# Digital Eavesdropper

*Acoustic speech characteristics as markers of exacerbations in COPD patients*

| | |
|---|---|
| **Name student** | Julia Merkus |
| **Student number** | 4462661 |
| **Study program** | MA Language and Speech Pathology |
| **Institution** | Radboud University, Nijmegen |
| **Document** | Master thesis |
| **Supervisor** | dr. W.A.J. Strik |
| **Second reader** | dr. E. Janse |
| **Title:** | Digital Eavesdropper. Acoustic speech characteristics as markers of exacerbations in COPD patients |
| **Date:** | August 14, 2019 |

# Preface

Before you lies the master thesis "Digital Eavesdropper: acoustic speech characteristics as markers of exacerbations in COPD patients". It has been written to fulfill the graduation requirements of the Language and Speech Pathology program at the Radboud University in Nijmegen. I was engaged in researching and writing this thesis from February to August 2019.

I noticed during my previous master thesis that I avoided stepping outside my comfort zone. This year I wanted to approach the thesis differently by choosing a subject which required skills I did not have yet. I have worked with unfamiliar computer clusters, operating systems and FTP applications, such as Ponyland, Linux and FileZilla, and I have gained more experience with programs I was already familiar with. Also, I have learned that struggling is part of the process. Therefore, this thesis has taught me valuable lessons both professionally and personally.

I would like to thank my supervisor, dr. Helmer Strik, for the excellent guidance and support during the process. I deliberately chose you to be my supervisor, because I knew you would provide me with challenges. This has maximized the learning opportunities, for which I am grateful. I would also like to thank dr. Esther Janse: not only for reading my thesis, but also for the interesting courses. The master program would not be the same without you. Thirdly, I want to thank dr. Hanneke van Helvoort from the department of Pulmonary Diseases of the Radboud University Medical Center for her contribution to the data collection for this study. Next, I want to express my gratitude towards Wei Xue, Msc. I was clueless at the start, but you have mentored me with great patience and kindness. Moreover, I would like to thank Mario Ganzeboom, Msc, Ferdy Hubers, MA, and Maarten Vos, Bsc, for providing me with scripts to align, analyze and transform my data.

Finally, I want to thank my (step) parents, brother, Marije, Gaby and Emma for being there for me.

I hope you enjoy your reading.

Julia Merkus

Nijmegen, August 14, 2019

## Abstract

Research suggests that speech deterioration indicates an exacerbation in patients with chronic obstructive pulmonary disease (COPD). This study provides a comparison of read speech of 9 stable COPD patients and 5 healthy controls (I) and 9 stable COPD patients and 9 COPD patients in exacerbation (II). Results showed a significant effect of condition on the number of (non-linguistic) in- and exhalations per syllable (I, II) and the ratio of voiced and silence intervals (II). Also, sustained vowels by 10 COPD patients in exacerbation were compared with 10 vowels in stable condition (III). Results showed an effect of condition on duration, shimmer, harmonics-to-noise ratio (HNR) and voice breaks. It was concluded that HNR, vowel duration and the number of (non-linguistic) in- and exhalations per syllable show potential for remote monitoring. Further research is needed to examine the validity of the results for natural speech and larger sample sizes.

# Contents

# 1. Introduction

## 1.1 Speech and speech disorders

Each human language consists of a set of vowels and consonants which are combined to form words. During the speech production process, thoughts are converted into spoken utterances to convey a message. The appropriate words and their meanings are selected in the mental lexicon (Dell & Burger, 1997). This pre-verbal message is then grammatically encoded, during which a syntactic representation of the utterance is built. The sounds are yet to be specified, but the abstract word symbols are assigned to their grammatical function before they are structured in a syntactic frame to determine the order (Cho-Reyes, Mack, & Thompson, 2016). Subsequently, the message is phonologically encoded. During this stage, a phonetic or articulatory plan is retrieved for each individual lemma and the utterance as a whole. Finally, the speaker produces the utterance according to the phonetic plan (Levelt, 2002).

