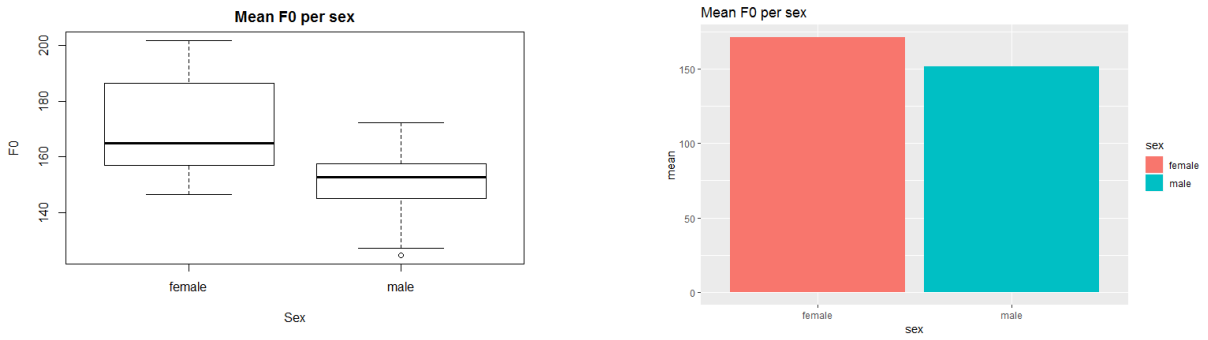


Figure 33: f0 per gender

6.2.4.4 Average f0 across the levodopa cycle

Figure 34 displays the effect of state (ON / OFF) on the mean *f0* of PD patients as a group (boxplot, left) and on the mean *f0* of each individual PD patient (bar chart, right). The mean *f0* for PD patients as a group is *148.9 ± 10.8* in the OFF state and *152.1 ± 8.6* in the ON state.

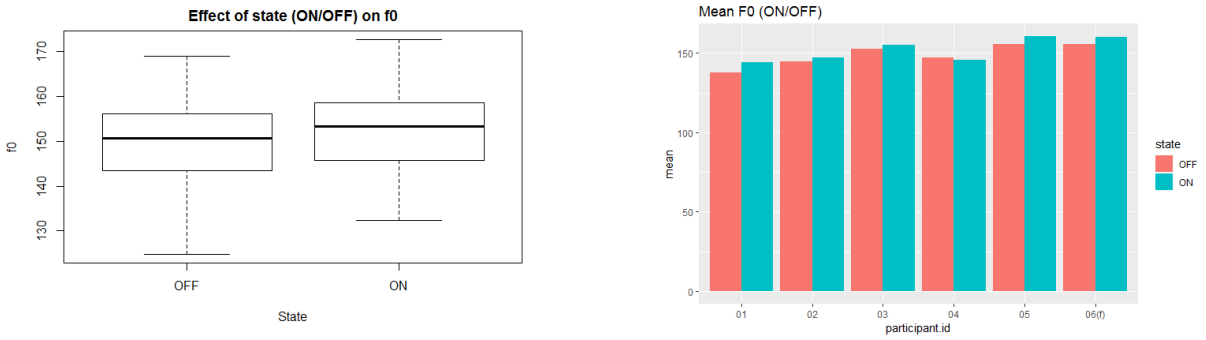


Figure 34: Effect of state on f0

Figure 35 displays the effect of levodopa cycle (OFF / + 1h / +2h) on the mean *f0* of PD patients as a group (boxplot, left) and on the mean tVSA score of each individual PD patient (bar chart, right). The mean *f0* for PD patients as a group is *148.9* ± *10.8* in the OFF state, *152.0* ± *9.5* one hour after levodopa intake, and *152.3* ± *7.8* two hours after the intake.

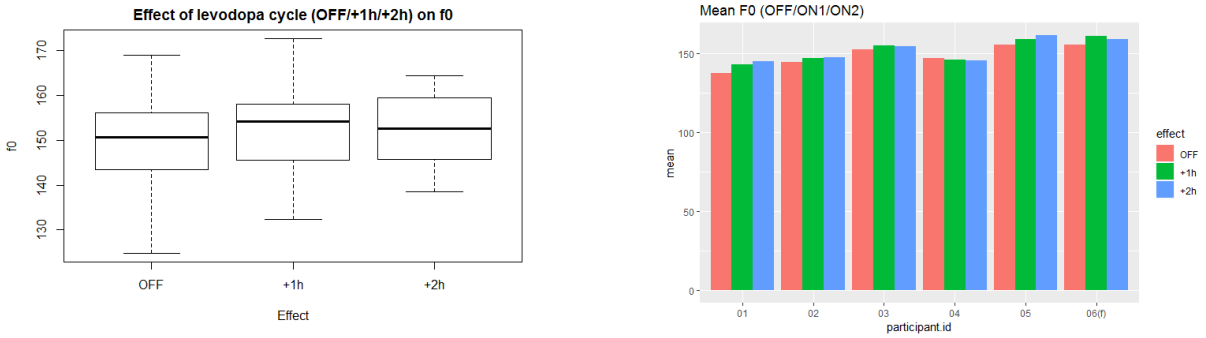


Figure 35: Effect of levodopa cycle on f0

6.2.5 Voice quality: CPPS

We obtained the cepstral peak prominence smoothed (CPPS) using the method described in Section 5.3.

6.2.5.1 Individual measures

Figure 36 displays mean CPPS score per participant (boxplot, left) and for each of the five sessions per participant (bar chart, right).

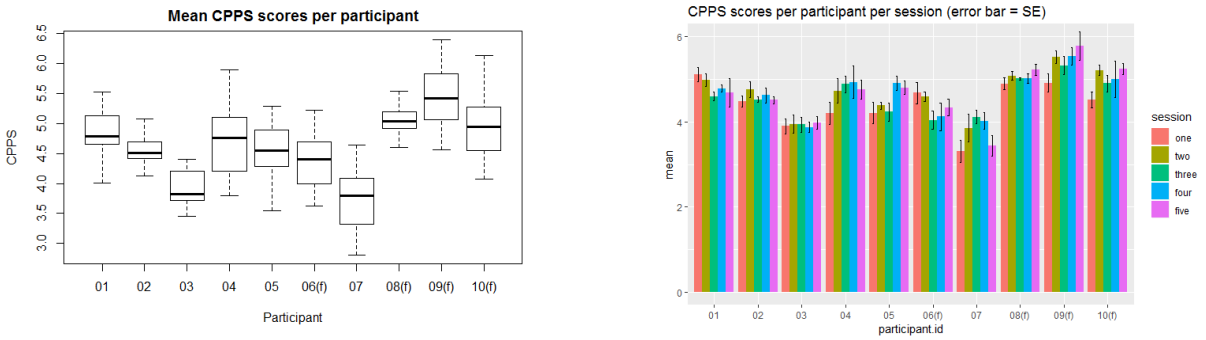


Figure 36: Mean CPPS per participant (left) and per session (right)

6.2.5.2 Average CPPS measures for PD and HC

Figure 37 displays the mean CPPS scores per group, comparing PD patients and healthy control speakers. Mean CPPS scores are higher for healthy control speakers (*4.80 ± 0.78*) than for PD patients (*4.48 ± 0.50*).

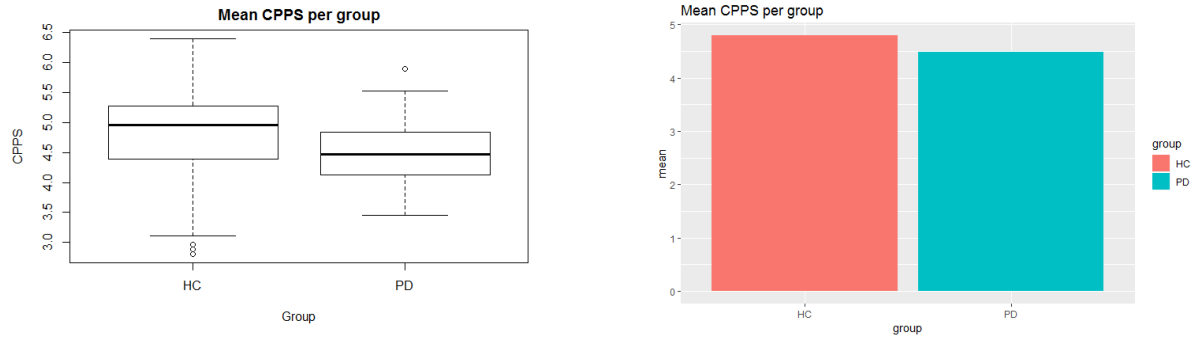


Figure 37: Mean CPPS per group (PD patients versus healthy control speakers)

6.2.5.3 Average CPPS measures for men and women

Figure 38 displays mean CPPS scores per gender. Mean CPPS scores are higher for women (*4.95 ± 0.58*) than men (*4.38 ± 0.58*).

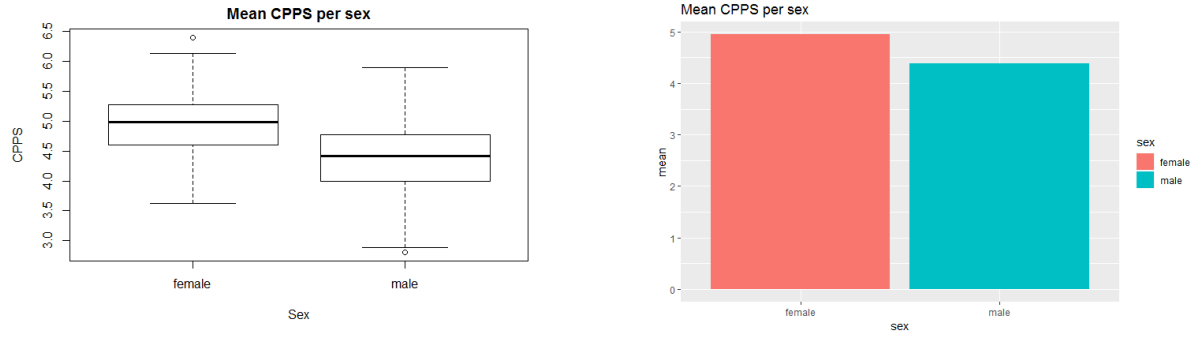


Figure 38: Mean CPPS per gender

6.2.5.4 Average CPPS measures across the levodopa cycle

Figure 39 displays the effect of state (ON / OFF) on the mean CPPS score of PD patients as a group (boxplot, left) and on the mean CPPS score of each individual PD patient (bar chart, right). The mean CPPS score for PD patients as a group is *4.43 ± 0.55* in the OFF state and *4.50 ± 0.49* in the ON state.

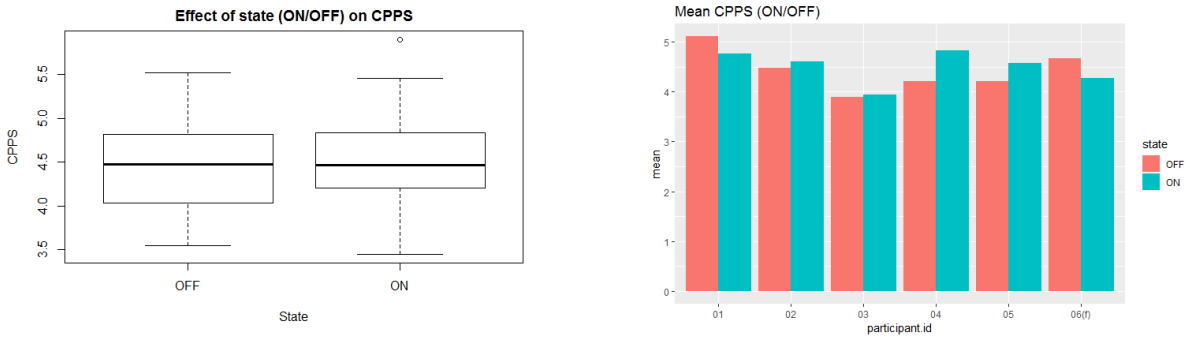


Figure 39: Effect of state on mean CPPS

Figure 40 displays the effect of levodopa cycle (OFF / + 1h / +2h) on the mean CPPS score of PD patients as a group (boxplot, left) and on the mean CPPS score of each individual PD patient (bar chart, right). The mean CPPS score for PD patients as a group is *4.43 ± 0.55* in the OFF state, *4.55 ± 0.53* one hour after levodopa intake, and *4.44 ± 0.46* two hours after the intake.

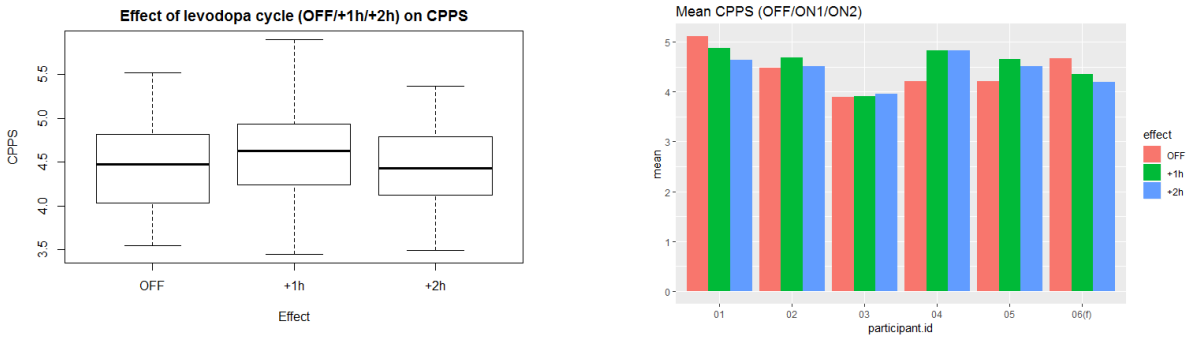


Figure 40: Effect of levodopa cycle on mean CPPS

6.2.6 Individual variability

An important observation in the acoustic parameters we measured is high individual variability, i.e. the randomness in the patterns across different days, especially in the speech of Parkinson's disease patients. Figure 41 (VAI scores of Participant 5) and Figure 42 (CPPS scores of Participant 5) below are examples of such variability across days for a male PD patient. Looking at Figure 41, we can see that VAI scores are higher in the first session compared to the second

session on Day 1 and Day 2. As the first session of each day is the only OFF state and higher VAI values indicate improved speech (i.e. increased vowel articulation), this would indicate that speech in the OFF state is better than speech in the ON state. Consequently, this could lead to the assumption that, at least for this patient, levodopa makes speech worse. However, looking at Day 3 and Day 4, we can see that the pattern is flipped and the VAI measure is in fact lower in the first session (OFF state) compared to the second session, indicating an improvement in speech. Only considering these two days could lead to the assumption that levodopa positively affects speech (or, better, vowel articulation).

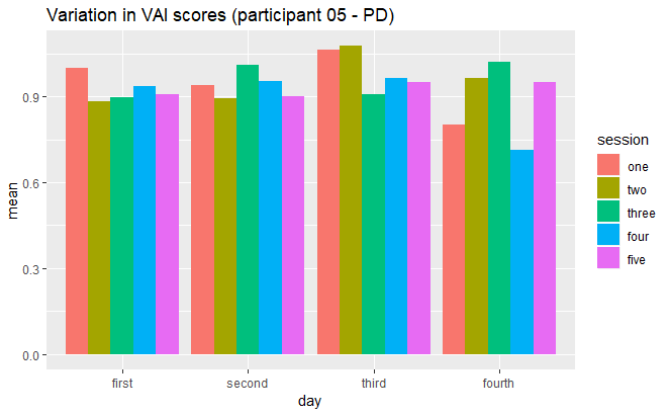


Figure 41: Mean VAI scores across all sessions for participant 05

Looking at CPPS measures (Figure 42) is similarly tricky. When the CPPS measure is lower, it means that speech is more pathological. For this patient, the CPPS value is at its lowest in the OFF state for the first three days and increases in the first ON state. The pattern for the first three days thus suggests that levodopa, in fact, improves voice quality as expressed by the CPPS measure. This pattern, however, is flipped on Day 4, as it is lower in the second session. If the only recording had been made on Day 4, this could have led to the assumption that levodopa makes voice quality worse, although the pattern on other days suggests the opposite.

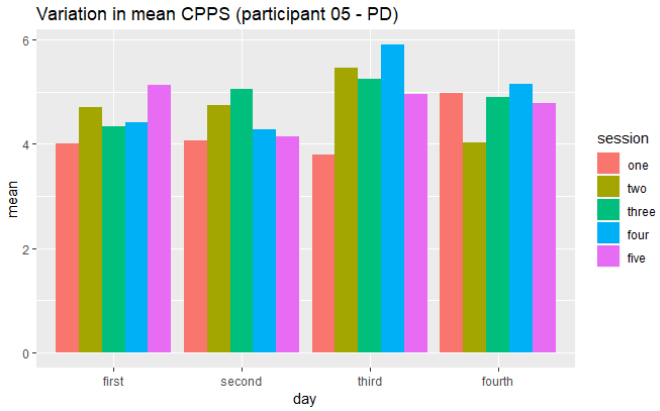


Figure 42: Mean CPPS scores for participant 7 across all sessions (PD)

Variability across measures occurs for all participants, both healthy speakers as well as PD patients. While such variability is difficult to account for in its entirety, it is crucial to know that it exists and needs to be considered when interpreting results.

6.3 Hypothesis 1: Comparing PD patients and healthy speakers

The first hypothesis we tested concerned the difference between PD patients and healthy speakers.

6.3.1 Hypothesis test

We tested whether there was an effect of group on the score, namely with "group" as a fixed effect and participants as a random effect. The predictor was significant (b = -0.98, $t(8)$ = -2.6, p = 0.03, d = -1.8), indicating that there is a significant difference in the speech of PD patients compared to healthy speakers (Figure 43). The effect size was very large (1.8 standard deviations difference).

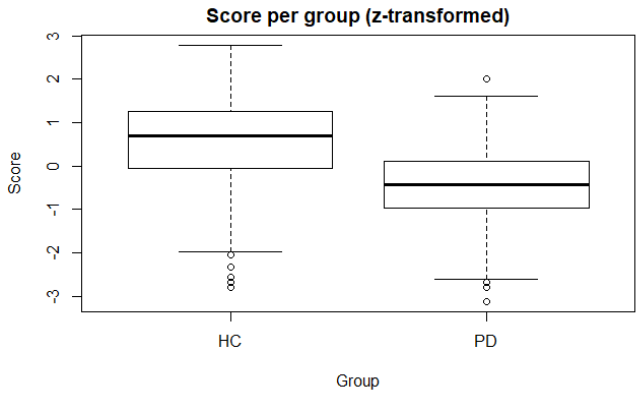


Figure 43: The effect of group on general score

However, from our visualizations of individual scores (see Section 6.2), we know that variable sex (male / female) affects speech as much as variable group (PD / HC) and that, due to our sample, the two variables are highly collinear. Therefore, a further exploratory analysis is needed to assess whether these results are also present when taking into account potentially important covariates.

6.3.2 Exploratory analysis: effect of gender

We performed our exploratory analysis as described in Section 5.4. Model comparison revealed that sex is a better predictor of score than group ($p < 0.001$, AIC lower by 18 units in model with *sex*). Figure 44 shows the effect of sex and group on *z*-transformed scores.

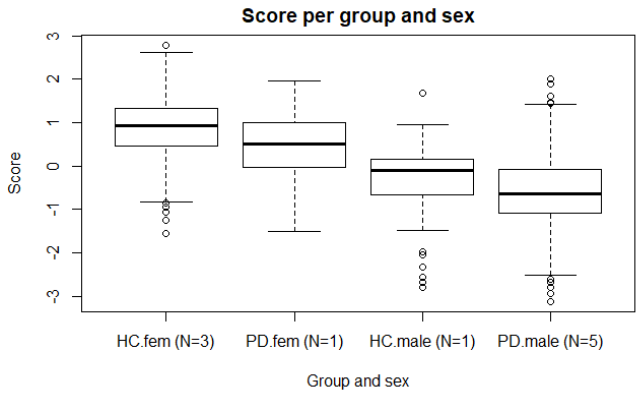


Figure 44: Score per group and sex

Due to our hypothesis, however, we chose to include group as a fixed factor alongside sex, even though this did not significantly improve the model ($p = 0.5$, AIC lower by 1.5 units). The final model thus included both sex ($b = -1.9$, $t(7.5) = -11.5$, $p < 0.001$, $d = -8.4$) and group ($b = -0.1$, $t(7.5) = -0.6$, $p = 0.54$, $d = -0.5$) as fixed factors, with type and participants as correlated random intercept and slope. The effect size was very large for sex ($d = -8.4$) and medium for group ($d = -0.5$). There were no other significant predictors found in the exploratory analysis.

6.3.3 Summary of best-fitting model

We built a linear mixed-effects model to test our null hypothesis that there is no difference between healthy Slovenian speakers and Slovenian speakers with Parkinson's disease. We first built a hypothesis-testing model, which assessed the effect of group on score. This model indicated a significant difference between the two groups ($b = -0.98$; $t(8) = -2.6$; $p = 0.03$, $d = -1.8$), leading to the interim conclusion that PD patients' speech is deteriorated compared to healthy speakers. However, when further proceeding with the exploratory analysis, we found that gender is a better predictor. This is especially important considering that most our controls were female and most our patients were male.

After performing an exploratory analysis (described in Section 6.3.2 above), we determined that the best model describing our data included sex ($b = -1.9$, $t(7.5) = -11.5$, $p < 0.001$, $d = -8.4$)

and group (included due to our hypothesis; non-significant, b = -0.1, t(-0.6), p = 0.54, d = -0.5) as fixed factors, with type of measure and participants as correlated random intercept and slope.

6.4 Hypothesis 2: Comparing OFF/ON state of PD patients

The second hypothesis we tested concerned the difference between the OFF and ON states of PD patients.

6.4.1 Hypothesis test

We tested whether there was an effect of state (ON/OFF) on the score, namely with "state" as a fixed effect and participants as a random effect. The model indicated that the values were slightly higher during the ON state (Figure 45), however this was not significant with a small effect size (b = 0.05, t(678) = 0.519, p = 0.6, d = 0.04).