Speech, language and voice disorders, such as apraxia, aphasia and spasmodic dysphonia, affect the vocal cords, nerves, muscles and brain structures, which results in a distorted language reception or speech production (Sataloff & Hawkshaw, 2014). The symptoms vary from adding superfluous words and taking pauses to hoarseness of the voice, depending on the type of disorder (Dodd, 2005). However, distortions of the speech may also occur as a result of a disease that seems unrelated to speech, such as multiple sclerosis (which limits articulatory movements and respiratory functions) or chronic obstructive pulmonary disease (which limits respiratory functions).

This study aims to determine which acoustic parameters are suitable for the automatic detection of exacerbations in patients suffering from chronic obstructive pulmonary disease (COPD) by investigating which aspects of speech differ between COPD patients and healthy speakers and which aspects differ between COPD patients in exacerbation and stable COPD patients.

## 1.2 Chronic obstructive pulmonary disease

### 1.2.1 Background

COPD is an umbrella term used to describe progressive lung diseases characterized by airflow limitation. According to the guidelines provided by the Global Initiative for Chronic Obstructive Lung Disease (GOLD, 2019), the official definition of COPD is "a common, preventable and treatable disease that is characterized by persistent respiratory symptoms and airflow limitation that is due to airway and/or alveolar abnormalities usually caused by significant exposure to noxious particles or gases" (p. 2).

### 1.2.2 Prevalence

The prevalence of COPD worldwide is estimated at roughly 12%, but the percentage differs greatly between different subgroups (Lopez et al., 2014). Most COPD patients are suffering from stage II COPD (70%), while stage I, III and IV make up respectively 16%, 11% and 3% of the COPD population. The four greatest predictors of COPD are years and intensity of smoking, age, sex and BMI. Most patients suffering from COPD are smokers with a low BMI, over 50 years old and male (Lopez et al., 2014).

Taking into account the three million annual deaths globally, COPD is currently the fourth leading cause of death in high-income countries and it is expected to be the third leading cause in 2020 due to a higher life expectancy and increasing air pollution (Buist et al., 2007; GOLD, 2019; Postma, Bush, & Van den Berge, 2015). However, the lung disease has been overlooked and neglected for a long time by both the public and the pharmaceutical industry. This neglect might be caused in part by the assumption that COPD is a self-inflicted health condition caused by smoking. Although smoking is the leading cause of COPD in high-income countries, over 15% of the patients are nonsmokers (Buist et al., 2007). These nonsmokers often develop COPD as a result of job or living circumstances that require them to be exposed to polluted air. COPD is associated with an economic burden, since the disease accounts for approximately 55% of the costs for respiratory diseases in Europe (GOLD, 2019).

### 1.2.3 Causes

COPD is the result of long-term exposure to noxious gases and particles, often influenced by host factors, such as genetics, poor lung growth and hyper-responsiveness (GOLD, 2019). The chronic airflow limitation in COPD is caused by a combination of small airways disease and parenchymal destruction or emphysema (Cosio-Piqueras & Cosio, 2001). Chronic inflammation causes narrowing of the small airways and destroys the lung parenchyma. As a result, the alveolar attachments to the small airways weaken or disappear and the lung elastic recoil decreases (GOLD, 2019). This disables the airways to remain fully open during expiration (Barnes, 2004).

The understanding of risk factors for COPD is still incomplete and requires further investigation (GOLD, 2019). COPD seems to be associated with a severe hereditary deficiency of an inhibitor of serine proteases, namely alpha-1 antitrypsin, and a familial risk of airflow limitation has been observed (GOLD, 2019). It remains unknown whether these genetic factors are directly responsible for COPD or markers of causal genes (Cho et al., 2010; GOLD, 2019).

### 1.2.4 Diagnosis

The Standards for the Diagnosis and Treatment of Patients with COPD document (Celli et al., 2004) is an updated version of the position papers on COPD. Both professionals and patients requested an update of the documents for the following reasons:

1. The prevalence of COPD is increasing;
2. There have been discoveries in the field that could increase the quality of life for patients;
3. There is a need for an online document to ensure its accessibility;
4. The care of COPD is now viewed as multidisciplinary;
5. There are new insights regarding the habit of smoking and its effect on COPD.

COPD is diagnosed by measuring the extent of airflow limitation. In most cases, spirometry is used to measure the lung function (GOLD, 2019). Spirometry is more reliable than clinical descriptions, such as ‘patient produces sputum for at least 3 months in 2 consecutive years’, since COPD is a heterogeneous disease in its clinical expression (Postma et al., 2015). Individuals suffering from COPD generally show a variety of symptoms, commonly including shortness of breath, tightness on the chess and coughing (with mucus).

Stable COPD is interrupted by episodes or exacerbations during which the respiratory symptoms acutely worsen (GOLD, 2019). These exacerbations are sometimes triggered by respiratory infections, but the cause or trigger of the exacerbation often remains unknown (Rutschmann et al., 2007). During exacerbations, the peripheral airway limitation causes gas to get trapped during expiration. This leads to hyperinflation which is associated with a limited inspiratory capacity and increased dyspnoea. As the disease progresses, the gas transfer for oxygen and carbon dioxide worsens, resulting in hypoxemia and hypercapnia. In addition, the submucosal glands enlarge because of the chronic airway irritation, leading to mucus hypersecretion (GOLD, 2019).

The morbidity increases with age because of the high rate of comorbidity in COPD patients. Other concomitant chronic conditions, such as diabetes and cardiovascular diseases, often interfere with COPD management and might compromise the patient's health even more. This causes COPD patients to be hospitalized frequently (GOLD, 2019). COPD is often a primary cause of death, but it is likely to be ruled a contributory cause of death, due to the comorbidity (McGarvey, John, Anderson, Zvarich, & Wise, 2007).

### 1.2.5 Treatment

The treatment for COPD differs per person. Currently, COPD cannot be reversed or cured. However, different types of medication provide patients with symptomatic relief and an improved quality of life. During exacerbations, patients are often treated with either Prednisone or oxygen treatment. Smoking patients are offered smoking cessation advice to encourage them to change, and ideally quit, their smoking habit (Soriano, Zielinski, & Price, 2009).

### 1.2.6 Speech characteristics

The upper airway functions primarily as part of the aerodigestive system, but it also functions as part of the speech production system (Monoson & Fox, 1987). An altered anatomy or physiology of the upper airway can cause abnormal speech. These abnormalities can be subdivided into resonance, articulation and phonation, which occur isolated or in combination. Abnormalities regarding resonance include deviations in vocal quality due to coupling of the vocal tract or vocal tract damping (Kummer, 2018). Articulatory deviances are characterized by mispronunciations of the speech sounds, caused by a lack of functional or organic control of the lips, hard palate, soft palate and the tongue (Kim, Kim, Jeon, Woo, & Kim, 2017). Abnormalities related to phonation cause deviant sound productions due to laryngeal dysfunctioning (Monoson & Fox, 1987). Patients suffering from COPD tend to show speech abnormalities during exacerbation, but the speech also becomes gradually more deviant in stable condition as the disease progresses (Mohklesi, 2003).

COPD patients generally show an abnormal swallowing pattern (Robinson et al., 2011). They swallow more often and more distinctively, also during respiration or speaking. Studies show that approximately 70 to 85% of patients experiences dysphagia to some extent (Good-Fratturelli, Curlee, & Holle, 2000; Robbins, Coyle, Rosenbek, Roecker, & Wood, 1999; Rosenbek, Robbins, Roecker, Coyle, & Wood, 1996). Mokhlesi, Logemann, Rademaker, Stangl and Corbridge (2002) argued that the altered swallowing physiology obtained in COPD patients might be the result of hyperinflation. Patients alter their swallow manoeuvres in order to protect themselves from aspiration, but patients are often unable to maintain the protective swallowing physiology during exacerbations (Mokhlesi et al., 2002). This could result in the increase of coughing and the occurrence of swallowing movements during speech.

Additionally, COPD patients might show signs of a hypotonic voice disorder. This type of disorder produces a weak and faint voice. The subglottal pressure does not rise and this causes asthenic hoarseness (Omori, 2011). Moreover, a study by Vyshedskiy and Murphy (2016) on the acoustic biomarkers of COPD showed that there are measurable differences between the lung sound patterns of COPD patients compared to age matched controls. The breathing difficulties could result in a deviant intensity, a lack of intonation, a decreased phonation time for sustained vowels, and a deviant pitch, depending on the compensatory strategy (Constantinescu et al., 2010).