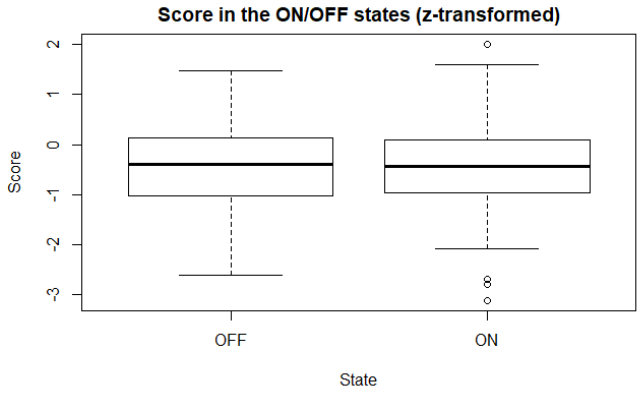


Figure 45: Effect of state on score

Before proceeding with the exploratory analysis, we tested whether medication cycle (i.e. OFF state, 1 hour and 2 hours after levodopa intake) better explains the data. The hypothesis-testing model including "effect" instead of "state" did not perform better (model comparison, p = 0.8, AIC difference: lower by 1.9 units for model 1), therefore we chose to proceed with "state".

6.4.2 Exploratory analysis

We performed our exploratory analysis as described in Section 5.4. No models were found that significantly better predicted the effect of levodopa on score.

6.4.3 Summary of best-fitting model

We built a linear mixed-effects model to test our null hypothesis that there is no difference between the ON and OFF state of PD patients. We first built a hypothesis-testing model, which assessed the effect of levodopa on score. The model – which had state as a fixed effect and random intercepts for participants – was not significant (b = 0.05, t(678) = 0.519, p = 0.6, d = 0.04). No better models were found in the exploratory analysis.

6.5 Case study: Twins with PD

One participant pair consisted of two PD patients, namely a pair of monozygotic (identical) twins. The participants in question were 57 years old at the time of testing. They grew up in the same household with an older brother (also diagnosed with PD) and followed the same educational path (same primary school, high school and university track). They both hold jobs as forest rangers; they live together and are not married. In sum, they are identical in their genes, upbringing and lifestyle. However, they do differ in one important respect, namely the duration and severity of their PD diagnosis.

One twin (participant code 01) was diagnosed with PD in 2012 but has shown few symptoms until recently. At the time of testing, he had started taking levodopa 10 days ago, did not experience any ON/OFF states nor has he felt any influence of the disease on speech. The second twin (participant code 02) was diagnosed with PD in 2011 and faces typical PD symptoms, predominantly tremor (also easily visible when in contact with him). He started taking levodopa 3 years ago, experiences ON/OFF states, and feels that the disease has influenced his speech. He reported tight muscles around his mouth in the OFF state, and softer muscles and faster speech in the ON state.

The present section is thus dedicated to examining the differences in speech between this pair of twins. In general, twins are more similar to each other than to the general population, however they do still show inter-speaker differences (Loakes, 2008). This also seems to be the case for our twins. They are similar in the formants of corner vowels (Table 13) and in the shape of their vowel space areas (see Figure 46 for individual vowel space areas).

Vowel →	/i/		/a/		/u/	
Participant ↓	f1	f2	f1	f2	f1	f2
01	307	1915	585	1184	346	962
02	286	1941	641	1190	361	1036

Table 13: Vowel formants for twins

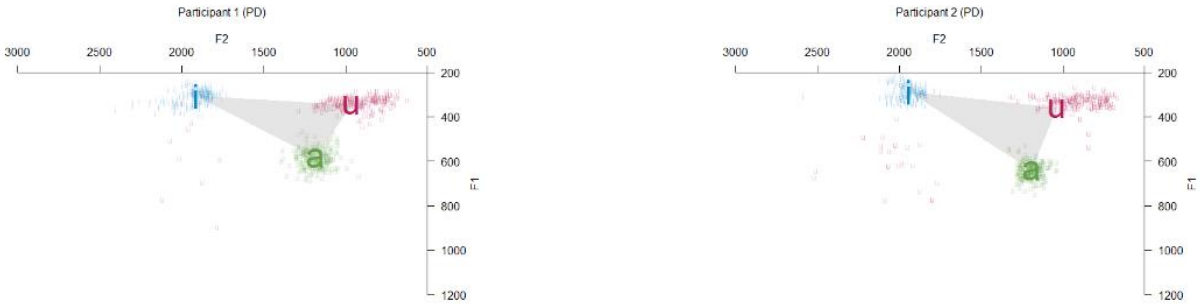
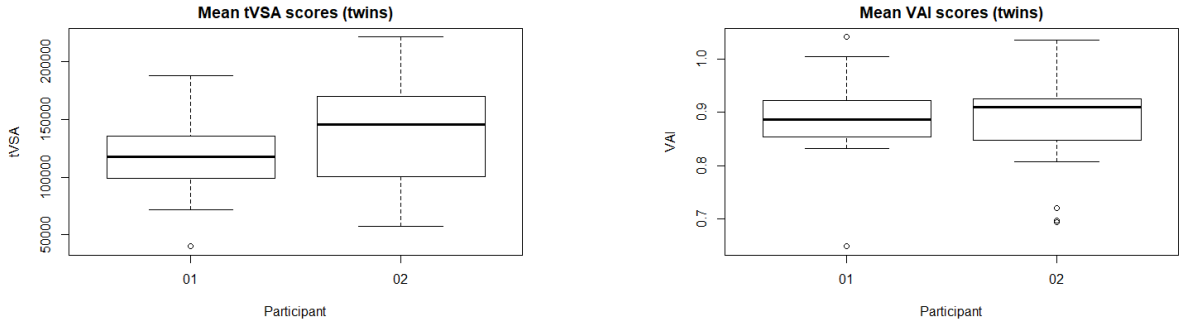


Figure 46: Differences in the vowel space areas of twins

However, they differ in some of our measured parameters (Table 14). This is especially noticeable with the vowel measures (Figure 47, upper left and upper right) and fundamental frequency (Figure 47, lower left), which are higher for the twin with more severe Parkinson's disease. As he had complained about muscle tension, this could also mean his laryngeal muscles are more rigid, leading to higher speech frequencies. Finally, the measure of pathological voice quality, CPPS, is lower for him (Figure 47, lower right).

ID	tVSA ± SD	VAI ± SD	*f0* ± SD	CPPS ± SD
1	121,584 ± 10,837	0.89 ± 0.020	142.8 ± 7.8	4.80 ± 0.38
2	139,494 ± 40,568	0.90 ± 0.048	146.7 ± 5.0	4.58 ± 0.27

Table 14: Mean values for the measured acoustic parameters



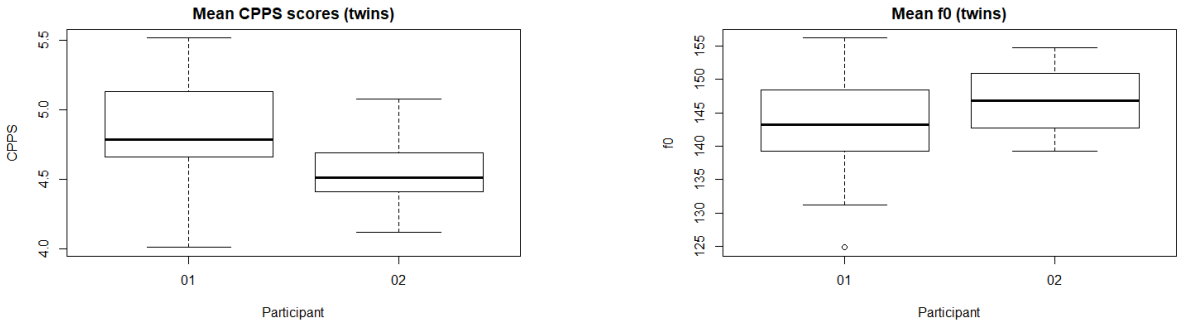


Figure 47: Comparison of individual measured parameters (twins)

Furthermore, they show different patterns across sessions (Figure 48).

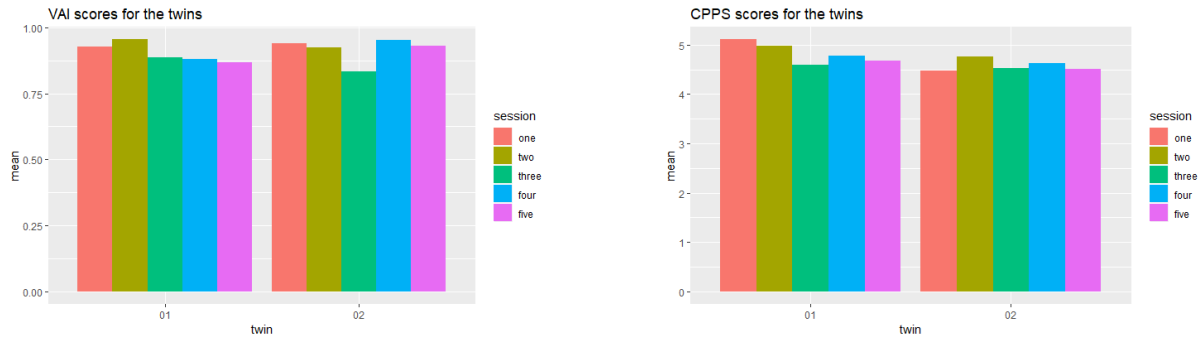


Figure 48: Vowel articulation and voice quality across sessions (twins)

They are seemingly similarly affected by levodopa in all measures excepting CPPS scores. More specifically, both twins have a lower vowel articulation index and a smaller vowel space area in the ON state (Figure 49, left), and higher fundamental frequency in the ON state (Figure 49, right).

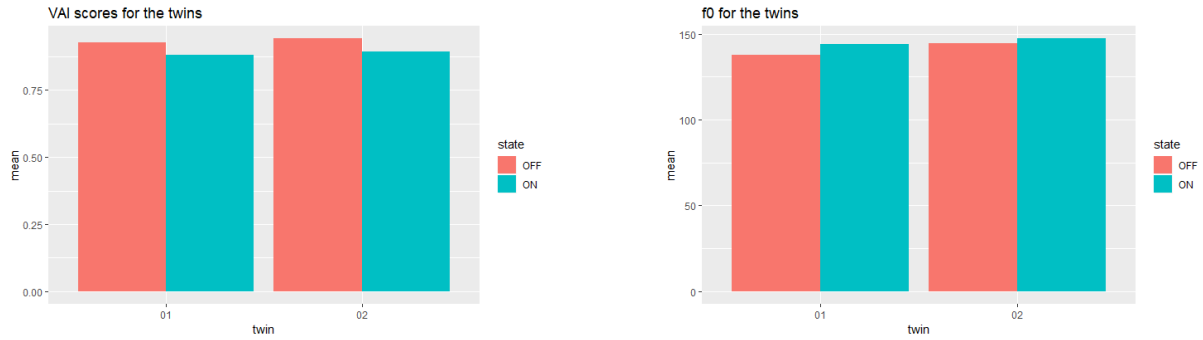


Figure 49: Vowel measure and voice quality measures in OFF and ON state (twins)

Keeping in mind that twins can have different speech characteristics, our case study nonetheless shows that twin research could be a potential good avenue for assessing the effect of Parkinson's disease and levodopa on speech.

7 Discussion and conclusions

The present study investigated the effect of levodopa on speech of Slovenian patients with Parkinson's disease. In Section 7.1 of this chapter, we discuss our hypotheses in light of the results we presented in Chapter 6. First, we discuss the hypothesis concerning the difference between Parkinson's disease patients and healthy control speakers (Section 7.1.1), followed by a discussion of the hypothesis concerning the effect of levodopa on PD speech (Section 7.1.2). In the next section (7.2) we discuss individual variability, including the case study on twins. We conclude the chapter by considering the limitations of our study and providing directions for further research on speech in Parkinson's disease (Section 7.3).

7.1 Discussion: Hypotheses

For our study, we analysed the speech of 6 Slovenian PD patients (5 male) and 4 healthy control speakers (1 male) who had recorded themselves on 4 different days, 5 times a day. In each of the 20 sessions, the participants filled out a fatigue survey and performed four speech tasks with their recording partner, including reading a wordlist, playing the kwartet card game and spot-the-differences game, and doing the diadochokinesis task. Following is a discussion on our two hypotheses, based on the analysis we have done on two of these tasks (wordlist and kwartet). Specifically, we formed our conclusions by looking at several acoustic parameters, including vowel articulation measures (tVSA and VAI) and voice quality measures (*f0* and CPPS).

7.1.1 Hypothesis 1: Comparing PD patients and healthy control speakers

Our first main research question concerned the difference between Parkinson's disease patients and healthy control speakers. Specifically, we set the null hypothesis that there will be no difference in the speech of PD patients and healthy control speakers, and the alternative hypothesis that compared to healthy controls, Slovenian PD patients will show significantly deteriorated speech, indicated in smaller values in all the measured acoustic parameters. The alternative hypothesis was not supported, as there was no significant effect of group. Instead, we found that sex was very significant, meaning that the gender of our participants explained the obtained data better than whether they suffered from Parkinson's disease.

When looking at trends in individual measures on group level, we could see that there is a trend towards reduced vowels' articulation in PD patients, mirrored in lower tVSA and VAI values.

There is additionally a trend towards lower *f0* and CPPS values in PD patients. However, only two participants, namely one healthy control speaker (male) and one PD patient (male) showed CPPS values that were lower than 4.0, which had previously been indicated as a cut-off point for non-pathological CPPS values (Heman-Ackah et al., 2014).

However, as the analysis showed, our data was better explained by the gender of the participants rather than Parkinson's disease. The trends seen for the two groups (PD and HC) were mirrored when looking at gender. Men showed reduced vowel articulation compared to women. Furthermore, men had a lower fundamental frequency and lower CPPS values. Finally, this difference was also visible in vowel space areas. When we plotted vowel space areas separated by gender (male/female) and group (PD/HC), we could see that there were more similarities in gender than in group. Furthermore, when we plotted only group, we saw clear bimodality in the data, as a cloud of data points belonging to the female PD speaker lay outside the bounds of the corner vowels for male PD speakers.

None of the other variables were significant when predicting our results. The first subquestion we included was whether fatigue has an effect on dysarthric and healthy speech. We showed that it did not have an effect. The study included a fatigue survey, marking tiredness levels on a scale from 0 to 10, that the participants had to fill out before each session. Most participants marked their fatigue levels between 0 and 2 ("not at all tired" to "a little tired") and including fatigue in our analysis did not improve the linear mixed-effects model. As most participants did not feel extremely tired at any point during the twenty sessions – only one participant, a healthy control, marked her tiredness level at 9 for two sessions on the same day – we cannot conclude that fatigue plays no role in speech. However, it does show that slight fatigue in patients cannot account for large differences in acoustic measures. Consequently, although Goberman, Coelho and Robb (2002) suggest that some PD speech studies, especially those on levodopa, were influenced by the fact that participants provided several samples on a single day and were therefore fatigued, our study shows that this is unlikely.

The second subquestion we included was whether task plays a role in vowel articulation. It did not have an effect. We included vowel formant data from two different tasks (wordlist and kwartet) to check whether task is an important predicting factor for our data. While values were slightly higher in the kwartet task, including task as a predictor did not significantly improve our models. Additionally, task did not differently affect PD patients and healthy controls. This is in opposition to results from previous studies, which found that choice of task plays a crucial

role (Goberman, Coelho and Robb, 2002; Ho et al., 1998), also when measuring vowel articulation (Rusz et al., 2013). There could be several reasons for this, for example because other studies compared a reading task and spontaneous speech whereas we compared a reading tasks and semi-spontaneous (scripted) speech; or because we obtained more data points for the wordlist task than for the kwartet task, and only distinguished between tasks for vowel articulation measures, not voice quality.

Finally, we also checked for the existence of the so-called vocal warm-up effect, which states that fundamental frequency is higher in the evening than in the morning. While we saw a trend towards higher values in the afternoon for all individuals, time of day was not a significant predictor for our model).

7.1.2 Hypothesis 2: Comparing OFF/ON state of PD patients

Our second main research question concerned the effect of levodopa on speech of PD patients. Specifically, we set the null hypothesis that there will be no difference in parkinsonian speech in the patients' OFF state compared to their ON state, and the alternative hypothesis that there will be a significant difference between the two states. The alternative hypothesis was not supported, as speech was not significantly affected by levodopa. Furthermore, none of the other variables (i.e. fatigue, task, time of day) affected the speech of PD patients across different states.

When looking at individual measures, no clear patterns emerged. Vowel articulation measures were higher in the OFF state than the ON state, however the difference was too small (and the standard deviation too high) to draw any conclusions about trends. Regarding voice quality measures, both the mean fundamental frequency and CPPS measures were lower in the OFF state than in the ON state. The difference, again, was too small to draw any conclusions. Our study thus joins the ranks of others, e.g. study by Fabbri et al. (2017) or Goberman, Coelho and Robb (2005), which did not find any significant differences in voice and speech parameters. Should the trends be bigger and appear in a bigger sample of (homogeneous) PD patients, we would presume that articulation, reflected in vowel measures, deteriorates when patients are on levodopa, while prosody and voice quality, reflected in fundamental frequency and CPPS, improve when patients are on levodopa. This would then also follow the results of Sanabria et al. (2001) who found that fundamental frequency improves in the ON state.

We additionally tested whether speech changes across the levodopa *cycle*, namely by looking at the effect of the drug in the OFF state as well as 1 and 2 hours after intake, as others have done before (e.g. De Letter et al., 2010; Goberman et al., 2002; Ho et al., 2008). While there was no significant effect of the drug cycle on speech, it was possible to see that there are differences between the first ON state (1 hour after intake) and the second ON state (2 hours after intake). Measuring speech on different occasions during the levodopa cycle is thus important.

Finally, there were individual changes and improvements in PD patients, as also seen in Goberman, Coelho and Robb (2005). Two patients, for example, showed improvements in all measured acoustic parameters 1 hour after levodopa intake (reflected in higher tVSA, VAI, *f0* and CPPS values). Both of these patients have been on the levodopa pill for approximately three years and reported speech problems as well as ON/OFF states. Other patients showed a worsening of vowel articulation one hour after levodopa intake, but individual improvements in *f0* and CPPS. Unfortunately, due to a very heterogeneous sample (both in terms of disease duration and the amount of time the patients have been taking levodopa) and due to high variation across days, these potential trends are not to be relied on.

7.2 Discussion: Individual variability

It is known that speakers with dysarthria show large individual variation in their speech (Metter and Hanson, 1986), however in our case, it became apparent that healthy speakers show non-negligible variability as well.

Based on the obtained vowel instances, we first calculated the mean formants per gender in order to compare it to previous studies, followed by the two vowel articulation measures (tVSA and VAI). Additionally, we extracted voice quality data from each participant. The average formant values of Slovenian corner vowels found in other studies (introduced in Chapter 2.3.3) versus our results are compared in Table 15 below. The results of the present study are marked in bold. We can see that values are comparable across the studies, with the exception of the second formant of the vowel /u/, which is higher than previously indicated. In line with previous research done by thesis author (Rebernik, 2018, unpublished), /u/ now seems more fronted than other studies' results show.

Vowel →	/i/		/a/		/u/	
Study ↓	f1	f2	f1	f2	f1	f2
Srebot-Rejec (1988)	382	2116	726	1332	393	747
Šuštaršič et al. (1996)	301	2250	735	1362	317	621
Tivadar (2004) – women	385	2318	774	1578	423	770
Tivadar (2004) – men	351	2219	603	1324	385	754
Jurgec (2005)	280	2309	717	1256	321	853
Present study – women	**312**	**2311**	**858**	**1512**	**352**	**944**
Present study – men	**297**	**1983**	**630**	**1245**	**337**	**988**

Table 15: Comparison of our vowels' formants and those obtained in previous studies

On an individual level, when plotting vowel space areas, the difference between female and male participants was immediately visible. More specifically, the vowel space areas of male participants 1-5 (PD patients) and male participant 7 (HC) are smaller than vowel space areas of participant 6 (PD patient) and 8-10 (HC). In individual vowels, participants differ most in corner vowel /a/, which can be either produced as a back vowel or more centralized. It is especially interesting to look at vowel spaces of the twins (participants 1 and 2) as they are more similar in shape and corner vowel placement compared to other participants.

When studying vowel articulation measures (both tVSA and VAI) for individuals we can see that the trends across sessions are similar for the two measures. However, whereas it is obvious that tVSA measures are higher for women than men, VAI seems to be able to better account for differences in gender (as already proposed by Skodda, Visser and Schlegel, 2010). It might therefore be a more suitable measure for studying vowel articulation in pathological speakers. Furthermore, our results for tVSA measures mirror the results of Skodda, Visser and Schlegel (2010), as the tVSA values for their participants were only smaller in male PD speakers compared to male healthy controls, while female PD speakers and healthy controls had similar values.

When studying voice quality measures (*f0* and CPPS), the trends across sessions differ depending on the measure we look at. The fundamental frequency is consistently lower in the morning sessions compared to the evening sessions, no matter the speaker (remember, however, that time of day wasn't a statistically significant predictor in our models). Additionally, although we included fundamental frequency, it is not a highly informative measure due to a small sample of PD patients. In retrospect, measuring variability in *f0* (as measured in Goberman, Coelho and Robb, for example) would have been more suitable. We chose against it because the software we used for measuring both CPPS and *f0* (Hillenbrand and Houde, 1996)

provided mean measurements and adding variability in *f0* would have demanded a significantly higher time investment.

On the other hand, CPPS values are seemingly higher in healthy female speakers compared to PD patients, regardless of the gender, although only two reach the threshold of lower than 4.0 necessary for qualifying pathological speech. In this, however, one participant stands out: the healthy male control speaker, who was older than other speakers, also had the lowest CPPS value. We do not know how age affects CPPS values. Furthermore, the higher values could also be due to language differences – the original article by Heman-Ackah et al. (2002) used English and the threshold of 4.0 is thus set on the basis of English, whereas we used Slovenian. Since CPPS is, at its basis, a measure of breathiness (or, rather, harmonic organization) it should not be language-dependent, but as of yet, no studies have been done on this, so we cannot be certain whether language plays a role.