## 1.3 E-health

### 1.3.1 Background

E-health knows many synonyms or near-synonyms, such as telehealth, telemedicine, e-medicine and telepractice. Even though these terms are frequently used as if they are interchangeable, there are subtle differences in definition. E-health refers to a combination of concepts, including health, technology and marketing (Oh, Rizo, Enkin, & Jadad, 2005). A qualitative analysis of 87 papers regarding e-health by Oh et al. (2005) revealed a tendency towards the following definition:

> E-health is an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology (Eysenbach, 2001, p. 2).

There are three common types of e-practice according to ASHA (2014):

1. Synchronous (client-interactive) services are conducted with video and audio connections in real time. This type is most similar to face-to-face encounters (ASHA, 2014). An example of a synchronous service is a situation in which patients receive feedback on their speech during a real time connection.
2. Asynchronous (store-and-forward) services relate to situations in which data are obtained and transmitted to a professional for interpretation (ASHA, 2014). An example of an asynchronous service is a situation in which patients record speech samples, which are then transmitted to a professional who provides the patient with an evaluation of the speech at a different time.
3. Hybrid services include a combination of synchronous and asynchronous services (ASHA, 2014). An example of a hybrid service is a situation in which patients receive feedback on their pre-recorded speech samples during an interactive appointment.

These three types are often used to assist both patients and medical professionals in the treatment of patients with chronic diseases. However, a fourth type of e-health services is more suitable for remotely diagnosing or monitoring patients instead of treating them. “Remote patient monitoring” includes a sensor or tracking device, which analyses several parameters related to the patient’s condition (Ong et al., 2016). Any changes in these parameters could result in an automatic warning for the patient, medical professional or both. Two potential benefits of remote patient monitoring are the restoration of a patient’s sense of anatomy to a certain extent, and the decrease in medical costs (Klersy et al., 2014). An example of remote patient monitoring is the use of a continuous glucose sensor, which provides glucose readings every few minutes. The sensor warns the user in case of an alarming blood glucose level (Adolfsson, Parkin, Thoman, & Krinelke, 2018).

### 1.3.2 Goals

According to Eysenbach (2001), the ‘e’ in e-health represents the ten goals or characteristics, namely:

1. *Efficiency*: e-health aims to increase the efficiency in healthcare and to reduce the medical costs for all groups in society.
2. *Enhancement of quality*: e-health aims to increase the quality of care by allowing individuals to compare different care providers.
3. *Empowerment:* e-health aims to empower consumers or patients by making medical knowledge accessible and by enabling patient choice.
4. *Encouragement:* e-health aims to encourage both patients and health care professionals to seek partnership and make decisions in a shared manner.
5. *Extension of the scope of health care:* e-health aims to extent the scope of health care beyond the conventional boundaries in both a geographical and conceptual matter.
6. *Education:* e-health aims to educate both patients and health care professionals by providing them with medical education and preventive information.
7. *Enablement:* e-health aims to enable standardized information exchange and communication between health care providers.
8. *Equity:* e-health aims to make health care more equitable by providing online access.
9. *Ethics:* e-health should solve ethical issues regarding informed consent and privacy.
10. *Evidence based:* e-health should be evidence based and the effectiveness and efficiency should not be assumed.

### 1.3.3 Benefits and disadvantages

E-health is considered both feasible and desirable. It could decrease the number of patient transfers and facilitates meeting the needs of underserved populations, such as people with a low socio-economic status and inhabitants of rural areas (Jennett & Andruchuck, 2001). Moreover, e-health contributes to providers’ and patients’ medical knowledge, because medical information is now easier accessible and not constrained by geography, nationality or institutional boundaries (Kaplan & Litewka, 2008). This patient-centred care enables patients to regain their self-anatomy, since more medical services are shifted from medical institutions to the homes of patients (Kaplan & Litewka, 2008). E-health contributes to the reduction of costs for many services, such as therapy or consults, and it enables patients to recover in the comfort of their own home (Greenhalgh et al., 2013). E-health might increase the satisfaction of patients if the care is of the same quality as face-to-face care (Molini et al., 2015).

However, e-health also knows its drawbacks. Patients are able to inform themselves by reading medical articles or instructions online, but many of them are unable to filter the information (De Castro Corrêa, Ferrari, & Berretin-Felix, 2013). Patients might be unable to distinguish between a relevant, evidence-based article or intervention and an unreliable source. In addition to this, individuals with a low income might receive care of less quality, because they tend to replace face-to-face care with cheaper and easier accessible e-health services (Young, Foster, Silander, & Wakefield, 2011). People can become more isolated because they see fewer specialists (Reynolds, Vick, & Haak, 2009). It is necessary for the health care professional to establish and maintain an empathic relationship with patients, even if there is no face-to-face contact (Brandt, Søgaard, Clemensen, Søndergaard, & Nielsen, 2018). Finally, several problems have yet to be overcome. These problems often relate to technology, regulation, training or the acceptance and recognition of benefits by both professionals and the public (Molini-Avejonas, Rondon-Melo, De La Higuera Amato, & Samelli, 2015).

### 1.3.4 E-health for language and speech pathology

#### *1.3.4.1 Previous studies*

Edwards, Stredler-Brown and Houston (2012) conducted a systematic research to investigate the application of e-health in the field of speech, language and hearing disorders. The majority of studies showed a preference for diagnostic and therapeutic e-health services over traditional, face-to-face methods (Edwards et al., 2012). Molini-Avejonas et al. (2015) conducted a systematic review to research the effectiveness and efficiency of e-health services for the language and speech domain. They analyzed 17 papers on language and 33 papers on hearing, focussing on assessment, intervention, screening, education and methodologies. Most studies showed an advantage for e-health over traditional services (85%), while 14% was unable to determine whether e-health offered more advantages. Less than 1% of the studies suggested that the traditional alternatives were preferred over the e-health approach, confirming the results of Edwards et al. (2012) (Molini-Avejonas et al., 2015).

Many studies regarding e-health focussed on apraxia of speech, aphasia and speech and language problems in patients with Alzheimer's disease. Hill, Theodoros, Russell and Ward (2009) researched the use of videoconferencing to assess speech and language skills in apraxic patients, based on the primary results of the pilot study by Theodoros, Russell, Hill, Cahill and Clark (2003). Language and speech pathologists assessed 11 participants both face-to-face and during a video call. The results revealed that there were no differences between the test scores obtained face-to-face and during the video call. The patients and professionals indicated their satisfaction during a questionnaire on the video call, even though professionals struggled with the assessment of severe apraxic patients (Hill et al., 2009). Palsbo (2007) previously obtained the same results for patients suffering from aphasia. She found that the agreement among speech and language pathologists ranged from 92% to 100% for the functional communication measures used to assess the aphasia. Palsbo did not conduct a survey to explore patients' opinions on video calls in comparison with face-to-face assessment.

Vestal, Smith-Oline, Hicks, Hutton and Hart (2006) confirmed these findings for the assessment of language and speech dysfunction in individuals with Alzheimer's disease. They examined the effectiveness of language assessment during in person assessments and during video calls. Vestal et al. also conducted a survey among the 10 patients to determine the general attitude with respect to e-health. The analysis revealed no differences in attitude regarding the face-to-face assessment and the online assessment of language and speech. The overall acceptance rate of the video calls was rated high for both the patients and professionals, provided that the video call quality was sufficient.

Little research has been done on the effects of audio compressing in relation to the recognition of acoustic cues in pathological speech (Frühholz, Marchi, & Schuller, 2016). A study by Sáenz-Lenchón et al. (2008) on the effect of audio compression in automatic detection of voice pathologies explored the performance of an automatic system for voice pathology detection. Saénz-Lenchón et al. used voice samples that were compressed in MP3 format and had different binary rates. The detectors characterized the voice signals by cepstral and noise measures, along with their derivatives. The results indicated no significant difference between the different rates, which suggested voice pathology detection does not require great amounts of storage space or transmission over wide-band communication channels.

### *1.3.4.2 Outcome measures*

Previous studies regarding language and speech pathologies have provided a broad range of outcome measures, which are suitable for the detection, or monitoring of pathological speech. In most studies, a distinction has been made between continuous (read or natural) speech and sustained vowels, since voices with different pathologies can be rated differently during these varying speech types (Maryn, De Bodt, & Roy, 2010). Sustained vowels have been used for the detection or monitoring of pathological speech due to the benefit of a clean voice signal (Boyanov & Hadjitodorov, 1997). This allows algorithms to extract voice features more easily, but a sustained vowel does not commonly occur in natural speech (Nathan, Vatanparvar, Rahman, Nemati, & Kuang, 2019). It is more complicated to use continuous speech for voice analysis, but the ecological validity is higher compared to sustained vowels (Maryn, & Roy, 2012).