Interestingly, CPPS is also the only measure in which the twins differed in the effect of levodopa. Specifically, CPPS is lower in the ON state for the twin with less severe PD, and higher in the ON state for the twin with more severe PD. Presuming that levodopa has a stronger effect on the twin with more severe PD (who is taking 450mg daily dose) compared to the twin who just started taking levodopa (75mg daily dose), it is possible that as a measure, only CPPS is affected by levodopa. Consequently, this could indicate that voice breathiness is improved by levodopa. CPPS, thus, can be characterized as a promising measure for determining the effect of levodopa.

Finally, we need to consider individual variability not just in the individual measures, but also across the four days and twenty sessions. If there are any conclusions that can be drawn from our study (except, of course, that gender plays a crucial role in acoustic research) they should concern variation in speech. In particular, speakers might show one pattern on one day (e.g. higher tVSA values in the morning) and a different pattern on another day (e.g. lower tVSA values in the morning). Speech is variable and, as our study shows, does not depend on only one factor, like fatigue, but rather several combined.

7.3 Limitations and further research

The present study has given us many insights, also as the first study on the acoustic properties of Slovenian parkinsonian speech. However, it still has considerable limitations that need to be acknowledged, predominantly related to our population sample. Due to recruitment and time constraints, the study had an uneven number of healthy control participants and PD patients. Further, due to the experimental design, where the recruited PD patients would record with their partners (namely, spouses), the healthy control participants were predominantly female while PD patients were predominantly male. The opposite was only true for one participant pair, with a female PD patient and her husband as a healthy control. However, both participants in that pair were markedly older (in their 70s) than the rest of our participants (in their 50s and 60s). Such a population sample also impacted our results, as we found it difficult to distinguish between the effect of gender and effect of disease (see Section 7.1.1 above).

Second, participant heterogeneity extended to our PD sample as well. While we recruited within our parameters (see the exclusion criteria in Chapter 4), the PD patients varied in terms of how long they have had Parkinson's disease and how long they have been on levodopa. With a study such as ours, it is difficult to draw any conclusions about the effect of levodopa, as the measured acoustic parameters did not significantly differ in the ON and OFF states, but our sample is too varied to find this a conclusive result.

Finally, there is a limitation in the measured acoustic parameters. Again, due to time constraints, we focused on analysing vowels' articulation (namely a measure of vowel space area and vowel centralization) and voice quality (namely a measure of breathiness and fundamental frequency). First, this meant we neglected other acoustic parameters, such as the voice onset time of plosives, which are also said to be affected in Parkinson's disease and which our study had also measured. Second, we did not study parameters not directly related to acoustics, such as articulatory rate (as measured by the DDK task) or rate of acceleration, both affected in PD. Finally, we chose to focus our efforts on two tasks, which contained read speech (i.e. the wordlist task) and semi-spontaneous speech (i.e. the kwartet task), and all our analysis stems from that. More studies, not just ours, suffer from this problem.

Future studies should continue to focus on the effect of levodopa on parkinsonian speech, as many unknowns still remain. On the one hand, they should make sure to choose a homogeneous sample in terms of disease duration and amount of time on levodopa. The healthy control

speakers should be matched to the PD patients in gender and age as closely as possible. On the other hand, it remains vital to choose different measures (from vowels and plosives to fundamental frequency, voice quality and rate), as it is likely that articulatory measures, such as vowel space, are affected differently by the disease than more prosodic measures, such as fundamental frequency, or voice quality measures, such as CPPS. Finally, our study has shown the necessity of recording the speech of PD patients on separate occasions, both within the same day and across different days. It is our hope that future studies will benefit from our insights and recommendations.

8 References

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences*, *12*(10), p. 366–375.

Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Audelman, J. U., Hoel, R. L., & Lassman, F. M. (1970). The effect of l-DOPA treatment on speech of patients with Parkinson's disease. *Neurology*, *20*(4), 410 – 411.

Azevedo, L. L., Cardoso, F., & Reis, C. (2003). Acoustic analysis of prosody in females with Parkinson's disease: effect of L-dopa. *Arq Neuropsiquiatr.*, *61*(4), 995–998.

Azevedo, L. L., Reis, C. A., Souza, I. S., & Cardoso, F. E. (2013). Prosody and levodopa in Parkinson's disease. *Arq Neuropsiquiatr*, *71*(11), 835–840.

Balasubramanium, R. K., Bhat, J. S., Fahim III, S., & Raju III., R. (2011). Cepstral Analysis of Voice in Unilateral Adductor Vocal Fold Palsy. *Journal of Voice*, *25*(3), 326–329.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Boersma, P., & Weenink, D. (2018). *Praat: doing phonetics by computer*. Retrieved from: http://www.praat.org/.

Bolmgren, M., Robb, M., & Chen, Y. (1998). A note on vowel centralization in stuttering and non-stuttering individuals. *Journal of Speech, Language, and Hearing Research*, *41*, p. 1942–1051.

G. J. Borden, K. S. Harris & L. J. Raphael (Eds.) (1994). *Speech science primer: physiology, acoustics, and perception of speech*. Baltimore, MD, United States: Williams and Wilkins.

Braak, H., Ghebremedhin, E., Rüb, U., Bratzke, H., Del Tredici, K. (2004). Stages in the development of Parkinson's disease-related pathology. *Cell and Tissue Research*, *318*(1), p. 121–134.

Brabenec, L., Mekyska, J., Galaz, Z., & Rektorova, I. (2017). Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *J Neural Trans*, *124*, p. 303–334.

Burk, B. R., & Watts, C. R. (2018). The effect of Parkinson disease tremor phenotype on cepstral peak prominence and transglottal airflow in vowels and speech. *Journal of Voice*, available online.

Campolounghi-Pegan, P. (2008). Parkinsonova bolezen (Združenje zdravnikov družinske medicine). Retrieved from http://www.drmed.org/wp-content/uploads/2014/06/VI-53.pdf

Chen, J. J. (2010). Parkinson's disease: health-related quality of life, economic cost, and implications of early treatment. *The American Journal of Managed Care*, *16*, S87–S93.

Connolly, B. S., & Lang, A. E. (2014). Pharmacological Treatment of Parkinson Disease: A review. *JAMA*, *311*(16), p. 1670–1683.

Critchley, E. (1981). Speech disorders of Parkinsonism: a review. *Journal of Neurology, Neurosurgery, and Psychiatry*, *44*, 751–758.

Daniels, N., Oates, J., Phyland, D., Feiglin, A., & Hughes, A. J. (1996). Vocal characteristics and response to levodopa in Parkinson's disease. *Movement Disorders*, *11*(Suppl. 1), 117.

Darley, Aronson and Brown (1969) Differential diagnostic patterns of dysarthria.

Lau, L. M. L. de, & Breteler, M. M. B. (2006). Epidemiology of Parkinson's disease. *Lancet Neurology*, *5*, p. 525–535.

Letter, M. de, Santens, P., & Borsel, J. van (2005). The effects of levodopa on word intelligibility in Parkinson's disease. *Journal of Communication Disorders*, 38, 187–196.

Letter, M. de, Santens, P., Bodt, M. de, Boon, P., & Borsel, J. van (2006). Levodopa-induced alterations in speech rate in advanced Parkinson's disease. *Acta neurologica Belgica*, *106*, 19–22.

Letter, M. de, Santens, P., Bodt, M. de, Maele, G. van, Borsel, J. van, & Boon, P. (2007a). The effect of levodopa on respiration and word intelligibility in people with advanced Parkinson's disease. *Clinical Neurology and Neurosurgery*, *109*, 495–500.

Letter, M. de, Santens, P., Estercam, I., Maele, G. van, Bodt, M. de, Boon, P., & Borsel, J. van (2007b). Levodopa-induced modifications of prosody and comprehensibility in advanced Parkinson's disease as perceived by professional listeners. *Clinical Linguistics and Phonetics*, *21*, 783–791.

Letter, M. de, Borsel, J. van, Boon, P., Bodt, M. de, Dhooge, I., & Santens, P. (2010). Sequential changes in motor speech across a levodopa cycle in advanced Parkinson's disease. *International Journal of Speech-Language Pathology, 12*(5), 405–413.

Letter, M. de, Santens, P., & Borsel, J. van (2005). The effects of levodopa on word intelligibility in Parkinson's disease. *Journal of Communication Disorders, 38*, 187–196.

Defazio, G., Guerrieri, M., Liuzzi, D., Gigante, A. D., & Di Nicola, V. (2016). Assessment of voice and speech symptoms in early Parkinson's disease by the Robertson dysarthria profile. *Neurological Sciences, 37*(3), 443–449.

Dimitrova, S. (1997). Bulgarian Speech Rhythm: Stress-Timed or Syllable-Timed. *Journal of the International Phonetic Association, 27*(1/2), 27–33.

Djaldetti, R., & Melamed, E. (1998). Management of response fluctuations: Practical guidelines. *Neurology, 51*(2), 36–40.

Dorsey, E. R., & Bloem, B. R. (2017). The Parkinson Pandemic – A call to action. *JAMA Neurology,* published online November 2017.

Dorsey, E. R., Constantinescu, R., …, Siderowf, A., & Tanner, C. M. (2007). Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology, 68*, p. 384–386.

Dorsey, E. R., Elbaz, A., …, Vos, T., & Murray, C. J. L. (2018). Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology, 17*(11), p. 939–953.

Doyle, P., Raade, A., St. Pierre, A., & Desai, S. (1995). Fundamental frequency and acoustic variability associated with production of sustained vowels by speakers with hypokinetic dysarthria. *Journal of Medical Speech–Language Pathology, 3*, 41–50.

Dykstra, A. D., Hakel, M. E., & Adams, S. G. (2007). Application of the ICF in reduced speech intelligibility in dysarthria. *Seminars in Speech and Language,* 28(4), p. 301–311.

Eichhorn, J. T., Kent, R. D., Austin, D., & Vorperian, H. K. (2018). Effects of aging on vocal fundamental frequency and vowel formants in men and women. *Journal of Voice, 32*(5), p. e1–e644.

EPDA – European Parkinson's Disease Association. (2019). About Parkinson's. Retrieved from www.epda.eu.com.

Fabbri, M., Guimaraes, I., Cardoso, R., …, Ferreira, J. J. (2017). Speech and Voice Response to a Levodopa Challenge in Late-Stage Parkinson's Disease. *Frontiers in Neurology, 22*(8): 432.

S. A. Factor & W. J. Weiner (Eds.) (2008). *Parkinson's disease: Diagnosis & Clinical Management, Second Edition*. New York, United States: Demos Medical Publishing.

Farrus, M., Hernando, J., & Ejarque, P. (2008). Jitter and shimmer measurements for speaker recognition. Retrieved from: http://www.cs.upc.edu.

Feigin, V. L., Abajobir, A. A., …, Murray, C. J. L., & Vos, T. (2017). Global, regional, and national burden of neurological disorders during 1990 – 2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Neurology, 16*(11), p. 877–897.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *J. psychiat. Res., 12*, p. 189–198.

Fraille, R., & and Godino-Llorente, J. I. (2014). Cepstral peak prominence: a comprehensive analysis. Retrieved from oa.upm.es.

Frota, S., Oliveira, P., Cruz, M., Vicente, S., Cardoso, R., Guimaraes, I., Ferreira, J., Pinto, S., & Vigario, M. (2018). (Dys)prosody in Parkinson's disease: effects of medication and disease duration on different prosodic functions. *European Journal of Neurology, 25*, 422–423.

Garrett, K. L., & Healey, C. E. (1987). An acoustic analysis of fluctuations in the voices of normal adult speakers across three times of day. *The Journal of the Acoustical Society of America, 82*(1), 58–62.

Granda, G., Mlakar, J., & Vodušek, D. B. (2003). Kratek preizkus spoznavnih sposobnosti – usmerjanje pri preiskovancih, startih od 55 do 75 let (I). *Zdravniški vestnik, 72*(10), 575–581.

Gentil, M., Tournier, P., & Pollak, A. L. (1999). Effect of bilateral subthalamic nucleus stimulation and dopatherapy on oral control in Parkinson's disease. *European Neurology, 42*, 136–140.

Goberman, A. M., & Coelho, C. (2002). Acoustic analysis of Parkinsonian speech II: L-Dopa related fluctuations and methodological issues. *NeuroRehabilitation, 17*, 247–254.

Goberman, A., Coelho, C., & Robb, M. (2002). Phonatory characteristics of Parkinsonian speech before and after morning medication: the ON and OFF states. *Journal of Communication Disorders, 35*, 217–239.

Goberman, A. M., Coelho, C. A., & Robb, M. P. (2005). Prosodic Characteristics of Parkinsonian Speech: The Effect of Levodopa-Based Medication. *Journal of Medical Speech-Language Pathology, 13*(1), 51–68.

Gorham-Rowan, M. M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders, 39*(3), 171–184.

Greenberg, M. L. (2003). Word prosody in Slovene from a typological perspective. *STUF - Sprachtypologie und Universalienforschung, 56*(3), p. 234–251.

Harel, B. T., Cannizzaro, M. S., Cohen, H., Reilly, N., & Snyder, P. J. (2004). Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment. *Journal of Neurolinguistics, 17*, 439–453.

Hartelius, L., & Svensson, P. (1994). Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: a survery. *Folia Phoniatrica et Logopaedica, 46*(1), 9–17.

Heman-Ackah, Y. D., Sataloff, R. T., …, Johnson, J., & Hillenbrand, J. (2014). Quantifying the cepstral peak prominence, a measure of dysphonia. *Journal of Voice, 28*(6), p. 783–788.

Heman-Ackah, Y. D., Michael, D. D., and Goding, G. S. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice, 16*, p. 20–27.

Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Baroody, M. M., Hillenbrand, J., & Sataloff, R. T. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Ann. Otol. Rhinol. Laryngol. 112*, p. 324–333.

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *Journal of Speech and Hearing Research, 39*, p. 311–321.

Ho, A. K., Bradshaw, J. L., & Iansek, R. (2008). For better or worse: the effect of levodopa on speech in Parkinson's disease. *Movement Disorders, 23*(4), p. 574–580.

Ho, A. K., Iansek, R., Marigliani, C., Bradshaw, J. L., & Gates, S. (1998). Speech impairment in a large sample of patients with Parkinson's disease. *Behavioural Neurology, 11*(3), p. 131–137.

Holmes, R. J., Oates, J. M., Phyland, D. J., & Hughes, A. J. (2000). Voice characteristics in the progression of Parkinson's disease. *International Journal of Language and Communication Disorders, 35*(3), 407–418.

Jankovic, J. (2007). Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry, 79*, p. 368–376.

Jian, J., Lin, E., Wang, J., & Hanson, D. G. (1999). Glottographic measures before and after levodopa treatment in Parkinson's disease. *The Laryngoscope, 109*, 1287–1294.

Josipović, V. (1994). English and Croatian in the typology of rhythmic systems. *Studia romanica et anglica zagrabiensia (SRAZ), 39*, 25–37.

Jurgec, P. (2004). Fonologija samoglasniških nizov v slovenščini. *Slavistična revija, 52*(2), 119–140.

Jurgec, P. (2005). Formant frequencies of standard Slovene vowels. *Govor, 22*(2), 127–144.

Jurgec, P. (2011). Slovenian has nine vowels. *Slavistična revija, 59*(3), 243–268.

Kempster, P. A., Gibb, W. R., Stern, G. M., & Lees, A. J. (1989). Asymmetry of substantia nigra neuronal loss in Parkinson's disease and its relevance to the mechanism of levodopa related motor fluctuations. *Journal of Neurology, Neurosurgery & Psychiatry, 52*(1), 72–76.

Khodadadian, A., Hemmati-Finarvand, M., Kalantary-Charvadeh, A., Ghobadi, A., & Mazaheri, M. (2018). Candidate biomarkers for Parkinson's disease. *Biomedicine & Pharmacotherapy, 104*, 699–704.

Klingelhoefer, L., & Reichmann, H. (2017). Parkinson's disease as a multisystem disorder. *Journal of Neural Transmission, 124*(6), 709–713.

Komar, S. (2007). Strategies to improve reading performance in English. In *Proceedings of the Phonetics Teaching and Learning Conference (PTLC), 24-26 August 2007.* University College London.

Larson, K., Ramig, L., & Scherer, R. (1994). Acoustic and glottographic voice analysis during drug-related fluctuations in Parkinson's disease. *Journal of Medical Speech-language Pathology, 2*, 227–239.

Letter, M. de, Santens, P., Bodt, M. de, Maele, G. van, Borsel, J. van, & Boon, P. (2007a). The effect of levodopa on respiration and word intelligibility in people with advanced Parkinson's disease. *Clin Neurol Neurosurg., 109*(6), 495–500.

Letter, M. de, Santens, P., Estercam, I., Maele, G. van, Bodt, M. de, Boon, P., & Borsel, J. van. (2007b). Levodopa-induced modifications of prosody and comprehensibility in advanced Parkinson's disease as perceived by professional listeners. *Clinical Linguistics and Phonetics, 21*(10), 783–791.

Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders,* 43(1), 47–57.

Loakes, D. (2008). A forensic phonetic investigation into the speech patterns of identical and non-identical twins. *The International Journal of Speech, Language and the Law, 15*(1), 97–100.

Louis, E. D., Winfield, L., Fahn, S., & Ford, B. (2001). Speech dysfluency exacerbated by levodopa in Parkinson's disease. *Movement disorders, 16,* 562–565.

Maryn, Y., Corthals, P., Cauwenberge, P. van, Roy, N., & Bodt, M. de (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *Journal of Voice, 24*(5), 540–555.

Maetzler, W., Liepelt, I., & Berg, D. (2009). Progression of Parkinson's disease in the clinical phase: potential markers. *The Lancet Neurology,* 8(12), 1158–1171.

Maillet, A., Krainik, A., Debû, Troprès, I., Lagrange, C., Thobois, S., Pollak, P., & Pinto, S. (2012). Levodopa Effects on Hand and Speech Movements in Patients with Parkinson's Disease: A fMRI Study. *PLOS One, 7*(10), e46541.

Marsden, C. D., Parkes, J. D., & Quinn, N. (1981). Fluctuations in disability in Parkinson's disease – clinical aspects. In C. D. Marsden & S. Fahn (Eds.): *Movement Disorders,* pp. 96–122. London, United Kingdom: Buttersworths.

Martinez-Martin, P., Gil-Nagel, A., Gracia, L. M., Gómez, J. B., Martinez-Sarriés, J., Bermejo, F. (1994). Unified Parkinson's disease rating scale characteristics and structure. *Movement Disorders, 9*(1), 76–83.

MATLAB. 2018. *Version 9.4.0 (r2018a)*. Natick, Massachusetts: The MathWorks Inc.

Mawdsley, C., & Gamsu, C. V. (1971). Periodicity of speech in Parkinsonism. *Nature, 231,* 315–316.

Metter, E. J., & Hanson, W. R. (1986). *Journal of Communication Disorders, 19*(5), 247–366.

Michael J. Fox Foundation. (2019). *Our promise.* Retrieved from https://www.michaeljfox.org

Miller, I. N., & Cronin-Golomb, A. (2010). Gender differences in Parkinson's disease: clinical characteristics and cognition. *Movement disorders, 25*(16), 2695–2703.

Miller, N., & Lowit, A. (2011). *Motor Speech Disorders: A Cross-Language Perspective.* Bristol, United Kingdom: Multilingual Matters.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. (2003). The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. *Movement Disorders, 18*(7), 738–750.

Müller, J., Wenning, G. K., Jellinger, K., McKee, A., Poewe, W., & Litvan, I. (2000). Progression of Hoehn and Yahr stages in Parkinsonian disorders: a clinicopathologic study. *Neurology, 55*(6), 888–891.

Nakano, K. K., Zubick, H., & Tyler, H. R. (1973). Speech defects of parkinsonian patients: effects of levodopa therapy on speech intelligibility. *Neurology, 23,* 865–870.

Ogrin, M. (2000). Logopedska obravnava bolnikov s Parkinsonovo boleznijo. *Medicinski razgledi, 39*(3), 313–317.

Okada, Y., Murata, M., & Toda, T. (2015). Effects of levodopa on vowel articulation in patients with Parkinson's disease. *Kobe J Med Sci, 61*(5), e144–154.

Olanow, C. W., Gauger, L. L., & Cederbaum, J. M. (1991). Temporal relationships between plasma and cerebrospinal fluid pharmacokinetics of levodopa and clinical effect in Parkinson's disease. *Annals of Neurology, 29*(5), 556–559.

Parkinson, J. (2002). An essay on the shaking palsy. 1817. *The Journal of Neuropsychiatry and Clinical Neurosciences, 14*(2), 223–236.

Parkinson's Foundation. (2019). *What is Parkinson's?* Retrieved from https://parkinson.org

Pavšič, K., & Pirtošek, Z. (2015). Nemotorični simptomi Parkinsonove bolezni. *Medicinski razgledi, 54*(1), 95–115.

R. F. Pfeiffer & I. Bodis-Wollner (Eds.) (2005). *Parkinson's Disease and Nonmotor Dysfunction*. Totova, NJ, United States: Humana Press, Inc.

Pinto, S., Gentil, M., Krack, P., Sauleau, P., Fraix, V., Benabid, A.-L., & Pollak, P. (2005). Changes Induced by Levodopa and Subthalamic Nucleus Stimulation on Parkinsonian Speech. *Movement Disorders, 20*(11), 1507–1515.

Pinto, S., Ozsanak, C., Tripoliti, E., Thobois, S., Limousin-Dowsey, P., & Azou, P. (2004). Treatments for dysarthria in Parkinson's disease. *The Lancet Neurology, 3*, p. 547–556.

Pinto, S., Chan, A., Guimaraes, I., Rothe-Neves, R., & Sadat, J. (2017). *Journal of Phonetics, 64*, p. 156–167.

Pinto, S., Gentil, M., Krack, P., Sauleau, P., Fraix, V., Benabid, A., & Pollak, P. (2005). Changes induced by levodopa and subthalamic nucleus stimulation on Parkinsonian speech. *Movement disorders, 20*(11), 1507–1515.

Plowman-Prine, E. K., Okun, M. S., Sapienza, C. M., Shrivastav, R., Fernandez, H. H., Foote, K. D. Ellis, C., Rodriguez, A. D., Burkhead, L. M., & Rosenbek, J. C. (2009). Perceptual characteristics of Parkinsonian speech: A comparison of the pharmacological effects of levodopa across speech and non-speech motor systems. *NeuroRehabilitation, 24*(2), 131–144.

Poluha, P. C., Teulings, H. L., & Brookshire, R. H. (1998). Handwriting and speech changes across the levodopa cycle in Parkinson's disease. *Acta Psychologica, 100*, 71–84.

Pringsheim, T., Jette, N., Frolkis, A., & Steeves, T. D. L. (2014). The prevalence of Parkinson's disease: a systematic review and meta-analysis. *Movement disorders, 29*(13), p. 1583–1590.

Quaglieri, C. E., & Celesia, G. G. (1977). Effect of thalamotomy and levodopa therapy on the speech of Parkinson patients. *European Neurology, 15*, 34–39.

R Core Team. (2013). *R: A language and environment for statistical computing*. Retrieved from: http://www.R-project.org/.

Ramig, L. O., Fox, C., & Sapir, S. (2004). Parkinson's disease: Speech and voice disorders and their treatment with the Lee Silverman Voice Treatment. *Seminars in Speech and Language, 25*(2), 169–180.

Roy, N., Nissen, S. L., Dromey, C., & Sapir, S. (2009). Articulatory changes in muscle tension dysphonia: evidence for vowel expansion following manual circumlaryngeal therapy. *Journal of Communication Disorders, 42*, p. 124 – 135.

Rusz, J., Cmejla, R., Tykalova, T., Ruzickova, H., Klemir, J., Majerova, V., Picmausova, J., Roth, J., & Ruzicka, E. (2013). Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task. *J. Acoust. Soc. Am.,* 134(3), p. 2171–2181.

Sanabria, J., Ruiz, P. G., Gutierrez, R., Marquez, F., Escobar, P., Gentil, M., & Cenjor, C. (2001). The Effect of Levodopa on Vocal Function in Parkinson's Disease. *Clinical Neuropharmacology, 24*(2), 99-102.

Santens, P., Letter, M. de, Borsel, J. van, Reuck, J. de, & Camaert, J. (2003). Lateralized effects of subthalamic nucleus stimulation on different aspects of speech in Parkinson's disease. *Brain and Language, 87*, p. 253–258.

Schapira, A. H. V. (2009). Neurobilogy and treatment of Parkinson's disease. *Trends in pharmacological sciences, 30*(1), 41–47.

Schulz, G. M., Greer, M., & Friedman, W. (2000). Changes in vocal intensity in Parkinson's disease following pallidotomy surgery. *Journal of Voice, 14*(4), p. 589–606.

Schulz, G. M., & Grant, M. K. (2000). Effects of speech therapy and pharmacologic and surgical treatments on voice and speech in Parkinson's disease: a review of the literature. *Journal of Communication Disorders, 33*, 59–88.

Širca, P. (2012). *Govor oseb s Parkinsonovo boleznijo* (Bachelor's thesis, University of Ljubljana, Slovenia). Retrieved from http://pefprints.pef.uni-lj.si/1259/1/Govor_oseb_s_Parkinsonovo_boleznijo_(diplomsko_delo)_Patricija_%C5%A0irca.pdf

Skodda, S., Visser, W., & Schlegel, U. (2010). Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease. *Journal of Neural Transmission, 117*(2), 197–205.

Skodda, S., Grönheit, W., Mancinelli, N., & Schlegel, U. (2013). Progression of voice and speech impairment in the course of Parkinson's disease: a longitudinal study. *Parkinson's Disease, 2013*, ID 389195.

Skodda, S., Grönheit, W., & Schlegel, U. (2011). Intonation and Speech Rate in Parkinson's Disease: General and Dynamic Aspects and Responsiveness to Levodopa Admission. *Journal of Voice, 25*(4), e199–e205.

Skodda, S., Visser, W., & Schlegel, U. (2011). Vowel articulation in Parkinson's disease. *Journal of Voice, 25*(4), 467–472.

Škofic, J., Horvat, M., & Kenda-Jež, K. (Eds.). (2016). *Slovenski lingvistični atlas 2.* Ljubljana, Slovenia: Založba ZRC, ZRC SAZU.

Solomon, N. P., & Hixon, T. J. (1993). Speech breathing in Parkinson's disease. *Journal of Speech and Hearing Research, 36*, 294–310.

Srebot Rejec, T. (1988). Word accent and vowel duration in Standard Slovene: An acoustic and linguistic investigation. *Slavistische Beiträge, 226*. Munich, Germany: Verlag Otto Sagner.

Srebot Rejec, T. (1992). Initial and final sonorant clusters in Slovene. *Linguistica (Ljubljana), 32*(2), 227–230.

Srebot Rejec, T. (1998). On the vowel system in present-day standard Slovene. In *Jezikovne tehnologije za slovenski jezik: zbornik konference* (p. 22–24)

Šuštaršič, R., Komar, S., & Petek, B. (1996). An acoustic analysis of contemporary vowels of the Standard Slovenian language. *4th International Conference on Spoken Language Processing (ICSLP 96)*, Delaware, 133 – 136.

Šuštaršič, R., & Tivadar, H. (2005). Perception of tonemicity in Standard Slovene. *Govor, 22*(1), 23–35.

Swigert, N. B. (1997). *The source for dysarthria*. East Moline, Illinois: LinguiSystems.

Tivadar, H. (2004). Fonetično-fonološke lastnosti samoglasnikov v sodobnem knjižnem jeziku. *Slavistična revija, 52*(1), 31–48.

Tjaden, K., & Wilding, G. E. (2004). Rate and loudness manipulations in dysarthria: acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research, 47*, p. 766–783.

Toporišič, J. (2006). *Besedjeslovne razprave.* Ljubljana, Slovenia: Založba ZRC, ZRC SAZU.

Trepetlika. (2019). *Kaj je parkinsonova bolezen?* Retrieved from http://www.trepetlika.si

Tropea, T. F., & Chen-Plotkin, A. S. (2018). Unlocking the mystery of biomarkers: a brief introduction, challenges and opportunities in Parkinson disease. *Parkinsonism and Related Disorders, 46*, S15–S18.

Trošt, M. (2008). Parkinsonova bolezen. *Farmacevtski vestnik, 59*, 60–63.

Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 113*(5), 609–623.

Tysnes, O.-B., & Storstein, A. (2017). Epidemiology of Parkinson's disease. *Journal of Neural Transmission, 124*, 901–905.

Walsh, B., & Smith, A. (2012). Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. *Movement Disorders, 27*(7), 843–850.

Watts, C. R., Awan, S. N., & Maryn, Y. (2017). A comparison of Cepstral Peak Prominence measures from two acoustic analysis programs. *Journal of Voice, 31*(3), 387.e1-e10.

Watson, P. J., & Munson, B. (2008). Parkinson's disease and the effect of lexical factors on vowel articulation. *The Journal of the Acoustical Society of America, 124*, EL291.

Wieling, M. B. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics, 70*, p. 86–116.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv:1308.5499. [http://arxiv.org/pdf/1308.5499.pdf]

Wolfe, V. I., Garvin, J. S., Bacon, M., & Waldrop, W. (1975). Speech changes in Parkinson's disease during treatment with L-dopa. *Journal of communication disorders, 8*, 271–279.

Woznicki, K. (2006). *An acoustic analysis of word prosody in Ljubljana Slovene*. The Ohio State University.

Zwirner, P., & Barnes, G. (1992). Vocal tract steadiness: A measure of phonatory stability and upper airway motor control during phonation in dysarthria. *Journal of Speech and Hearing Research, 35*, 761–768.

Appendix 1: Recruitment Letter

university of groningen

Teja Rebernik
Center za jezik in kognicijo
Univerza v Groningenu
E-pošta: t.rebernik@rug.nl

Vpliv levodope in utrujenosti na govor bolnikov s parkinsonovo boleznijo

OPIS RAZISKAVE

Kratek opis in pomen raziskave

Eden od simptomov parkinsonove bolezni je tudi slabšanje govora oziroma hipokinetična dizartrija, za katero so značilni momljanje, nenatančna izgovorjava soglasnikov, ovirana izgovorjava samoglasnikov ter tih in hripav govor. Kljub razširjenosti hipokinetične dizartrije pa je zaenkrat nejasno, kako nanjo vplivata levodopa in utrujenost, kar bomo proučevali v tej raziskavi.

Raziskava je pomembna iz dveh razlogov. Prvič, raziskave govora bolnikov s parkinsonovo boleznijo so se v zadnjem desetletju razširile, vendar pa raziskovalci velikokrat govor merijo pri posameznem bolniku le enkrat, zato ne vedo, ali ga merijo v najprimernejšem trenutku. Drugič, večina raziskav poteka na maternih govorcih angleščine, nemščine in nizozemščine, manj pa na drugih jezikih, ki imajo izrazito drugačne značilnosti.

Z raziskavo govora slovenskih bolnikov s parkinsonovo boleznijo bomo lahko argumentirano prispevali k razpravi o tem, katere značilnosti veljajo le za nekatere jezike, katere pa so univerzalne. Pričakujemo pa tudi, da bomo lahko izpostavili, da proučevanje govora slovenskih bolnikov prinaša uporabne rezultate in prispeva k boljšemu razumevanju vpliva levodope in utrujenosti na govor.

Vzorec bolnikov

Vzorec raziskave bo zajel 5 do 10 bolnikov. Sodelujejo lahko le bolniki s parkinsonovo boleznijo, ki:

- nimajo možganskih poškodb ali v preteklosti niso doživeli možganske kapi;
- nimajo depresije oziroma imajo blažjo obliko depresije, za katero ne potrebujejo zdravljenja;
- levodopo jemljejo v obliki tablet;
- niso prestali operacije za globoko stimulacijo možganov;
- nimajo drugih govornih motenj (npr. jecljanje);
- jim je slovenščina prvi, materni, jezik.

Potek raziskave

Sodelujoči bolniki bodo s pomočjo snemalne opreme (tj. pametnega telefona in naglavnih mikrofonov) snemali svoj govor, in sicer štiri dni v dveh tednih. Na dan se posnamejo petkrat; natančni časi so odvisni od urnika vnosa levodope. Raziskava zahteva od bolnika skupaj 5 ur

časa. Raziskovalka pred prvim snemanjem na dom bolnika prinese opremo in natančno razloži celoten potek raziskave.

Vsako posamično snemanje traja približno 15 minut, sodelujoči bolniki pa ga opravljajo s partnerjem. Oba najprej prebereta spisek stavkov, nato pa igrata dve igri: v eni iščeta razlike med dvema slikama, v drugi pa igrata igro s kartami, katere cilj je, da igralec zbere 4 karte iste vrste (npr. vse karte, ki imajo slike psov različnih pasem).

Varstvo in obdelava podatkov
Po pravilih Univerze v Groningenu mora vsak raziskovalec pred začetkom raziskave pridobiti etično dovoljenje od univerzitetne raziskovalne komisije (*CETO*), ki obenem tudi potrdi, da raziskava ne potrebuje širšega zdravstvenega etičnega dovoljenja. Anonimnost podatkov je zagotovljena v skladu z novo uredbo GDPR, ki stopa v uporabo 25. 5. 2018. Sodelujoči v raziskavi lahko s snemanjem govora kadarkoli prenehajo in jim pri tem ni treba navesti nobenega razloga.

Rezultati
Rezultati bodo objavljeni v magistrski nalogi, ki bo napisana za namene programa *Jezik in kognicija* na Univerzi v Groningenu (Nizozemska), ter v znanstvenem članku, izpeljanem iz magistrske naloge. Izsledki bodo prav tako objavljeni na spletni strani Društva Trepetlika, po dogovoru in želji pa se lahko za Društvo na isto temo pripravi tudi kratko predavanje.

O raziskovalki
Teja Rebernik je magistrska študentka in raziskovalna asistentka na Univerzi v Groningenu na Nizozemskem. Dela na presečišču jezikoslovja in kognitivnih znanosti, ter med drugim sodeluje v projektu, ki se ukvarja z govorom nizozemskih bolnikov s parkinsonovo boleznijo. Za seboj ima zaključen dodiplomski študij na Univerzi v Ljubljani ter podiplomski študij na Univerzi v Groningenu. V svojem prostem času veliko bere in igra violino, ko je v Sloveniji pa jo najdete v Mariboru, kjer rada zaide v gozdove Pohorja.

Teja Rebernik
Center for Language and Cognition
University of Groningen
E-mail: t.rebernik@rug.nl

Effect of levodopa and fatigue on the speech of patients with Parkinson's Disease

RESEARCH STUDY DESCRIPTION

Short description and importance of research study

One of the symptoms of Parkinson's disease is also disordered speech, i.e. hypokinetic dysarthria, the symptoms of which include mumbling, imprecise pronunciation of consonants and vowels, and quiet and hoarse speech. However, despite the pervasiveness of hypokinetic dysarthria, it is currently unclear how it is affected by levodopa and fatigue, which I will study in this research study.

The research study is important for two reasons. Firstly, studies on speech of Parkinson patients have been frequent over the past two decades, but researchers often measure speech only once for each patient, so they do not know whether they are measuring it at the most suitable moment. Secondly, most studies are done with native speakers of English, German and Dutch, but less so on other languages, which have significantly different characteristics.

By studying the speech of Slovenian patients with Parkinson's Disease, I will be able to contribute to the discussion on which characteristics appear only in certain languages, and which are universal. I also expect to be able to highlight how studying the speech of Slovenian patients brings useful results and contributes towards a better understanding of the effect of levodopa and fatigue on speech.

Participant sample

The sample will consist of between 5 to 10 patients. Eligible participants with Parkinson's Disease are those who:

- don't have any brain injuries and haven't had any strokes in the past;
- aren't diagnosed with depression or have a milder form of depression for which they do not need medication;
- take levodopa in the form of pills;
- didn't undergo deep brain stimulation;
- don't have any other speech disorders (e.g. stuttering);
- have Slovenian as their first, native, language.

Course of research study

The participants will use recording equipment (i.e. a smartphone and headset microphones) to record their speech, namely four days across two weeks. They record themselves five times a day; the exact times depend on the schedule of levodopa intake. The study demands

approximately five hours of the patients' time altogether. Before the first recording session, the researcher brings equipment to the participant's home and explains the proceedings of the study.

Each individual recording session takes approximately 15 minutes, and the participants do it with a partner. Both first read a list of sentences, then play two games: in one game, they look for differences between two pictures, in the other, they play a card game whose goal is for the player to collect 4 cards in the same category (e.g. all cards that have pictures of dogs of different breeds).

Data safety and analysis

Permission to conduct the study will be obtained from the research ethics review committee (*CETO*) at the University of Groningen. This also confirms that the research study does not need a broader medical ethical permission. Data anonymity is ensured in accordance with the new GDPR regulations, which came into effect on 25. 5. 2018. Participants can stop recording their speech at any time and drop out of the study. They do not need to state any reason for doing so.

Results

The results will be published in a Master's thesis, written for purposes of the programme *Language and Cognition*, an in a scientific article, derived from the Master's thesis. The results will also be published on the website from Association Trepetlika. If so desired, a short lecture on the topic can also be prepared for the Association.

About the researcher

Teja Rebernik is a Master student and research assistant at the University of Groningen, Netherlands. She works at the crossroads of linguistics and cognitive sciences, and is among other things also collaborating on a project on the speech of Dutch patients with Parkinson's Disease. She has finished her Bachelor's at the University of Ljubljana and a previous Master's at the University of Groningen. In her free time, she reads a lot and plays the violin. When she's in Slovenia, you can find her in Maribor, where she loves to hike across the forests of Pohorje.

Appendix 2: CETO approval

university of groningen / faculty of arts / commissie ethische toetsing onderzoek (ceto) research ethics committee

Prof. dr. John C.J. Hoeks
ceto@rug.nl

To Whom it May Concern

Date
19 July 2018

Dear Sir/Madam,

The Research Ethics Committee (CETO) of the Faculty of Arts, University of Groningen has reviewed the proposal *'Speech Deterioration in Parkinson's Disease (Slovenian)' (by reference to 53197881)* submitted by *Teja Rebernik*. The CETO has established that the research protocol follows internationally recognized standards to protect the research participants. We therefore have no objection against this proposal.

Respectfully yours,

Prof. Dr. J.C.J. Hoeks

1 › 1

Appendix 3: Information letter

university of groningen

Teja Rebernik
Center za jezik in kognicijo
Univerza v Groningenu
E-pošta: t.rebernik@rug.nl

Vpliv levodope in utrujenosti na govor bolnikov s parkinsonovo boleznijo

OPIS RAZISKAVE

Kratek opis in pomen raziskave

Eden od simptomov parkinsonove bolezni je tudi slabšanje govora oziroma hipokinetična dizartrija, za katero so značilni momljanje, nenatančna izgovorjava soglasnikov, ovirana izgovorjava samoglasnikov ter tih in hripav govor. Kljub razširjenosti hipokinetične dizartrije pa je zaenkrat nejasno, kako nanjo vplivata levodopa in utrujenost, kar bomo proučevali v tej raziskavi.

Raziskava je pomembna iz dveh razlogov. Prvič, raziskave govora bolnikov s parkinsonovo boleznijo so se v zadnjem desetletju razširile, vendar pa raziskovalci velikokrat govor merijo pri posameznem bolniku le enkrat, zato ne vedo, ali ga merijo v najprimernejšem trenutku. Drugič, večina raziskav poteka na maternih govorcih angleščine, nemščine in nizozemščine, manj pa na drugih jezikih, ki imajo izrazito drugačne značilnosti.

Z raziskavo govora slovenskih bolnikov s parkinsonovo boleznijo bomo lahko argumentirano prispevali k razpravi o tem, katere značilnosti veljajo le za nekatere jezike, katere pa so univerzalne. Pričakujemo pa tudi, da bomo lahko izpostavili, da proučevanje govora slovenskih bolnikov prinaša uporabne rezultate in prispeva k boljšemu razumevanju vpliva levodope in utrujenosti na govor.

Vzorec bolnikov

Vzorec raziskave bo zajel 5 do 10 bolnikov. Sodelujejo lahko le bolniki s parkinsonovo boleznijo, ki:

- nimajo možganskih poškodb ali v preteklosti niso doživeli možganske kapi;
- nimajo depresije oziroma imajo blažjo obliko depresije, za katero ne potrebujejo zdravljenja;
- levodopo jemljejo v obliki tablet;
- niso prestali operacije za globoko stimulacijo možganov;
- nimajo drugih govornih motenj (npr. jecljanje);
- jim je slovenščina prvi, materni, jezik.

Potek raziskave

Sodelujoči bolniki bodo s pomočjo snemalne opreme (tj. pametnega telefona in naglavnih mikrofonov) snemali svoj govor, in sicer štiri dni v dveh tednih. Na dan se posnamejo petkrat;

natančni časi so odvisni od urnika vnosa levodope. Raziskava zahteva od bolnika skupaj 5 ur časa. Raziskovalka pred prvim snemanjem na dom bolnika prinese opremo in natančno razloži celoten potek raziskave.

Vsako posamično snemanje traja približno 15 minut, sodelujoči bolniki pa ga opravljajo s partnerjem. Oba najprej prebereta spisek stavkov, nato pa igrata dve igri: v eni iščeta razlike med dvema slikama, v drugi pa igrata igro s kartami, katere cilj je, da igralec zbere 4 karte iste vrste (npr. vse karte, ki imajo slike psov različnih pasem).

Varstvo in obdelava podatkov

Po pravilih Univerze v Groningenu mora vsak raziskovalec pred začetkom raziskave pridobiti etično dovoljenje od univerzitetne raziskovalne komisije (*CETO*), ki obenem tudi potrdi, da raziskava ne potrebuje širšega zdravstvenega etičnega dovoljenja. Anonimnost podatkov je zagotovljena v skladu z novo uredbo GDPR, ki stopa v uporabo 25. 5. 2018.

Osebni podatki sodelujočih bodo obravnavani tajno. V roku štirih tednov od prenehanja raziskave bodo podatki v celoti anonimizirani, kar pomeni, da sodelujoči ne bodo več mogli zahtevati polnega dostopa do svojih podatkov ali svojih podatkov odstraniti iz nabora podatkov raziskave.

Sodelujoči v raziskavi lahko s snemanjem govora kadarkoli prenehajo in jim pri tem ni treba navesti nobenega razloga.

Rezultati

Rezultati bodo objavljeni v magistrski nalogi, ki bo napisana za namene programa *Jezik in kognicija* na Univerzi v Groningenu (Nizozemska), ter v znanstvenem članku, izpeljanem iz magistrske naloge. Izsledki bodo prav tako objavljeni na spletni strani Društva Trepetlika, po dogovoru in želji pa se lahko za Društvo na isto temo pripravi tudi kratko predavanje.

O raziskovalki

Teja Rebernik je magistrska študentka in raziskovalna asistentka na Univerzi v Groningenu na Nizozemskem. Dela na presečišču jezikoslovja in kognitivnih znanosti, ter med drugim sodeluje v projektu, ki se ukvarja z govorom nizozemskih bolnikov s parkinsonovo boleznijo. Za seboj ima zaključen dodiplomski študij na Univerzi v Ljubljani ter podiplomski študij na Univerzi v Groningenu. V svojem prostem času veliko bere in igra violino, ko je v Sloveniji pa jo najdete v Mariboru, kjer rada zaide v gozdove Pohorja.

Teja Rebernik
Center for Language and Cognition
University of Groningen
E-mail: t.rebernik@rug.nl

Effect of levodopa and fatigue on the speech of patients with Parkinson's Disease

RESEARCH STUDY DESCRIPTION

Short description and importance of research study

One of the symptoms of Parkinson's disease is also disordered speech, i.e. hypokinetic dysarthria, the symptoms of which include mumbling, imprecise pronunciation of consonants and vowels, and quiet and hoarse speech. However, despite the pervasiveness of hypokinetic dysarthria, it is currently unclear how it is affected by levodopa and fatigue, which I will study in this research study.

The research study is important for two reasons. Firstly, studies on speech of Parkinson patients have been frequent over the past two decades, but researchers often measure speech only once for each patient, so they do not know whether they are measuring it at the most suitable moment. Secondly, most studies are done with native speakers of English, German and Dutch, but less so on other languages, which have significantly different characteristics.

By studying the speech of Slovenian patients with Parkinson's Disease, I will be able to contribute to the discussion on which characteristics appear only in certain languages, and which are universal. I also expect to be able to highlight how studying the speech of Slovenian patients brings useful results and contributes towards a better understanding of the effect of levodopa and fatigue on speech.