Alemami and Almazadydeh (2018) investigated the potential of voice breaks in the detection of pathological speech. The degree of voice breaks corresponds with the absence or presence of periodic vibrations of the acoustic chords (Maryn, Corthals, De Bodt, Van Cauwenberge, & Deliyski, 2009). The number or degree of unvoiced segments can indicate vocal instability, which is associated with pathological speech (Maryn, Corthals, Van Cauwenberge, Roy, & De Bodt, 2010). Healthy speakers do not experience difficulty maintaining voicing during a sustained vowel. Therefore, the normative value for the fraction of unvoiced frames is zero. Any values higher than zero are considered a sign of pathology (Boersma, 2004). Hunter and Titze (2016) also suggested a degree of voice breaks of less than 1% for healthy adults. However, they indicated that the degree of voice breaks increases gradually to over 10% by age 90 (Hunter & Titze, 2016).

Jitter is defined as the parameter of frequency variation from cycle to cycle and is affected by the lack of control of vibration of the cords. Shimmer relates to the amplitude variation of the sound wave and occurs with the reduction of glottal resistance and mass lesions on vocal cords or with the presence of noise emission and breathiness (Datta, Sengupta, Banerjee, & Ghosh, 2019). Programs for the acoustic analysis of speech offer numerous options to measure jitter and shimmer, but jitter (local), jitter ppq3, jitter rap, shimmer (local), shimmer apq3 and shimmer apq5 are most commonly used for speech technology software (Farrús, Hernando, & Ejarque, 2007; Silva, Oliveira, & Andrea, 2009; Woubie, Luque, & Hernando, 2014). Jitter and shimmer measures have proven to be highly sensitive in the detection of pathological voices. The measures have been integrated into software programs containing algorithms to distinguish pathological voices from healthy voices (Texeira & Gonçalves, 2016; Vieira, 2002). Jiang, Zhang, MacCallum, Sprecher and Zhou (2009) investigated the potential of objective acoustic analysis methods with dynamic and traditional perturbation measures to describe pathological voices. They found that measures related to jitter and shimmer could potentially improve objective clinical analyses (Jiang et al., 2009). Pathological speakers often show higher jitter values and lower shimmer values compared to healthy speakers (Teixeira & Fernandes, 2014).

Jiang et al. (2009) also found evidence for the potential of the harmonic-to-noise ratio (HNR). HNR provides the ratio between periodic and non-periodic components, which reflects the efficiency of speech. A healthy, sonorant and harmonic voice is characterized by a high HNR, whereas a value lower than 7 dB is considered pathological (Jiang et al., 2009). A low HNR indicates hoarseness and breathiness, as seen in smokers. Previous studies suggest a strong relationship between the perception of voice quality and HNR (Ferrand, 2002, 2007).

### 1.3.5 The future of e-health for COPD patients

COPD patients would benefit from a smart, automated remote system which is able to correctly identify changes in speech acoustics that indicate an upcoming exacerbation. Previous research by Wormald, Moran, Reilly and Lacy (2008) on the diagnostic accuracy of a smart system for the identification of vocal fold paralysis showed that the speech analysis system demonstrated a 92% sensitivity and a 75% specificity. The system assessed the pitch perturbation, amplitude perturbation and HNR in the 78 speech samples obtained by telephone calls. These findings suggest that a smart system might be a reliable and convenient screening method for distorted speech, as observed during COPD exacerbations, provided that the system is further trained (Wormald et al., 2008).

El Elmary, Fezari and Amara (2014) have investigated acoustic voice analysis methods based on adaptive features in order to develop a detection system for speech pathologies. Their results suggested that it is possible to differentiate between many diseases based on acoustic speech measures. Elmary et al. emphasized the need for efficient features to obtain a high classification rate, but they also mentioned that the accuracy of the system increased with the number of parameters. An accurate detection or monitoring application could prevent COPD patients from being hospitalized by warning them in case of deterioration.