Participant sample

The sample will consist of between 5 to 10 patients. Eligible participants with Parkinson's Disease are those who:

- don't have any brain injuries and haven't had any strokes in the past;
- aren't diagnosed with depression or have a milder form of depression for which they do not need medication;
- take levodopa in the form of pills;
- didn't undergo deep brain stimulation;
- don't have any other speech disorders (e.g. stuttering);
- have Slovenian as their first, native, language.

Course of research study

The participants will use recording equipment (i.e. a smartphone and headset microphones) to record their speech, namely four days across two weeks. They record themselves five times a day; the exact times depend on the schedule of levodopa intake. The study demands

approximately five hours of the patients' time altogether. Before the first recording session, the researcher brings equipment to the participant's home and explains the proceedings of the study.

Each individual recording session takes approximately 15 minutes, and the participants do it with a partner. Both first read a list of sentences, then play two games: in one game, they look for differences between two pictures, in the other, they play a card game whose goal is for the player to collect 4 cards in the same category (e.g. all cards that have pictures of dogs of different breeds).

Data safety and analysis

Permission to conduct the study will be obtained from the research ethics review committee (*CETO*) at the University of Groningen. This also confirms that the research study does not need a broader medical ethical permission. Data anonymity is ensured in accordance with the new GDPR regulations, which came into effect on 25. 5. 2018.

The participants' personal data will be treated confidentially. Within 4 weeks after the data collection period, the data will be completely anonymized. After the 4 weeks have passed, the participants will not be able to ask for full access to their data nor have their data be removed from the data set.

Participants can stop recording their speech at any time and drop out of the study. They do not need to state any reason for doing so.

Results

The results will be published in a Master's thesis, written for purposes of the programme *Language and Cognition,* an in a scientific article, derived from the Master's thesis. The results will also be published on the website from Association Trepetlika. If so desired, a short lecture on the topic can also be prepared for the Association.

About the researcher

Teja Rebernik is a Master student and research assistant at the University of Groningen, Netherlands. She works at the crossroads of linguistics and cognitive sciences, and is among other things also collaborating on a project on the speech of Dutch patients with Parkinson's Disease. She has finished her Bachelor's at the University of Ljubljana and a previous Master's at the University of Groningen. In her free time, she reads a lot and plays the violin. When she's in Slovenia, you can find her in Maribor, where she loves to hike across the forests of Pohorje.

Appendix 4: Consent form

SOGLASJE

Motnje govora v Parkinsonovi bolezni

Prebral/a sem informacijsko pismo za sodelujoče. Lahko sem vprašal/a dodatna vprašanja. Na svoja vprašanja sem dobil/a zadovoljive odgovore. Imel/a sem dovolj časa, da se odločim, ali želim sodelovati pri raziskavi.

Vem, da je sodelovanje v raziskavi popolnoma prostovoljno. Vem, da se lahko v kateremkoli trenutku odločim, da ne želim več sodelovati. Za to ne rabim podati nobenega razloga.

Vem, da raziskovalci lahko vidijo moje osebne podatke. Pravico imam vedeti, kako so moji osebni podatki shranjeni in zavarovani.

Raziskovalcem dovolim, da moje podatke uporabijo za namene, ki so zapisani v informacijskem pismu. Če obstaja razlog, da bi moji podatki bili ponovno uporabljeni za kak drug namen, bodo me raziskovalci ponovno vprašali za dovoljenje.

Vem, da bodo moji raziskovalni podatki (tj. posnetki) poslani preko aplikacij Google Drive in MetaCtrl. Ti podjetji imata dostop do teh posnetkov.

Ime sodelujoči:

Podpis: Datum: ___ | ___ | _____

S spodnjim podpisom potrjujem, da sem sodelujočega/sodelujočo obvestil o tej raziskavi.

Če med potekom raziskave pride do sprememb, ki bi lahko vplivale na soglasje sodelujoče osebe, potem bom ga/jo o tem pravočasno obvestil.

Ime raziskovalca (ali nadomestnega raziskovalca):

Podpis: Datum: ___ | ___ | _____

CONSENT FORM

Speech problems in Parkinson's disease

I have read the information letter for the participants. I was able to ask additional questions. I obtained satisfactory answers to my questions. I had enough time to decide whether I wish to participate in the study.

I know that participating in the study is completely voluntary. I know that at any moment, I can decide that I do not wish to participate anymore. I do not need to give any reason for this.

I know that the researchers can see my personal information. I have the right to know how my personal information is stored and protected.

I allow the research to use my data for the purposes written in the information letter. If there is a reason for my data to be used for a different purpose, the researchers will ask me for permission again.

I know that my research data (namely recordings) will be sent through Google Drive and MetaCtrl apps. These two companies have access to these recordings.

Name of participant:

Signature: Date: ___ | ___ | _____

With the signature below, I confirm that I have sufficiently informed the participant about this study.

If, during the study, any changes occur that could affect the participants' consent, I will inform him/her about this on time.

Name of researcher (or substitute researcher):

Signature: Date: ___ | ___ | _____

Appendix 5: Instruction sheet for participants

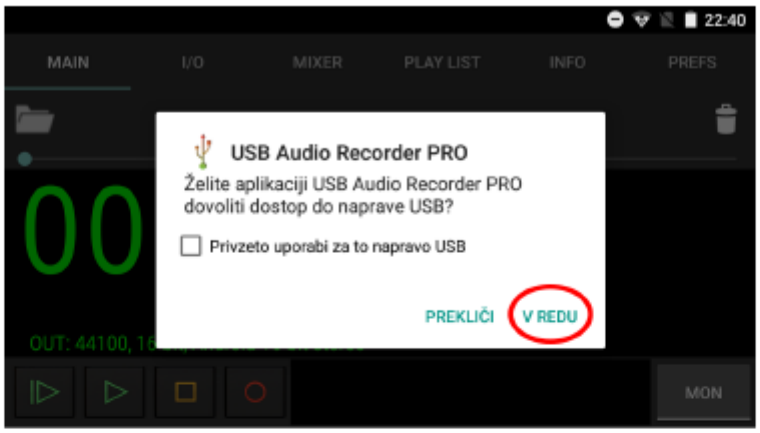
Navodila

1 Kako povezati napravo

- Poskrbite, da ima telefon polno baterijo.
- Telefon in napravo vedno nastavite na isto mesto.
- Črno škatlico povežite s telefonom (preko kabla USB). Na škatlici bodo zasvetile modre lučke.

- Na telefonu odprite aplikacijo "USB Aud..."
- Aplikacija bo zdaj prosila za dovoljenje za dostop do naprave USB. Kliknite V REDU.

1



- Nataknite si mikrofon. Poskrbite, da je kroglica (končni del mikrofona) oddaljena približno 1 cm od vaših ust. Med snemanjem poskusite **ne** premikati mikrofona. Če je mikrofon nameščen neudobno in ga morate premakniti, potem na glas recite, da ga boste premaknili, preden to storite.
- Kliknite na gumb "snemaj" (rdeči krog). Dokler teče čas na ekranu, telefon snema. Ekran se ne bo izklopil.

2 Ocena utrujenosti

- Na listu *Ocena utrujenosti* označite, kako utrujeno se počutite v tistem trenutku. Te ocene ne rabite povedati na glas.

3 Spisek besed

- Vzemite enega od dveh kupov. Premešajte karte in jih položite s hrbtom navzdol. Potegnite vsako karto posebej in preberite stavek na karti v normalnem bralnem tempu.

4 Iskanje razlik (5 minut)

- Na sliki so razlike v lokaciji, barvi in podrobnostih predmetov. Najdite 10 razlik med svojo sliko in sliko soigralca. Poskusite tvoriti polne stavke in izogibajte se pomanjševalnicam (npr. recite žaba in ne žabica). Soigralec ne sme videti vaše slike. Razlike iščete 5 minut oziroma manj, če vseh 10 razlik najdete prej kot v 5 minutah.
- Seznam razlik je priložen v kuverti, če po koncu snemanja želite pogledati rešitve.

5 Kvartet (5 minut)

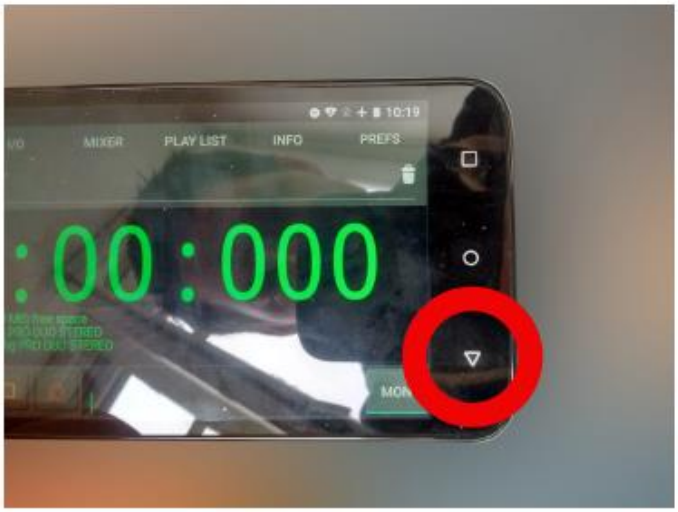
- Vsak začne s 6 kartami. Svojega soigralca/soigralko povprašajte o njegovih/njenih kartah na naslednji način: **Lahko od tebe iz kategorije [ime kategorije] dobim [ime karte]?** (npr. "Lahko od tebe iz kategorije *kapa* dobim *klobuk*?")
- Cilj igre je zbrati čim več kategorij v celoti (npr. vse štiri kape ali vse štiri rublje).
- Kvartet igrate 5 minut.

6 Ponavljanje

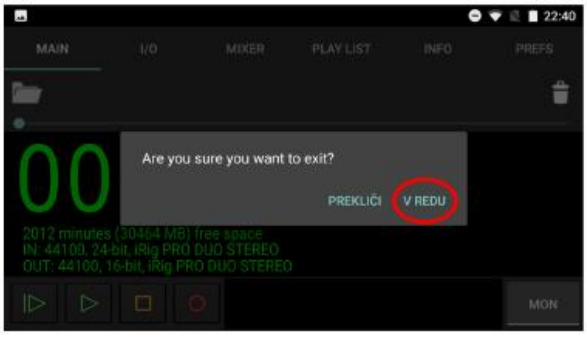
- Čim večkrat in čim hitreje ponovite zloge "pa", "ta", "ka" in "pataka" v eni sapi.

7 Kako izklopiti napravi

- Ko končate s snemanjem, pritisnite na gumb "stop" (rumeni kvadrat).
- Ob koncu vsakega snemanja pritisnite na trikotnik na dnu telefona.



Nato pritisnite V REDU.



- Črno škatlico odklopite od telefona. Mikrofona in kabel USB pustite priklopljene na črno škatlico.
- Telefon pustite vklopljen in pri miru 15 minut.
- Telefon dajte polniti.

8 Pomoč

- Če imate kakršnakoli vprašanja, lahko mirno pokličete ali mi napišete SMS. Tudi kratka vprašanja niso čisto noben problem! Kličete/pišete lahko na 041 532 024.

Appendix 6: MMSE (Slovenian version)

Koda sodelujočega: **Ime raziskovalca/raziskovalke:** Teja Rebernik

Orientacijska vprašanja

1.	Katerega leta smo?	
2.	V katerem letnem času smo?	
3.	Katerega meseca smo?	
4.	Kateri dan v tednu je danes?	
5.	Kateri datum je danes?	
6.	V kateri državi živite?	
7.	V kateri regiji živite?	
8.	V katerem kraju ste sedaj?	
9.	Kdaj ste bili rojeni?	
10.	Koliko ste stari?	

Pomnjenje

»Povedala vam bom nekaj besed. Zapomnite si jih in jih ponovite, ko jih končam.« *(Besede izgovarjamo v razmaku ene sekunde.* Žoga. Drevo. Zastava. *Besede bolnik ponavlja, dokler jih ne zna.)*

11.	**Žoga**	
12.	**Drevo**	
13.	**Zastava**	

Pozornost in računanje

»Odštevajte od 100 po 7. Od dobljenega števila spet 7 in tako naprej. Torej: 100 minus 7 je …?«.
Alternativno: »Povedala vam bo besedo. Črkujte jo v obratnem vrstnem redu. Beseda je: **lonec**.«

14.	**93**	**C**	
15.	**86**	**E**	
16.	**79**	**N**	
17.	**72**	**O**	
18.	**65**	**L**	

Priklic

»Ponovite prosim tri stvari, za katere sem vam prej naročila, da si jih zapomnite.«

19.	**Žoga**	
20.	**Drevo**	
21.	**Zastava**	

Jezik

22.	Kaj je to? (Ura)	
23.	Kaj je to? (Svinčnik)	
24.	Ponovite za menoj stavek: **Nobenih in, če, ampak.**	
25.	Prepognite tale list papirja na polovico,	
26.	ga položite na kolena,	
27.	nato pa ga izročite meni.	
28.	Preberite, kaj piše na tem listu papirja in potem to naredite: **Zaprite oči.**	
29.	Napišite, prosim, nek stavek.	
30.	Prerišite tole.	

_____/30

Datum preizkusa: __ / __ / ____

Appendix 7: Questionnaire

Demografski podatki *[demographic information]*

Spol: Ž M *[gender]*

Stopnja izobrazbe: *[educational level]*

I II III IV V VI/1 VI/2 VII VIII/1 VIII/2

Kraj rojstva: _____ *[Place of birth. Where have you spent most of your life?]*

V katerem kraju ste preživeli večino svojega življenja? _____

Jezikovni podatki *[language information]*

Govorite še kak drug jezik? DA NE *[Do you speak any other languages? Yes/No. If yes, which?]*

Če da, katerega? _____

Diagnoza parkinsonove bolezni *[Parkinson's disease diagnosis]*

Kdaj so vam postavili diagnozo parkinsonove bolezni? _____
[When were you diagnosed with Parkinson's disease?]
Ali parkinsonova bolezen vpliva na vaš govor? DA NE
[Does Parkinson's disease influence your speech? Yes / No. If yes, how?]
Če da, kako? _____
_____

Zdravila *[medication]*

Katera zdravila jemljete? *[Which medication do you take?]*

_____
_____
_____

Kolikokrat na dan vzamete levodopo? _____ *[How frequently do you take levodopa?]*

Ob katerih časih jemljete levodopo? *[When do you take levodopa?]*

_____
_____

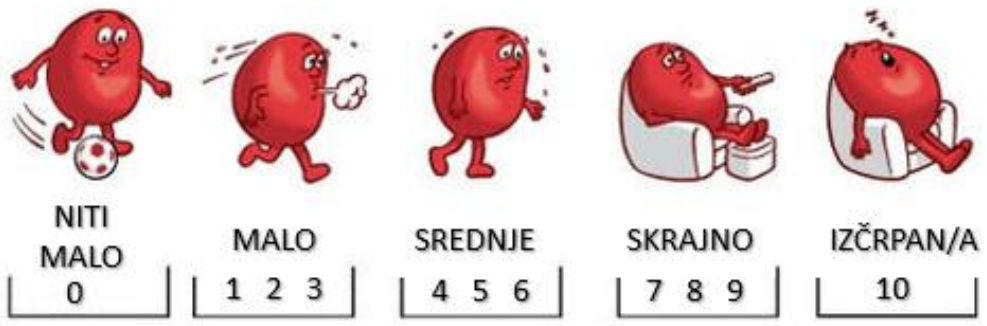
Izkusite stanje vklopa in izklopa? DA NE *[Do you experience ON/OFF states?]*

Koliko let jemljete levodopo? _____ *[How long have you been taking levodopa?]*

Appendix 8: Fatigue survey

OCENA UTRUJENOSTI – PRVI DAN

Pred pričetkom snemanja prosim obkrožite, kako utrujeno se počutite v tistem trenutku.



PRVO (1) snemanje

0 1 2 3 4 5 6 7 8 9 10

DRUGO (2) snemanje

0 1 2 3 4 5 6 7 8 9 10

TRETJE (3) snemanje

0 1 2 3 4 5 6 7 8 9 10

ČETRTO (4) snemanje

0 1 2 3 4 5 6 7 8 9 10

PETO (5) snemanje

0 1 2 3 4 5 6 7 8 9 10

Appendix 9: List of carrier phrases

Bilabial voiceless sound (/p/)
Beseda kapa ima več kot en zlog.
Beseda papež ima več kot en zlog.
Beseda pipa ima več kot en zlog.
Beseda ekipa ima več kot en zlog.
Beseda lupa ima več kot en zlog.
Beseda pupa ima več kot en zlog.

Bilabial voiced sound (/b/)
Beseda žaba ima več kot en zlog.
Beseda kabel ima več kot en zlog.
Beseda riba ima več kot en zlog.
Beseda šiba ima več kot en zlog.
Beseda rubelj ima več kot en zlog.
Beseda tuba ima več kot en zlog.

Dental voiceless sound (/t/)
Beseda solata ima več kot en zlog.
Beseda vrata ima več kot en zlog.
Beseda kita ima več kot en zlog.
Beseda pita ima več kot en zlog.
Beseda ruta ima več kot en zlog.
Beseda valuta ima več kot en zlog.

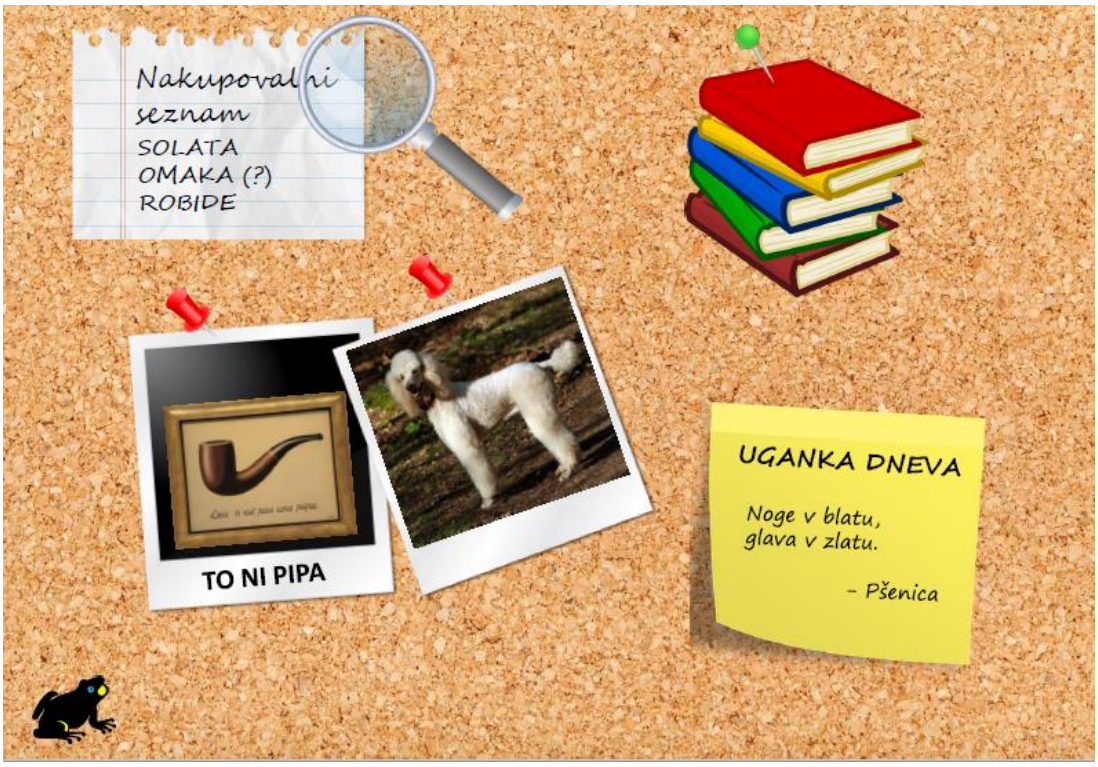
Dental voiced sound (/d/)
Beseda brada ima več kot en zlog.
Beseda čelada ima več kot en zlog.
Beseda robida ima več kot en zlog.
Beseda piramida ima več kot en zlog.
Beseda pudelj ima več kot en zlog.
Beseda Buda ima več kot en zlog.

Velar voiceless sound (/k/)
Beseda omaka ima več kot en zlog.
Beseda mlaka ima več kot en zlog.
Beseda pika ima več kot en zlog.
Beseda slika ima več kot en zlog.
Beseda bukev ima več kot en zlog.
Beseda kljuka ima več kot en zlog.

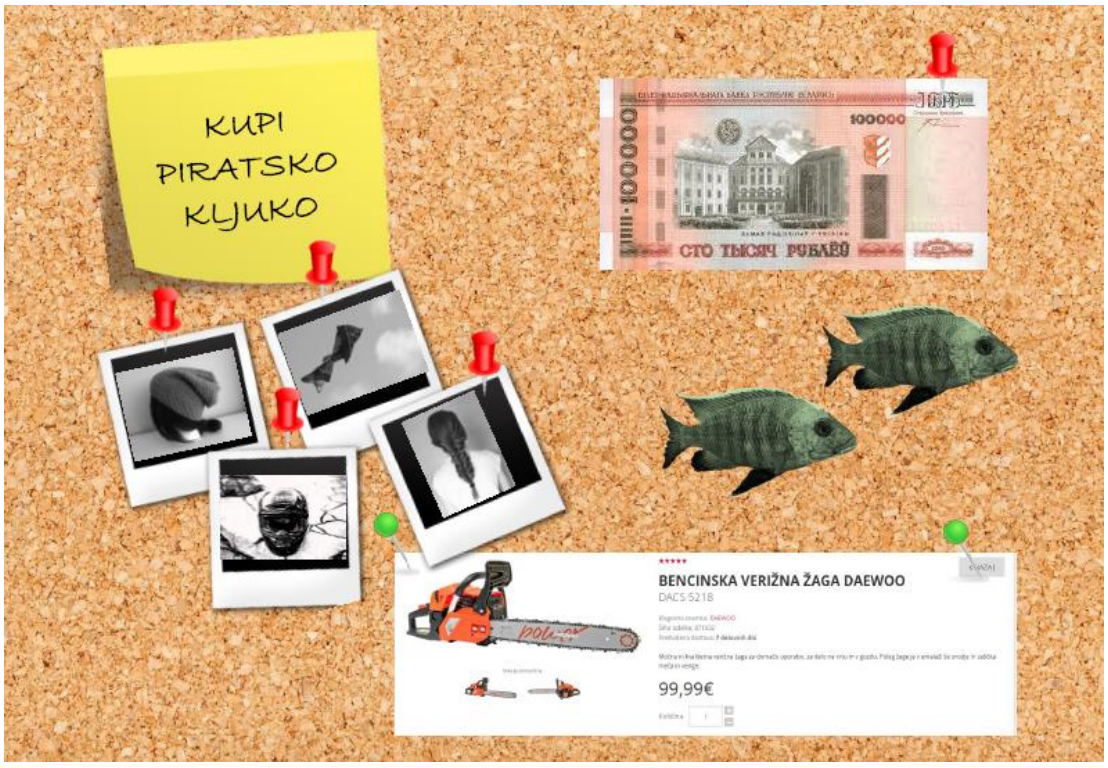
Velar voiced sound (/g/)
Beseda žaga ima več kot en zlog.
Beseda glagol ima več kot en zlog.
Beseda figa več kot en zlog.
Beseda knjiga ima več kot en zlog.
Beseda vijuga ima več kot en zlog.
Beseda uganka ima več kot en zlog.

Appendix 10: Game sheets (Spot-the-Differences)

Below are example game sheets for the Spot-the-Differences task. Participants received a different game sheet for every session. We created altogether 40 game sheets.




BENCINSKA VERIŽNA ŽAGA DAEWOO
DACS 5218
99,99€
KUPI
NOVO
KLJUKO
100000
СТО ТЫСЯЧ РУБЛЁЙ


KUPI
PIRATSKO
KLJUKO
100000
СТО ТЫСЯЧ РУБЛЁЙ
BENCINSKA VERIŽNA ŽAGA DAEWOO
DACS 5218
99,99€

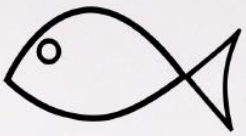

LP iz Egipta!
ZMAGOVALNA EKIPA
Buda pošilja lepe pozdrave
Pozdrav z vijugastih cest Irske.


LP od piramid!
ANGLEŠKA EKIPA
Buda pošilja lepe pozdrave

Francoski glagoli?

BRADA ALI BRKI?

BUKEV ALI LIPA?



VALUTE
Rubelj
Avstralski dolar
Turške lire

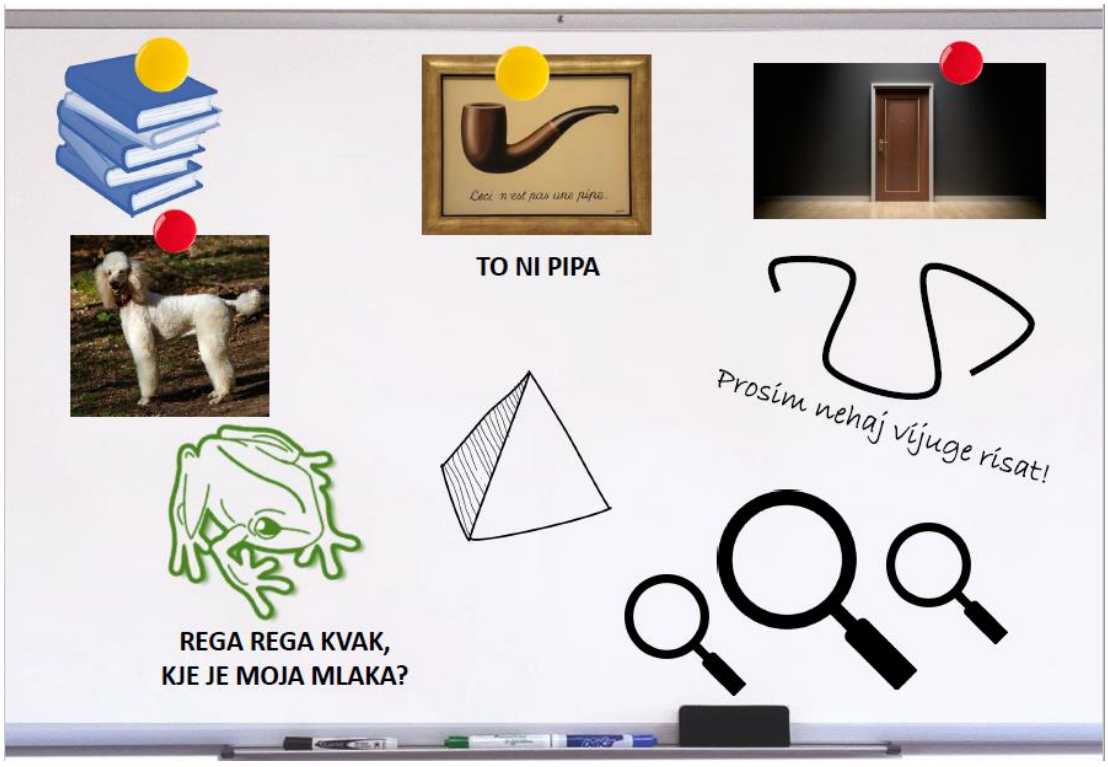
Nemški glagoli?

Brada je znak pravega moškega.

LIPA ALI BUKEV?

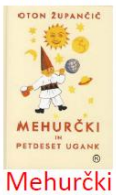
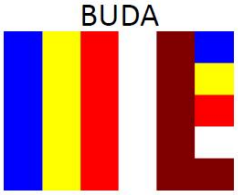
VALUTE
Tolar
Rubelj
Franki


Ceci n'est pas une pipe.
TO NI PIPA
Daj, nehaj vijuge risat!
LUPA
ŽABA DELA REGA REGA KVAK
V SVOJI MLAKI


Ceci n'est pas une pipe.
TO NI PIPA
Prosim nehaj vijuge risat!
REGA REGA KVAK,
KJE JE MOJA MLAKA?

Appendix 11: Kwartet cards

PITA  Jabolčna pita Marelična pita Francoska pita Bučkina pita	PITA  Jabolčna pita Marelična pita Francoska pita Bučkina pita	ČELADA  Motoristična čelada Kolesarska čelada Plezalna čelada Varovalna čelada	ČELADA  Motoristična čelada Kolesarska čelada Plezalna čelada Varovalna čelada
PITA  Jabolčna pita Marelična pita Francoska pita Bučkina pita	PITA  Jabolčna pita Marelična pita Francoska pita Bučkina pita	ČELADA  Motoristična čelada Kolesarska čelada Plezalna čelada Varovalna čelada	ČELADA  Motoristična čelada Kolesarska čelada Plezalna čelada Varovalna čelada

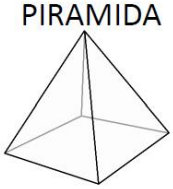
KNJIGA  Mehurčki Romeo in Julija Maček Muri Povodni mož	KNJIGA  Mehurčki Romeo in Julija Maček Muri Povodni mož	BUDA  Buda Budistična filozofija Budistični tempelj Budistična zastava	BUDA  Buda Budistična filozofija Budistični tempelj Budistična zastava
KNJIGA  Mehurčki Romeo in Julija Maček Muri Povodni mož	KNJIGA  Mehurčki Romeo in Julija Maček Muri Povodni mož	BUDA  Buda Budistična filozofija Budistični tempelj Budistična zastava	BUDA  Buda Budistična filozofija Budistični tempelj Budistična zastava

BUKEV	BUKEV	EKIPA	EKIPA
			
Drevo Bukov list Deblo Jesenska bukev	Drevo Bukov list Deblo Jesenska bukev	Španska ekipa Srbska ekipa Italijanska ekipa Madžarska ekipa	Španska ekipa Srbska ekipa Italijanska ekipa Madžarska ekipa
BUKEV	**BUKEV**	**EKIPA**	**EKIPA**
			
Drevo Bukov list Deblo Jesenska bukev	Drevo Bukov list Deblo Jesenska bukev	Španska ekipa Srbska ekipa Italijanska ekipa Madžarska ekipa	Španska ekipa Srbska ekipa Italijanska ekipa Madžarska ekipa

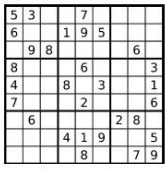
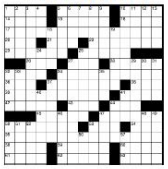
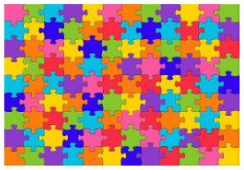
SLIKA	SLIKA	KAPA	KAPA
			
Mona Liza Krik Zvezdnato nebo Vztrajnost spomina	Mona Liza Krik Zvezdnato nebo Vztrajnost spomina	Klobuk Kapa s šiltom Baretka Slamnik	Klobuk Kapa s šiltom Baretka Slamnik
SLIKA	**SLIKA**	**KAPA**	**KAPA**
			
Mona Liza Krik Zvezdnato nebo Vztrajnost spomina	Mona Liza Krik Zvezdnato nebo Vztrajnost spomina	Klobuk Kapa s šiltom Baretka Slamnik	Klobuk Kapa s šiltom Baretka Slamnik

SOLATA	SOLATA	ŽABA	ŽABA
			
Kristalka Paradižnikova solata Motovilec Regrat	Kristalka Paradižnikova solata Motovilec Regrat	Krastača Česnovka Sekulja Zelena rega	Krastača Česnovka Sekulja Zelena rega
SOLATA	**SOLATA**	**ŽABA**	**ŽABA**
			
Kristalka Paradižnikova solata Motovilec Regrat	Kristalka Paradižnikova solata Motovilec Regrat	Krastača Česnovka Sekulja Zelena rega	Krastača Česnovka Sekulja Zelena rega

RIBA	RIBA	RUBELJ	RUBELJ
			
Losos Postrv Morski pes Zlata ribica	Losos Postrv Morski pes Zlata ribica	En rubelj Dva rublja Pet rubljev Dvesto rubljev	En rubelj Dva rublja Pet rubljev Dvesto rubljev
RIBA	**RIBA**	**RUBELJ**	**RUBELJ**
			
Losos Postrv Morski pes Zlata ribica	Losos Postrv Morski pes Zlata ribica	En rubelj Dva rublja Pet rubljev Dvesto rubljev	En rubelj Dva rublja Pet rubljev Dvesto rubljev

PIRAMIDA	PIRAMIDA	OMAKA	OMAKA
			
Egipčanska Prehrambena Geometrična Louvre	Egipčanska Prehrambena Geometrična Louvre	Bolonjeze Gobova omaka Bešamel Gorčična omaka	Bolonjeze Gobova omaka Bešamel Gorčična omaka
PIRAMIDA	**PIRAMIDA**	**OMAKA**	**OMAKA**
			
Egipčanska Prehrambena Geometrična Louvre	Egipčanska Prehrambena Geometrična Louvre	Bolonjeze Gobova omaka Bešamel Gorčična omaka	Bolonjeze Gobova omaka Bešamel Gorčična omaka

PUPA	PUPA	VALUTA	VALUTA
			
Rjavolasa pupa Blontna pupa Barbika Porcelanasta pupa	Rjavolasa pupa Blontna pupa Barbika Porcelanasta pupa	Evro Ameriški dolar Japonski jen Britanski funt	Evro Ameriški dolar Japonski jen Britanski funt
PUPA	**PUPA**	**VALUTA**	**VALUTA**
			
Rjavolasa pupa Blontna pupa Barbika Porcelanasta pupa	Rjavolasa pupa Blontna pupa Barbika Porcelanasta pupa	Evro Ameriški dolar Japonski jen Britanski funt	Evro Ameriški dolar Japonski jen Britanski funt

GLAGOL  Prepevati Telefonirati Kolesariti Brati	GLAGOL  Prepevati Telefonirati Kolesariti Brati	UGANKA  Sudoku Križanka Rubikova kocka Sestavljanka	UGANKA  Sudoku Križanka Rubikova kocka Sestavljanka
GLAGOL  Prepevati Telefonirati Kolesariti Brati	GLAGOL  Prepevati Telefonirati Kolesariti Brati	UGANKA  Sudoku Križanka Rubikova kocka Sestavljanka	UGANKA  Sudoku Križanka Rubikova kocka Sestavljanka

Appendix 12: Statistical analysis

```
#test the null hypothesis that there is no difference in score between PD patients and healthy speakers, with g
roup as fixed effect and random intercepts for subjects

# if the score is lower, it means that it is more pathological (lower tVSA, lower VAI, lower CPPS, lower F0)

modell <- lmer(z.score ~ group + (1|participant.id), data = data)
summary(modell)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2457.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5682 -0.6023 -0.0073  0.6097  3.7356
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.3305   0.5749
##  Residual                   0.4997   0.7069
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   0.5888     0.2894  7.9987   2.035   0.0763 .
## groupPD      -0.9750     0.3736  7.9974  -2.610   0.0311 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr)
## groupPD -0.775
```

```
lme.dscore(modell, data, type="lme4")
```

```
##                t       df         d
## groupPD -2.609776 7.997369 -1.845694
```

Our hypothesis test shows that the PD group has significantly lower scores compared to the healthy speakers.

```
# model with group and sex as fixed effect and random intercepts for subjects
model2 <- lmer(z.score ~ sex + group + (1|participant.id), data = data)
summary(model2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2448.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5546 -0.6051 -0.0056  0.6162  3.7366
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1171   0.3422
##  Residual                   0.4997   0.7069
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   0.8576     0.1876  7.0127   4.572  0.00256 **
## sexmale      -1.0754     0.2772  7.0270  -3.880  0.00601 **
## groupPD      -0.3474     0.2772  7.0270  -1.253  0.25017
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) sexmal
## sexmale -0.369
## groupPD -0.369 -0.584
```

```
lme.dscore(model2, data, type="lme4")
```

```
##                t       df         d
## sexmale -3.879636 7.027019 -2.9270855
## groupPD -1.253354 7.027019 -0.9456233
```

```
anova(model1, model2) # sex significantly improves the model (we continue with model 2)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model1: z.score ~ group + (1 | participant.id)
## model2: z.score ~ sex + group + (1 | participant.id)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model1  4 2463.2 2483.3 -1227.6   2455.2
## model2  5 2453.7 2478.9 -1221.9   2443.7 11.479      1   0.000704 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
modelsex <- lmer(z.score ~ sex*group + (1|participant.id), data = data)
summary(modelsex)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex * group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2447.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5513 -0.6061 -0.0056  0.6177  3.7365
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1356   0.3683
##  Residual                   0.4997   0.7069
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                  Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)       0.8782     0.2161  6.0077   4.064   0.0066 **
## sexmale          -1.1581     0.4324  6.0201  -2.678   0.0365 *
## groupPD          -0.4301     0.4324  6.0201  -0.995   0.3581
## sexmale:groupPD   0.1572     0.5960  6.0203   0.264   0.8008
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sexmal gropPD
## sexmale     -0.500
## groupPD     -0.500  0.250
## sexml:grpPD  0.363 -0.725 -0.725
```

```
lme.dscore(modelsex, data, type="lme4")
```

```
##                           t       df          d
## sexmale         -2.6783717 6.020065 -2.1832339
## groupPD         -0.9947810 6.020065 -0.8108806
## sexmale:groupPD  0.2637102 6.020261  0.2149558
```

```
anova(model2, modelsex) # sex in interaction with group does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## modelsex: z.score ~ sex * group + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2    5 2453.7 2478.9 -1221.9   2443.7
## modelsex  6 2455.6 2485.8 -1221.8   2443.6 0.1151      1     0.7344
```

Model with group*sex as fixed factor is not significantly better than model with sex and group as fixed factors.

```
modeltask <- lmer(z.score ~ task + sex + group + (1|participant.id), data = data)
summary(modeltask)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ task + sex + group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2453
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5482 -0.6068 -0.0087  0.6081  3.7405
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1170   0.3421
##  Residual                   0.5001   0.7072
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     0.86729    0.19049    7.46611   4.553  0.00223 **
## taskwordlist   -0.01349    0.04650 1115.10984  -0.290  0.77176
## sexmale        -1.07549    0.27714    7.02699  -3.881  0.00600 **
## groupPD        -0.34751    0.27714    7.02699  -1.254  0.24997
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) tskwrd sexmal
## taskwordlst -0.176
## sexmale     -0.364  0.001
## groupPD     -0.364  0.001 -0.584
```

```
lme.dscore(modeltask, data, type="lme4")
```

```
##                      t          df          d
## taskwordlist -0.290140 1115.109842 -0.01737716
## sexmale      -3.880708    7.026985 -2.92790108
## groupPD      -1.253939    7.026985 -0.94606675
```

```
anova(model2, modeltask) # adding task does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## modeltask: z.score ~ task + sex + group + (1 | participant.id)
##           Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2     5 2453.7 2478.9 -1221.9   2443.7
## modeltask  6 2455.7 2485.8 -1221.8   2443.7 0.0854      1     0.7701
```

```
# model with group+sex and type (i.e. tVSA, VAI, F0, CPPS) as fixed effects
model3 <- lmer(z.score ~ type + sex + group + (1|participant.id), data = data)
summary(model3)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ type + sex + group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2460.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5445 -0.6062 -0.0079  0.6122  3.7371
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.117    0.3421
##  Residual                   0.501    0.7078
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                 Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)    8.598e-01  1.900e-01  7.393e+00   4.524  0.00237 **
## typeVAI        1.013e-16  5.254e-02  1.113e+03   0.000  1.00000
## typeMeanF0    -6.052e-03  6.234e-02  1.113e+03  -0.097  0.92268
## typeMeanCPPS -6.052e-03  6.234e-02  1.113e+03  -0.097  0.92268
## sexmale       -1.075e+00  2.772e-01  7.027e+00  -3.880  0.00601 **
## groupPD       -3.474e-01  2.772e-01  7.027e+00  -1.253  0.25012
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) typVAI typMF0 tMCPPS sexmal
## typeVAI     -0.138
## typeMeanF0  -0.117  0.421
## typeMenCPPS -0.117  0.421  0.355
## sexmale     -0.364  0.000  0.000  0.000
## groupPD     -0.364  0.000  0.000  0.000 -0.584
```

```
lme.dscore(model3, data, type="lme4")
```

```
##                          t          df             d
## typeVAI       1.928485e-15 1113.008344  1.156105e-16
## typeMeanF0   -9.707973e-02 1113.019750 -5.819788e-03
## typeMeanCPPS -9.707973e-02 1113.019750 -5.819788e-03
## sexmale      -3.879896e+00    7.027251 -2.927233e+00
## groupPD      -1.253492e+00    7.027251 -9.457115e-01
```

```
anova(model2, model3) # type does not improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## model3: z.score ~ type + sex + group + (1 | participant.id)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2  5 2453.7 2478.9 -1221.9   2443.7
## model3  8 2459.7 2499.9 -1221.9   2443.7 0.0192      3     0.9993
```

```
# model with group+sex and fatigue as fixed effects
model4 <- lmer(z.score ~ fatigue + sex + group + (1|participant.id), data = data)
summary(model4)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ fatigue + sex + group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2455.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5520 -0.6063 -0.0047  0.6117  3.7310
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1180   0.3435
##  Residual                   0.5001   0.7072
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##              Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  8.620e-01  1.910e-01  7.380e+00   4.514  0.00241 **
## fatigue     -2.086e-03  1.495e-02  1.085e+03  -0.140  0.88902
## sexmale     -1.073e+00  2.786e-01  7.035e+00  -3.852  0.00621 **
## groupPD     -3.481e-01  2.783e-01  7.001e+00  -1.251  0.25122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) fatigu sexmal
## fatigue -0.167
## sexmale -0.355 -0.052
## groupPD -0.367  0.017 -0.584
```

```
lme.dscore(model4, data, type="lme4")
```

```
##                  t          df            d
## fatigue -0.1395809 1085.373307 -0.008473569
## sexmale -3.8523408    7.034566 -2.904932553
## groupPD -1.2507121    7.001444 -0.945351988
```

```
anova(model2, model4) # fatigue does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## model4: z.score ~ fatigue + sex + group + (1 | participant.id)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2  5 2453.7 2478.9 -1221.9   2443.7
## model4  6 2455.7 2485.9 -1221.9   2443.7 0.0078      1     0.9297
```

```
# model with group+sex and time of day (morning versus afternoon) as fixed effects
model5 <- lmer(z.score ~ time + sex + group + (1|participant.id), data = data)
summary(model5)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ time + sex + group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2450.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5669 -0.6015  0.0047  0.5990  3.6771
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1173   0.3424
##  Residual                   0.4989   0.7063
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)    0.82877    0.18851    7.13496   4.397  0.00303 **
## timeevening    0.07062    0.04285 1115.08093   1.648  0.09960 .
## sexmale       -1.07411    0.27738    7.02717  -3.872  0.00607 **
## groupPD       -0.34875    0.27739    7.02718  -1.257  0.24883
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) tmvnng sexmal
## timeevening -0.093
## sexmale     -0.368  0.003
## groupPD     -0.367 -0.003 -0.584
```

```
lme.dscore(model5, data, type="lme4")
```

```
##                     t          df           d
## timeevening  1.648155 1115.080929  0.09871314
## sexmale     -3.872275    7.027173 -2.92149999
## groupPD     -1.257278    7.027182 -0.94857323
```

```
anova(model2, model5) # time does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## model5: z.score ~ time + sex + group + (1 | participant.id)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2  5 2453.7 2478.9 -1221.9   2443.7
## model5  6 2453.0 2483.2 -1220.5   2441.0 2.7113      1    0.09964 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# model with group+sex and day as fixed effects
model6 <- lmer(z.score ~ day + sex + group + (1|participant.id), data = data)
summary(model6)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ day + sex + group + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2457.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4700 -0.6033 -0.0033  0.6016  3.7286
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1167   0.3417
##  Residual                   0.4994   0.7067
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)    0.90688    0.19106    7.59397   4.747  0.00167 **
## daysecond     -0.03965    0.06067 1113.12230  -0.653  0.51361
## daythird      -0.04274    0.06006 1113.10036  -0.712  0.47687
## dayfourth     -0.11036    0.06014 1113.08002  -1.835  0.06677 .
## sexmale       -1.07493    0.27680    7.02731  -3.884  0.00598 **
## groupPD       -0.34775    0.27679    7.02717  -1.256  0.24914
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) dyscnd dythrd dyfrth sexmal
## daysecond -0.162
## daythird  -0.163  0.520
## dayfourth -0.164  0.520  0.525
## sexmale   -0.361 -0.002 -0.003 -0.001
## groupPD   -0.362  0.000  0.001  0.001 -0.584
```

```
lme.dscore(model6, data, type="lme4")
```

```
##                   t          df          d
## daysecond -0.6534412 1113.122296 -0.03917104
## daythird  -0.7115798 1113.100356 -0.04265662
## dayfourth -1.8349888 1113.080023 -0.11000191
## sexmale   -3.8835005    7.027313 -2.92993996
## groupPD   -1.2563486    7.027167 -0.94787274
```

```
anova(model2, model6) # day does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## model6: z.score ~ day + sex + group + (1 | participant.id)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2  5 2453.7 2478.9 -1221.9   2443.7
## model6  8 2456.2 2496.4 -1220.1   2440.2 3.5376      3     0.3159
```

```
modelsession <- lmer(z.score ~ session + group + sex + (1|participant.id), data=data)
summary(modelsession)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ session + group + sex + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2458.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5647 -0.6200 -0.0054  0.6041  3.6781
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1173   0.3424
##  Residual                   0.4991   0.7064
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     0.76593    0.19248    7.75463   3.979  0.00433 **
## sessiontwo      0.09345    0.06662 1112.02515   1.403  0.16097
## sessionthree    0.09436    0.06753 1112.04735   1.397  0.16262
## sessionfour     0.13313    0.06634 1112.01984   2.007  0.04502 *
## sessionfive     0.13324    0.06662 1112.03058   2.000  0.04575 *
## groupPD        -0.34883    0.27739    7.02718  -1.258  0.24875
## sexmale        -1.07362    0.27739    7.02724  -3.870  0.00608 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sssntw sssnth sssnfr sssnfv gropPD
## sessiontwo  -0.177
## sessionthre -0.173  0.500
## sessionfour -0.177  0.509  0.502
## sessionfive -0.176  0.506  0.500  0.509
## groupPD     -0.360  0.000  0.000 -0.003 -0.001
## sexmale     -0.360  0.002  0.000  0.004  0.001 -0.584
```

```
lme.dscore(modelsession, data, type="lme4")
```

```
##                     t          df          d
## sessiontwo    1.402746 1112.025151  0.08413017
## sessionthree  1.397243 1112.047349  0.08379926
## sessionfour   2.006763 1112.019835  0.12035658
## sessionfive   1.999966 1112.030582  0.11994833
## groupPD      -1.257510    7.027181 -0.94874796
## sexmale      -3.870370    7.027242 -2.92004803
```

```
lme.dscore(modelsession, data, type="lme4")
```

```
##                     t          df          d
## sessiontwo    1.402746 1112.025151  0.08413017
## sessionthree  1.397243 1112.047349  0.08379926
## sessionfour   2.006763 1112.019835  0.12035658
## sessionfive   1.999966 1112.030582  0.11994833
## groupPD      -1.257510    7.027181 -0.94874796
## sexmale      -3.870370    7.027242 -2.92004803
```

```
anova(model2, modelsession) # adding session does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## modelsession: z.score ~ session + group + sex + (1 | participant.id)
##              Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## model2        5 2453.7 2478.9 -1221.9   2443.7
## modelsession  9 2456.4 2501.6 -1219.2   2438.4 5.331      4      0.255
```

Model 2 remains better (lower AIC, not significant).

We stick to model 2 and check for interactions.

```
# Check whether patients are more susceptible to fatigue.

modelgroupfatigue <- lmer(z.score ~ sex + group*fatigue + (1|participant.id), data = data)
summary(modelgroupfatigue)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + group * fatigue + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2460.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5496 -0.6085 -0.0061  0.6094  3.7335
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1180   0.3434
##  Residual                   0.5005   0.7075
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)      8.658e-01  1.936e-01  7.760e+00   4.472  0.00224 **
## sexmale         -1.075e+00  2.789e-01  7.020e+00  -3.854  0.00622 **
## groupPD         -3.562e-01  2.868e-01  7.823e+00  -1.242  0.25015
## fatigue         -3.502e-03  1.924e-02  1.119e+03  -0.182  0.85564
## groupPD:fatigue  3.575e-03  3.054e-02  1.013e+03   0.117  0.90684
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sexmal gropPD fatigu
## sexmale     -0.357
## groupPD     -0.391 -0.554
## fatigue     -0.232 -0.010  0.165
## groupPD:ftg  0.165 -0.047 -0.242 -0.630
```

```
lme.dscore(modelgroupfatigue, data, type="lme4")
```

```
##                         t          df           d
## sexmale         -3.8542994    7.020156 -2.909390779
## groupPD         -1.2420905    7.823010 -0.888170402
## fatigue         -0.1819748 1118.888954 -0.010880475
## groupPD:fatigue  0.1170542 1013.340495  0.007354268
```

```
modelgrouptask <- lmer(z.score ~ sex + group*task + (1|participant.id), data = data)
summary(modelgrouptask)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + group * task + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2455.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5511 -0.6050 -0.0099  0.6033  3.7420
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1170   0.3421
##  Residual                   0.5005   0.7075
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                       Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)            0.85880    0.19492    8.18496   4.406  0.00215 **
## sexmale               -1.07583    0.27714    7.02771  -3.882  0.00599 **
## groupPD               -0.33333    0.28559    7.92461  -1.167  0.27707
## taskwordlist          -0.00154    0.07444 1114.05646  -0.021  0.98350
## groupPD:taskwordlist  -0.01961    0.09535 1114.02602  -0.206  0.83712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sexmal gropPD tskwrd
## sexmale     -0.354
## groupPD     -0.396 -0.568
## taskwordlst -0.273 -0.004  0.189
## grpPD:tskwr  0.212  0.006 -0.242 -0.781
```

```
lme.dscore(modelgrouptask, data, type="lme4")
```

```
##                                t          df           d
## sexmale              -3.88185616    7.027715 -2.928615483
## groupPD              -1.16714706    7.924614 -0.829213796
## taskwordlist         -0.02068791 1114.056462 -0.001239633
## groupPD:taskwordlist -0.20563112 1114.026024 -0.012321715
```

```
modelsexfatigue <- lmer(z.score ~ group + sex*fatigue + (1|participant.id), data = data)
summary(modelsexfatigue)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ group + sex * fatigue + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2460.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5486 -0.6009 -0.0068  0.6135  3.7226
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1193   0.3454
##  Residual                   0.5005   0.7074
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                     Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)       8.561e-01  1.942e-01  7.691e+00   4.409  0.00248 **
## groupPD          -3.463e-01  2.799e-01  6.982e+00  -1.237  0.25594
## sexmale          -1.059e+00  2.888e-01  7.880e+00  -3.667  0.00651 **
## fatigue           6.148e-04  1.994e-02  1.120e+03   0.031  0.97541
## sexmale:fatigue -6.217e-03  3.016e-02  1.064e+03  -0.206  0.83670
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) gropPD sexmal fatigu
## groupPD     -0.367
## sexmale     -0.377 -0.559
## fatigue     -0.223  0.032  0.124
## sexmale:ftg  0.150 -0.030 -0.243 -0.661
```

```
lme.dscore(modelsexfatigue, data, type="lme4")
```

```
##                           t          df           d
## groupPD          -1.23739245    6.981684 -0.936606900
## sexmale          -3.66656466    7.880020 -2.612315919
## fatigue           0.03083585 1120.116862  0.001842698
## sexmale:fatigue -0.20616136 1064.029465 -0.012640389
```

```
modelsextask <- lmer(z.score ~ group + sex*task + (1|participant.id), data=data)
summary(modelsextask)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ group + sex * task + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2452.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5771 -0.6071  0.0025  0.6063  3.7168
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.117    0.3421
##  Residual                   0.499    0.7064
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                       Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)            0.94317    0.19488    8.18132   4.840  0.00121 **
## groupPD               -0.34448    0.27712    7.02747  -1.243  0.25371
## sexmale               -1.20221    0.28554    7.92175  -4.210  0.00302 **
## taskwordlist          -0.12026    0.07433 1114.05637  -1.618  0.10596
## sexmale:taskwordlist   0.17518    0.09521 1114.02602   1.840  0.06605 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) gropPD sexmal tskwrd
## groupPD     -0.354
## sexmale     -0.396 -0.568
## taskwordlst -0.272 -0.004  0.189
## sxml:tskwrd  0.212  0.006 -0.241 -0.781
```

```
lme.dscore(modelsextask, data, type="lme4")
```

```
##                                t          df           d
## groupPD              -1.243078    7.027467 -0.93784016
## sexmale              -4.210284    7.921754 -2.99178705
## taskwordlist         -1.617932 1114.056374 -0.09694749
## sexmale:taskwordlist  1.839894 1114.026021  0.11024913
```

```
# an effect sex and type (in interaction)
model7 <- lmer(z.score ~ group + sex*type + (1|participant.id), data = data)
summary(model7)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ group + sex * type + (1 | participant.id)
##    Data: data
##
## REML criterion at convergence: 2412.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8582 -0.5770  0.0207  0.5779  3.5960
##
## Random effects:
##  Groups         Name        Variance Std.Dev.
##  participant.id (Intercept) 0.1172   0.3424
##  Residual                   0.4777   0.6912
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                       Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)            1.12445    0.19347    7.94402   5.812 0.000410 ***
## groupPD               -0.34429    0.27714    7.02610  -1.242 0.253986
## sexmale               -1.51531    0.28385    7.73210  -5.338 0.000778 ***
## typeVAI               -0.37453    0.08174 1110.00797  -4.582 5.13e-06 ***
## typeMeanF0            -0.31036    0.09650 1110.01301  -3.216 0.001337 **
## typeMeanCPPS          -0.51216    0.09650 1110.01301  -5.307 1.34e-07 ***
## sexmale:typeVAI        0.61797    0.10500 1110.00797   5.886 5.25e-09 ***
## sexmale:typeMeanF0     0.50278    0.12436 1110.00985   4.043 5.65e-05 ***
## sexmale:typeMeanCPPS   0.83911    0.12436 1110.00985   6.747 2.42e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) gropPD sexmal typVAI typMF0 tMCPPS sx:VAI sx:MF0
## groupPD     -0.357
## sexmale     -0.391 -0.570
## typeVAI     -0.211  0.000  0.144
## typeMeanF0  -0.179 -0.001  0.123  0.424
## typeMenCPPS -0.179 -0.001  0.123  0.424  0.359
## sxml:typVAI  0.164  0.000 -0.185 -0.778 -0.330 -0.330
## sxml:typMF0  0.138  0.002 -0.157 -0.329 -0.776 -0.278  0.422
## sxml:tMCPPS  0.138  0.002 -0.157 -0.329 -0.278 -0.776  0.422  0.357
```

```
lme.dscore(model7, data, type="lme4")
```

```
##                              t          df          d
## groupPD             -1.242299    7.026103 -0.9373439
## sexmale             -5.338332    7.732096 -3.8396085
## typeVAI             -4.581987 1110.007967 -0.2750558
## typeMeanF0          -3.216016 1110.013007 -0.1930564
## typeMeanCPPS        -5.307120 1110.013007 -0.3185847
## sexmale:typeVAI      5.885669 1110.007967  0.3533156
## sexmale:typeMeanF0   4.042791 1110.009846  0.2426878
## sexmale:typeMeanCPPS 6.747219 1110.009846  0.4050339
```

```
anova(model2, model7) # type in interaction with sex is a significant and necessary predictor
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model2: z.score ~ sex + group + (1 | participant.id)
## model7: z.score ~ group + sex * type + (1 | participant.id)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model2  5 2453.7 2478.9 -1221.9   2443.7
## model7 11 2409.6 2464.9 -1193.8   2387.6 56.157      6  2.705e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model7a <- lmer(z.score ~ group + sex*type + (1 + type | participant.id), data = data)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(model7a)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ group + sex * type + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2026.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9075 -0.4989  0.0566  0.5940  3.3316
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1125   0.3355
##                 typeVAI      0.1062   0.3258    0.28
##                 typeMeanF0   0.4031   0.6349    0.17 -0.74
##                 typeMeanCPPS 0.6837   0.8269   -0.60  0.15  0.01
##  Residual                    0.3204   0.5660
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                       Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)            1.0714     0.1795  7.8041   5.969 0.000369 ***
## groupPD               -0.1138     0.1747 12.0815  -0.651 0.526988
## sexmale               -1.6517     0.2475  8.9845  -6.675 9.19e-05 ***
## typeVAI               -0.3752     0.1761  8.1507  -2.130 0.065135 .
## typeMeanF0            -0.3150     0.3271  7.9952  -0.963 0.363842
## typeMeanCPPS          -0.5168     0.4209  7.9986  -1.228 0.254465
## sexmale:typeVAI        0.6188     0.2272  8.1277   2.723 0.025730 *
## sexmale:typeMeanF0     0.5047     0.4223  7.9920   1.195 0.266250
## sexmale:typeMeanCPPS   0.8411     0.5434  7.9966   1.548 0.160272
## ---
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) gropPD sexmal typVAI typMF0 tMCPPS sx:VAI sx:MF0
## groupPD     -0.240
## sexmale     -0.584 -0.417
## typeVAI      0.172 -0.005 -0.122
## typeMeanF0   0.115 -0.002 -0.082 -0.629
## typeMenCPPS -0.581 -0.002  0.423  0.169  0.030
## sxml:typVAI -0.134  0.008  0.157 -0.775  0.488 -0.131
## sxml:typMF0 -0.089  0.003  0.106  0.488 -0.775 -0.023 -0.630
## sxml:tMCPPS  0.450  0.002 -0.546 -0.131 -0.023 -0.775  0.169  0.030
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
#compare the models with and without slopes
anova(model7, model7a, refit=F) # the model with (1 + type | participant.id) is significantly better
```

```
## Data: data
## Models:
## model7: z.score ~ group + sex * type + (1 | participant.id)
## model7a: z.score ~ group + sex * type + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model7  11 2434.4 2489.7 -1206.2   2412.4
## model7a 20 2066.4 2167.0 -1013.2   2026.5 385.99      9  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We drop sex*type as predictor and re-run the hypothesis test.

```
model10 <- lmer(z.score ~ group + (1 + type | participant.id), data = data)
```

```
summary(model10)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ group + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2045.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8929 -0.5034  0.0572  0.5876  3.3217
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.5427   0.7367
##                 typeVAI      0.1720   0.4148   -0.54
##                 typeMeanF0   0.3773   0.6142   -0.36 -0.20
##                 typeMeanCPPS 0.7120   0.8438   -0.57  0.45  0.21
##  Residual                    0.3206   0.5662
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   0.4492     0.2515  8.0714   1.786   0.1116
## groupPD      -0.7480     0.3247  8.0666  -2.304   0.0499 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr)
## groupPD -0.775
## convergence code: 0
## Model failed to converge with max|grad| = 0.0151918 (tol = 0.002, component 1)
```

```
lme.dscore(model10, data, type="lme4")
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0151918
## (tol = 0.002, component 1)
```

```
##                t     df         d
## groupPD -2.304043 8.0666 -1.622465
```

here is still a tendency for lower values in the PD group and it remains statistically significant (p = 0.05, d = - 1.6).

```
model11 <- lmer(z.score ~ sex + (1 + type | participant.id), data = data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00464743
## (tol = 0.002, component 1)
```

```
summary(model11)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2031.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9297 -0.5156  0.0600  0.5839  3.3457
```

```
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1313   0.3624
##                 typeVAI      0.1722   0.4149    0.52
##                 typeMeanF0   0.3783   0.6151    0.36 -0.19
##                 typeMeanCPPS 0.7127   0.8442   -0.22  0.45  0.21
##  Residual                    0.3205   0.5661
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   1.2556     0.1027  8.1296   12.22 1.62e-06 ***
## sexmale      -2.0670     0.1321  8.0046  -15.65 2.76e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr)
## sexmale -0.778
## convergence code: 0
## Model failed to converge with max|grad| = 0.00464743 (tol = 0.002, component 1)
```

```
lme.dscore(model11, data, type="lme4")
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00464743
## (tol = 0.002, component 1)
```

```
##                t       df        d
## sexmale -15.64658 8.004628 -11.0606
```

Model with sex is very significant ($p < 0.001$, $d = -11$)

```
anova(model10, model11)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model10: z.score ~ group + (1 + type | participant.id)
## model11: z.score ~ sex + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model10 13 2069.4 2134.8 -1021.7   2043.4
## model11 13 2051.4 2116.8 -1012.7   2025.4 18.024      0  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It turns out that sex (model 11) is a better predictor of our results than PD ($p < 0.001$, AIC is 18 points lower. We nonetheless keep group as a predictor alongside sex, due to our hypothesis.

```
model12 <- lmer(z.score ~ sex + group + (1 + type | participant.id), data=data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0204948
## (tol = 0.002, component 1)
```

```
summary(model12)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + group + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2032.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9402 -0.5116  0.0584  0.5813  3.3433
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1212   0.3482
##                 typeVAI      0.1718   0.4145    0.51
##                 typeMeanF0   0.3785   0.6152    0.32 -0.20
##                 typeMeanCPPS 0.7107   0.8431   -0.23  0.45  0.21
##  Residual                    0.3206   0.5662
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)   1.2620     0.1145  7.0573  11.025 1.06e-05 ***
## sexmale      -1.9690     0.1714  7.4621 -11.485 5.21e-06 ***
## groupPD      -0.1088     0.1714  7.4621  -0.635    0.544
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) sexmal
## sexmale -0.363
## groupPD -0.363 -0.599
## convergence code: 0
## Model failed to converge with max|grad| = 0.0204948 (tol = 0.002, component 1)
```

```
lme.dscore(model12, data, type="lme4")
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0204948
## (tol = 0.002, component 1)
```

```
##                  t       df          d
## sexmale -11.4849159 7.462066 -8.4086886
## groupPD  -0.6348429 7.462066 -0.4648006
```

```
anova(model11, model12) # adding group does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model11: z.score ~ sex + (1 + type | participant.id)
## model12: z.score ~ sex + group + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model11 13 2051.4 2116.8 -1012.7   2025.4
## model12 14 2052.9 2123.3 -1012.5   2024.9 0.5081      1      0.476
```

```
model13 <- lmer(z.score ~ fatigue + sex + group + (1 + type | participant.id), data=data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0130574
## (tol = 0.002, component 1)
```

```
summary(model13)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ fatigue + sex + group + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2039.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9342 -0.5075  0.0547  0.5834  3.3230
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1256   0.3544
##                 typeVAI      0.1717   0.4143    0.51
##                 typeMeanF0   0.3791   0.6157    0.34 -0.20
##                 typeMeanCPPS 0.7124   0.8440   -0.23  0.45  0.21
##  Residual                    0.3208   0.5664
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##              Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)   1.286198   0.115516   7.681551  11.134 5.20e-06 ***
## fatigue      -0.007942   0.011870 942.035628  -0.669    0.504
## sexmale      -1.981609   0.169301   7.482236 -11.705 4.46e-06 ***
## groupPD      -0.102736   0.168961   7.422643  -0.608    0.561
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##         (Intr) fatigu sexmal
## fatigue -0.220
## sexmale -0.338 -0.066
## groupPD -0.358  0.020 -0.600
## convergence code: 0
## Model failed to converge with max|grad| = 0.0130574 (tol = 0.002, component 1)
```

```
anova(model12, model13) # adding fatigue does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model12: z.score ~ sex + group + (1 + type | participant.id)
## model13: z.score ~ fatigue + sex + group + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model12 14 2052.9 2123.3 -1012.5   2024.9
## model13 15 2054.4 2129.8 -1012.2   2024.4 0.4765      1       0.49
```

```
model14 <- lmer(z.score ~ day + group + sex + (1 + type | participant.id), data=data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00518859
## (tol = 0.002, component 1)
```

```
summary(model14)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ day + group + sex + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2041
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9685 -0.5045  0.0668  0.5818  3.3318
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1201   0.3466
##                 typeVAI      0.1717   0.4144    0.50
##                 typeMeanF0   0.3787   0.6154    0.32 -0.20
##                 typeMeanCPPS 0.7126   0.8442   -0.24  0.45  0.21
##  Residual                    0.3199   0.5656
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)    1.30110    0.11886    8.17085  10.947 3.65e-06 ***
## daysecond     -0.02725    0.04858 1083.59042  -0.561   0.5749
## daythird      -0.03316    0.04810 1083.54531  -0.689   0.4908
## dayfourth     -0.10446    0.04816 1083.43200  -2.169   0.0303 *
## groupPD       -0.11128    0.17202    7.52894  -0.647   0.5369
## sexmale       -1.96037    0.17203    7.52956 -11.396 5.14e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) dyscnd dythrd dyfrth gropPD
## daysecond -0.209
## daythird  -0.211  0.521
## dayfourth -0.214  0.520  0.525
## groupPD   -0.351  0.000  0.002  0.001
## sexmale   -0.349 -0.004 -0.007 -0.003 -0.599
## convergence code: 0
## Model failed to converge with max|grad| = 0.00518859 (tol = 0.002, component 1)
```

```
anova(model12, model14) # adding day does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model12: z.score ~ sex + group + (1 + type | participant.id)
## model14: z.score ~ day + group + sex + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model12 14 2052.9 2123.3 -1012.5   2024.9
## model14 17 2053.7 2139.1 -1009.8   2019.7 5.2515      3     0.1543
```

```
model15 <- lmer(z.score ~ sex + group + time + (1 + type | participant.id), data=data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00468732
## (tol = 0.002, component 1)
```

```
summary(model15)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + group + time + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2034.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8983 -0.4997  0.0762  0.5753  3.2780
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1212   0.3481
##                 typeVAI      0.1719   0.4146    0.50
##                 typeMeanF0   0.3772   0.6142    0.32 -0.20
##                 typeMeanCPPS 0.7137   0.8448   -0.24  0.45  0.21
##  Residual                    0.3197   0.5655
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)    1.23267    0.11602    7.34388  10.625 1.02e-05 ***
## sexmale       -1.96280    0.17241    7.52180 -11.384 5.22e-06 ***
## groupPD       -0.11130    0.17241    7.52185  -0.646   0.5378
## timeevening    0.06552    0.03431 1085.38708   1.909   0.0565 .
## ---
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sexmal gropPD
## sexmale     -0.361
## groupPD     -0.359 -0.599
## timeevening -0.122  0.006 -0.006
## convergence code: 0
## Model failed to converge with max|grad| = 0.00468732 (tol = 0.002, component 1)
```

```
lme.dscore(model15, data, type="lme4")
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00468732
## (tol = 0.002, component 1)
```

```
##                       t          df          d
## sexmale     -11.3844200    7.521797 -8.3019499
## groupPD      -0.6455513    7.521847 -0.4707589
## timeevening   1.9093875 1085.387083  0.1159129
```

```
anova(model12, model15) # adding time does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model12: z.score ~ sex + group + (1 + type | participant.id)
## model15: z.score ~ sex + group + time + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model12 14 2052.9 2123.3 -1012.5   2024.9
## model15 15 2051.3 2126.7 -1010.6   2021.3 3.6282      1    0.05681 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model16 <- lmer(z.score ~ task + sex + group + (1 + type|participant.id), data=data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00731097
## (tol = 0.002, component 1)
```

```
summary(model16)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ task + sex + group + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2037
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9604 -0.5208  0.0588  0.5862  3.3425
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1211   0.3480
##                 typeVAI      0.1718   0.4145    0.50
##                 typeMeanF0   0.3795   0.6161    0.32 -0.20
##                 typeMeanCPPS 0.7135   0.8447   -0.24  0.45  0.21
##  Residual                    0.3207   0.5663
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     1.27293    0.11718    7.75026  10.863 5.81e-06 ***
## taskwordlist   -0.02277    0.04205 1094.66765  -0.542    0.588
## sexmale        -1.96544    0.17187    7.53578 -11.436 4.98e-06 ***
## groupPD        -0.10972    0.17187    7.53578  -0.638    0.542
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) tskwrd sexmal
## taskwordlst -0.202
## sexmale     -0.356  0.002
## groupPD     -0.356  0.002 -0.599
## convergence code: 0
## Model failed to converge with max|grad| = 0.00731097 (tol = 0.002, component 1)
```

```
lme.dscore(model16, data, type="lme4")
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00731097
## (tol = 0.002, component 1)
```

```
##                         t         df          d
## taskwordlist  -0.5416215 1094.66765 -0.03274046
## sexmale      -11.4357963    7.53578 -8.33167443
## groupPD       -0.6383974    7.53578 -0.46511141
```

```
anova(model12, model16) # adding task does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: data
## Models:
## model12: z.score ~ sex + group + (1 + type | participant.id)
## model16: z.score ~ task + sex + group + (1 + type | participant.id)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## model12 14 2052.9 2123.3 -1012.5   2024.9
## model16 15 2054.6 2130.0 -1012.3   2024.6 0.2929      1     0.5883
```

```
# (re-)check for interactions in sex
model17 <- lmer(z.score ~ sex*task + group + (1 + type|participant.id), data=data)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00839925
## (tol = 0.002, component 1)
```

```
summary(model17)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex * task + group + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2040.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9547 -0.5169  0.0585  0.5848  3.3410
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1209   0.3477
##                 typeVAI      0.1718   0.4145    0.50
##                 typeMeanF0   0.3791   0.6157    0.32 -0.20
##                 typeMeanCPPS 0.7103   0.8428   -0.24  0.45  0.21
##  Residual                    0.3210   0.5666
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                          Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)           1.276e+00  1.208e-01  8.716e+00  10.559 2.93e-06 ***
## sexmale              -1.970e+00  1.789e-01  8.819e+00 -11.013 1.88e-06 ***
## taskwordlist         -2.863e-02  6.714e-02  1.094e+03  -0.426    0.670
## groupPD              -1.098e-01  1.721e-01  7.552e+00  -0.638    0.543
## sexmale:taskwordlist  9.644e-03  8.615e-02  1.093e+03   0.112    0.911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sexmal tskwrd gropPD
## sexmale     -0.398
## taskwordlst -0.309  0.214
## groupPD     -0.343 -0.579 -0.007
## sxml:tskwrd  0.239 -0.273 -0.779  0.011
## convergence code: 0
## Model failed to converge with max|grad| = 0.00839925 (tol = 0.002, component 1)
```

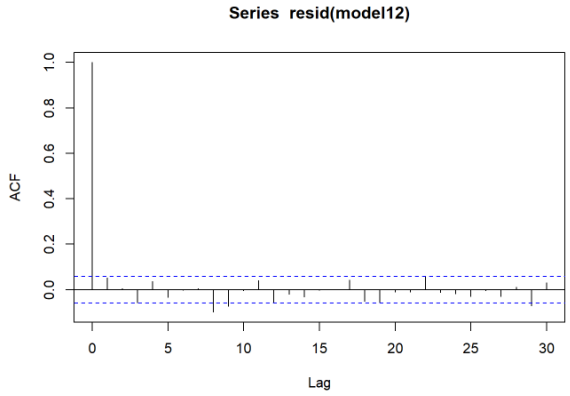
```
model18 <- lmer(z.score ~ sex*fatigue + group + (1 + type|participant.id), data=data)
summary(model18)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex * fatigue + group + (1 + type | participant.id)
##    Data: data
##
## REML criterion at convergence: 2044.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.9294 -0.5110  0.0545  0.5779  3.3084
##
## Random effects:
##  Groups         Name         Variance Std.Dev. Corr
##  participant.id (Intercept)  0.1281   0.3580
##                 typeVAI      0.1717   0.4144    0.51
##                 typeMeanF0   0.3784   0.6151    0.34 -0.20
##                 typeMeanCPPS 0.7129   0.8443   -0.23  0.45  0.21
##  Residual                    0.3210   0.5666
## Number of obs: 1126, groups:  participant.id, 10
##
## Fixed effects:
##                    Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)       1.279e+00  1.178e-01  8.436e+00  10.856 3.03e-06 ***
## sexmale          -1.966e+00  1.780e-01  9.157e+00 -11.041 1.35e-06 ***
## fatigue          -3.826e-03  1.595e-02  1.091e+03  -0.240    0.810
## groupPD          -9.637e-02  1.689e-01  7.509e+00  -0.571    0.585
## sexmale:fatigue  -9.458e-03  2.391e-02  8.890e+02  -0.396    0.693
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) sexmal fatigu gropPD
## sexmale     -0.377
## fatigue     -0.296  0.162
## groupPD     -0.358 -0.557  0.043
## sexmale:ftg  0.202 -0.313 -0.668 -0.041
```

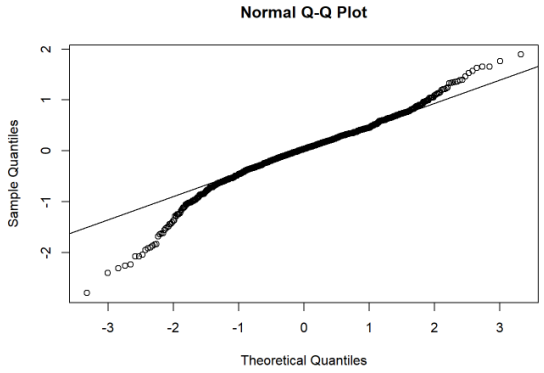
```
# check assumptions (multicollinearity)
vif(model12)
```

```
##      sex    group
## 1.559604 1.559604
```

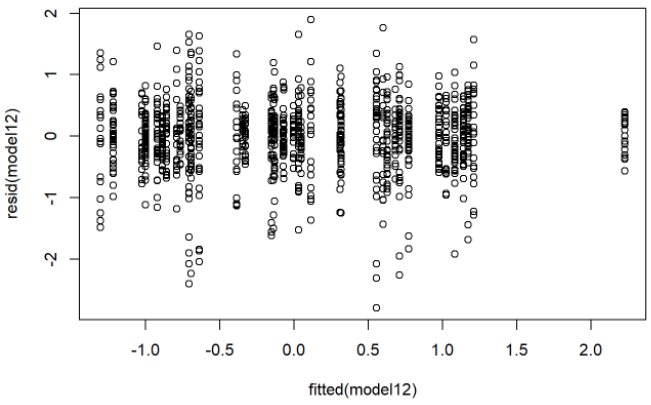
```
# check assumptions (autocorrelation)
acf(resid(model12))
```




```
#check assumptions (normality)
qqnorm(resid(model12))
qqline(resid(model12))
```




```
# check assumptions (heteroscedasticity)
plot(fitted(model12), resid(model12))
```



```
#testing the null hypothesis that there is no effect of medication on the speech of PD patients (ON/OFF state); with state
 as fixed effect and random intercepts for subjects

modelPD1 <- lmer(z.score ~ state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1900.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7166 -0.6383  0.0580  0.6413  2.8607
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.224e-32 2.688e-16
##  Residual                   9.495e-01 9.744e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)  -0.03917    0.08418 678.00000  -0.465    0.642
## stateON       0.04879    0.09394 678.00000   0.519    0.604
##
## Correlation of Fixed Effects:
##         (Intr)
## stateON -0.896
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD1, data=dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                t  df         d
## stateON 0.5193332 678 0.03988974
```

```
#first, we test whether it is better to use "effect" (OFF state, +1 hour, +2 hours) as opposed to "state" (only ON/OFF)
modelPD2 <- lmer(z.score ~ effect + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ effect + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1903.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7043 -0.6383  0.0633  0.6404  2.8587
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.234e-32 2.690e-16
##  Residual                   9.508e-01 9.751e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)  -0.03917    0.08424 677.00000  -0.465    0.642
## effect+1h     0.03929    0.10255 677.00000   0.383    0.702
## effect+2h     0.05863    0.10317 677.00000   0.568    0.570
##
## Correlation of Fixed Effects:
##             (Intr) effc+1
## effect+1h -0.821
## effect+2h -0.816  0.671
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD2, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                   t  df          d
## effect+1h 0.3831664 677 0.02945256
## effect+2h 0.5683407 677 0.04368622
```

```
anova(modelPD1, modelPD2) # the model with "effect" is not significantly different so we continue with "state" for the res
t of our analysis
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD2: z.score ~ effect + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD2  5 1902.4 1925.0 -946.22   1892.4 0.0539      1     0.8164
```

```
modelPDtask <- lmer(z.score ~ task + state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPDtask)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ task + state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1903.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7140 -0.6379  0.0580  0.6407  2.8585
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.235e-32 2.690e-16
##  Residual                   9.509e-01 9.751e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)  -3.952e-02  1.027e-01  6.770e+02  -0.385    0.701
## taskwordlist  4.879e-04  8.207e-02  6.770e+02   0.006    0.995
## stateON       4.879e-02  9.401e-02  6.770e+02   0.519    0.604
##
## Correlation of Fixed Effects:
##             (Intr) tskwrd
## taskwordlst -0.572
## stateON     -0.741  0.011
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPDtask, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                        t  df            d
## taskwordlist 0.005944851 677 0.0004569583
## stateON      0.518984102 677 0.0398923587
```

```
anova(modelPD1, modelPDtask) # adding task does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPDtask: z.score ~ task + state + (1 | participant.id)
##             Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## modelPD1     4 1900.5 1918.6 -946.25   1892.5
## modelPDtask  5 1902.5 1925.1 -946.25   1892.5     0      1     0.9952
```

```
modelPD3 <- lmer(z.score ~ sex + state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD3)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ sex + state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1902.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7140 -0.6378  0.0578  0.6407  2.8585
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.235e-32 2.690e-16
##  Residual                   9.509e-01 9.751e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept) -3.976e-02  1.211e-01  6.770e+02  -0.328    0.743
## sexmale      6.888e-04  1.023e-01  6.770e+02   0.007    0.995
## stateON      4.879e-02  9.402e-02  6.770e+02   0.519    0.604
##
```

```
##
## Correlation of Fixed Effects:
##         (Intr) sexmal
## sexmale -0.719
## stateON -0.632  0.013
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD3, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                   t  df           d
## sexmale 0.006731626 677 0.0005174348
## stateON 0.518993714 677 0.0398930976
```

```
anova(modelPD1, modelPD3) #adding sex does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD3: z.score ~ sex + state + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD3  5 1902.5 1925.1 -946.25   1892.5     0      1     0.9946
```

```
modelPD4 <- lmer(z.score ~ fatigue + state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD4)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ fatigue + state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1906
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7135 -0.6377  0.0577  0.6412  2.8588
##
```

```
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.235e-32 2.690e-16
##  Residual                   9.509e-01 9.751e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##              Estimate Std. Error         df t value Pr(>|t|)
## (Intercept) -3.867e-02  1.048e-01  6.770e+02  -0.369    0.712
## fatigue     -1.677e-04  2.090e-02  6.770e+02  -0.008    0.994
## stateON      4.872e-02  9.441e-02  6.770e+02   0.516    0.606
##
## Correlation of Fixed Effects:
##         (Intr) fatigu
## fatigue -0.595
## stateON -0.772  0.092
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD4, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                    t  df             d
## fatigue -0.008023182 677 -0.0006167119
## stateON  0.515988726 677  0.0396621155
```

```
anova(modelPD1, modelPD4) #adding fatigue does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD4: z.score ~ fatigue + state + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD4  5 1902.5 1925.1 -946.25   1892.5 1e-04      1     0.9936
```

```
modelPD5 <- lmer(z.score ~ type + state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD5)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ type + state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1908.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7083 -0.6370  0.0578  0.6397  2.8545
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.256e-32 2.694e-16
##  Residual                   9.537e-01 9.766e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept) -3.903e-02  1.167e-01  6.750e+02  -0.334    0.738
## typeMeanF0   3.978e-16  1.261e-01  6.750e+02   0.000    1.000
## typeVAI     -2.218e-04  1.108e-01  6.750e+02  -0.002    0.998
## typetVSA    -2.218e-04  1.108e-01  6.750e+02  -0.002    0.998
## stateON      4.879e-02  9.415e-02  6.750e+02   0.518    0.604
##
## Correlation of Fixed Effects:
##             (Intr) typMF0 typVAI typVSA
## typeMeanF0 -0.540
## typeVAI    -0.612  0.569
## typetVSA   -0.612  0.569  0.647
## stateON    -0.645  0.000 -0.004 -0.004
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD5, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                       t  df            d
## typeMeanF0  3.155086e-15 675  2.428786e-16
## typeVAI    -2.000960e-03 675 -1.540340e-04
## typetVSA   -2.000960e-03 675 -1.540340e-04
## stateON     5.181907e-01 675  3.989034e-02
```

```
anova(modelPD1, modelPD5) #adding type does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD5: z.score ~ type + state + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD5  7 1906.5 1938.2 -946.25   1892.5     0      3          1
```

```
modelPD6 <- lmer(z.score ~ day + state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD6)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ day + state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1904.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7475 -0.6573  0.0562  0.6657  2.8200
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.214e-32 2.686e-16
##  Residual                   9.482e-01 9.738e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##             Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)  0.01695    0.10858 675.00000   0.156   0.8760
## daysecond   -0.03152    0.10762 675.00000  -0.293   0.7697
## daythird    -0.01458    0.10644 675.00000  -0.137   0.8911
## dayfourth   -0.18316    0.10701 675.00000  -1.712   0.0874 .
## stateON      0.05182    0.09391 675.00000   0.552   0.5812
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) dyscnd dythrd dyfrth
## daysecond -0.527
## daythird  -0.527  0.524
## dayfourth -0.513  0.521  0.527
## stateON   -0.701  0.019  0.011 -0.005
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD6, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                     t  df           d
## daysecond -0.2928514 675 -0.02254371
## daythird  -0.1369641 675 -0.01054350
## dayfourth -1.7115831 675 -0.13175773
## stateON    0.5518371 675  0.04248044
```

```
anova(modelPD1, modelPD6) # adding day does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD6: z.score ~ day + state + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD6  7 1902.6 1934.2 -944.28   1888.6 3.931      3      0.269
```

```
modelPD7 <- lmer(z.score ~ time + state + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD7)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ time + state + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1902
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6678 -0.6357  0.0566  0.6244  2.8613
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.221e-32 2.687e-16
##  Residual                   9.491e-01 9.742e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##               Estimate Std. Error         df t value Pr(>|t|)
## (Intercept) -3.917e-02  8.416e-02  6.770e+02  -0.465    0.642
## timeevening  9.423e-02  8.341e-02  6.770e+02   1.130    0.259
## stateON      4.618e-04  1.032e-01  6.770e+02   0.004    0.996
##
## Correlation of Fixed Effects:
##             (Intr) tmvnng
## timeevening  0.000
## stateON     -0.815 -0.414
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD7, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                       t  df           d
## timeevening 1.129711993 677 0.0868367179
## stateON     0.004474792 677 0.0003439604
```

```
anova(modelPD1, modelPD7) # adding time of day does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD7: z.score ~ time + state + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD7  5 1901.2 1923.8 -945.61   1891.2 1.2807      1     0.2578
```

```
modelPD8 <- lmer(z.score ~ state*fatigue + (1|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD8)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ state * fatigue + (1 | participant.id)
##    Data: dataPD
##
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ state * fatigue + (1 | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1903.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7754 -0.6343  0.0536  0.6305  3.0073
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.
##  participant.id (Intercept) 7.182e-32 2.680e-16
##  Residual                   9.440e-01 9.716e-01
## Number of obs: 680, groups:  participant.id, 6
##
## Fixed effects:
##                  Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)       0.35618    0.19315 676.00000   1.844   0.0656 .
## stateON          -0.39588    0.20571 676.00000  -1.924   0.0547 .
## fatigue          -0.13244    0.05828 676.00000  -2.273   0.0234 *
## stateON:fatigue   0.15165    0.06240 676.00000   2.430   0.0153 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) statON fatigu
## stateON     -0.939
## fatigue     -0.901  0.846
## stateON:ftg  0.841 -0.889 -0.934
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
lme.dscore(modelPD8, dataPD, type="lme4")
```

```
## boundary (singular) fit: see ?isSingular
```

```
##                         t  df          d
## stateON         -1.924485 676 -0.1480373
## fatigue         -2.272669 676 -0.1748207
## stateON:fatigue  2.430325 676  0.1869481
```

```
anova(modelPD1, modelPD8) # adding fatigue does not significantly improve the model
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD8: z.score ~ state * fatigue + (1 | participant.id)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## modelPD1  4 1900.5 1918.6 -946.25   1892.5
## modelPD8  6 1898.6 1925.7 -943.29   1886.6 5.9157      2    0.05193 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# due to our first hypothesis analysis, we check whether model including "type" as correlated random slope is better
modelPD1a <- lmer(z.score ~ state + (1+type|participant.id), data=dataPD)
```

```
## boundary (singular) fit: see ?isSingular
```

```
summary(modelPD1a)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: z.score ~ state + (1 + type | participant.id)
##    Data: dataPD
##
## REML criterion at convergence: 1900.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7166 -0.6383  0.0580  0.6413  2.8607
##
## Random effects:
##  Groups         Name        Variance  Std.Dev.  Corr
##  participant.id (Intercept) 0.000e+00 0.0000000
##                 typeMeanF0  8.830e-08 0.0002972   NaN
##                 typeVAI     9.119e-08 0.0003020   NaN  0.67
##                 typetVSA    2.129e-08 0.0001459   NaN -0.12 -0.56
##  Residual                   9.495e-01 0.9744100
## Number of obs: 680, groups:  participant.id, 6
##
```

```
##
## Fixed effects:
##              Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)  -0.03917    0.08418 677.98934  -0.465    0.642
## stateON       0.04879    0.09394 678.00000   0.519    0.604
##
## Correlation of Fixed Effects:
##         (Intr)
## stateON -0.896
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
anova(modelPD1, modelPD1a) # model including "type" is not better
```

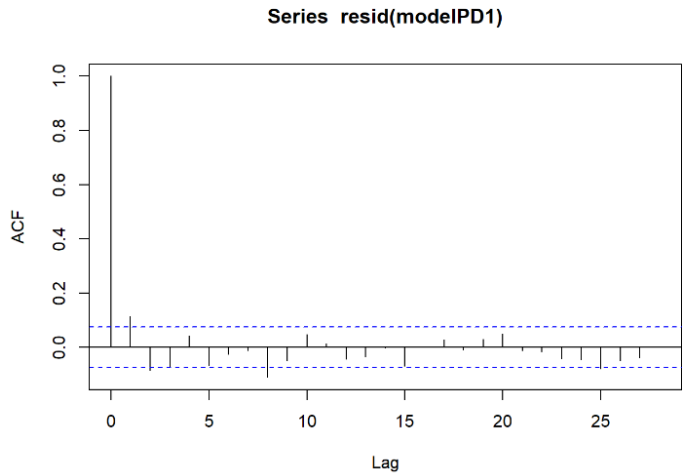
```
## refitting model(s) with ML (instead of REML)
```

```
## Data: dataPD
## Models:
## modelPD1: z.score ~ state + (1 | participant.id)
## modelPD1a: z.score ~ state + (1 + type | participant.id)
##           Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## modelPD1   4 1900.5 1918.6 -946.25   1892.5
## modelPD1a 13 1918.5 1977.3 -946.25   1892.5     0      9          1
```

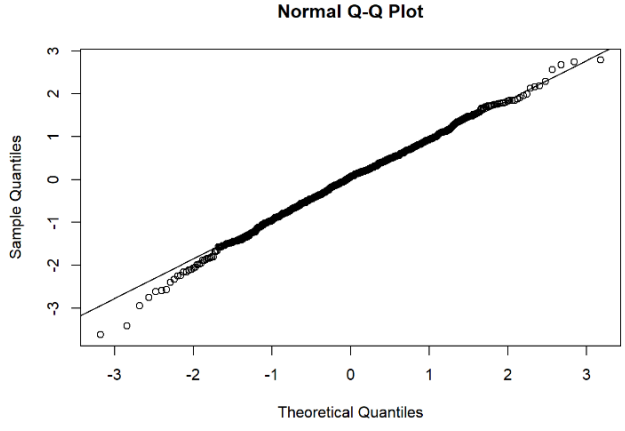
There are no models that would significantly better predict our results for PD patients.

```
# check assumptions (multicollinearity);
# cannot check for multicollinearity because there is only 1 term
```

```
# check assumptions (autocorrelation)
acf(resid(modelPD1))
```




```
#check assumptions (normality)
qqnorm(resid(modelPD1))
qqline(resid(modelPD1))
```



```
# check assumptions (heteroscedasticity)
plot(fitted(modelPD1), resid(modelPD1))
```