Brennan, Georgeadis, Baron and Barker (2004) investigated if patient performance during video calls differed from the performance during encounters in person, depending on the subgroup. They investigated the story telling skills of brain-injured patients with different ages, genders, education, and a different extent of technology experience. The study showed no differences in performance, suggesting age, gender, education and technology experience are no exclusionary factors for candacy for e-health (Brennan et al., 2004). A new health application for COPD patients should provide the patient and caregiver with a secure data connection, since medical information is highly sensitive (Mashima & Doarn, 2008). Furthermore, the application should be tested thoroughly to ensure patient outcomes are preserved or improved (Molini-Avejonas et al., 2015).

Recently, Shastry, Balasubramanium and Acharya (2014) and Nathan et al. (2019) have researched the assessment of COPD patients using biomarkers from natural speech with the help of mobile devices. Shastry et al. used acoustic voice analysis to compare the speech of COPD patients with that of healthy adults. All participants were asked to sustain a vowel for as long as they could. The study revealed a lower fundamental frequency, a lower pitch variability and higher jitter measurements for COPD patients compared to healthy speakers (Shastry et al., 2014). In the study by Nathan et al., the speech of 91 patients with COPD or asthma was compared with the speech of 40 healthy speakers. They used feature extraction algorithms to analyze read speech based on pause features and prosodic features, such as pause frequency, jitter and shimmer. Their algorithm showed a 68% accuracy for patients with a high disease severity, whereas the algorithm was accurate in 89% of the cases for patients with a moderate disease severity. Based on these results, Nathan et al. are currently attempting to improve the accuracy of the segmentation and extraction algorithms to adapt them for conversational speech and noisy environments (Nathan et al., 2019).

## 1.4 Current research

### 1.4.1 Problem definition and research goals

Exacerbations are often recognized at a late stage, which delays the treatment (Hurst et al., 2010). It is of great importance to diagnose the patient quickly and correctly, because early treatment with

Prednisone shortens the duration and the seriousness of the exacerbation and could prevent COPD patients from being hospitalized (Trappenburg et al., 2011). Moreover, many patients do not fully recover from an exacerbation. A study by Perera et al. (2007) on recovery after exacerbation in 73 adult COPD patients has shown that the symptoms did not recover to baseline in 23% of COPD patients (35 days post onset). In addition, 22% suffered a recurrent exacerbation which further deteriorated the patients' condition. Early detection and treatment could prevent further damage to the respiratory system (Perera et al., 2007). Previous studies on the use of e-monitoring to help manage patients with COPD have not yet lead to consensus regarding the most suitable acoustic parameters to detect exacerbations (Bolton, Waters, Peirce, & Elwyn, 2010).

Medical professionals from the Department of Pulmonary Diseases of Dekkerswald (Radboud University Medical Centre) have indicated that they are able to estimate the condition of a patient who suffers from COPD by listening to their speech (Pasman[1], 2018). This means they believe the speech of COPD patients during stable periods differs from the speech during exacerbations. The information regarding the acoustic speech characteristics of COPD patients could be of great value if the professionals' observations are correct. The acoustic markers could be implemented in speech pathology recognition software to prematurely identify signs of deterioration and send a warning to the patient (and health care professional).

The current research provides an exploratory analysis of the differences in acoustic measures for stable COPD patients and healthy controls. Furthermore, the current research aims to identify those acoustic speech characteristics that mark the beginning of an exacerbation in COPD patients.

### 1.4.2. Research questions

The following research questions have been formulated for this study:

1. Which (acoustic) measures extracted from read speech differ for COPD speakers in stable condition and healthy speakers?

2. In which aspects does the speech of COPD patients during an exacerbation differ from speech of COPD patients during stable periods?

[1] "Ideeën luisteren" is an internship report by Pasman (2018) which is available for download from a server that is not publicly accessible. The document can be consulted upon request.

2.1 Which (acoustic) measures extracted from read speech differ for COPD speakers during an exacerbation and COPD speakers in stable condition?

2.2 Which (acoustic) measures extracted from sustained vowels differ for COPD speakers during an exacerbation and COPD speakers in stable condition?

### 1.4.3 Hypotheses

COPD is not a speech pathology and most studies focussed on different medical conditions. Therefore, the hypotheses regarding the acoustic parameters in speech of COPD patients often had to be derived from pilot studies or studies with a different patient population. Based on those studies and the observations by Pasman (2018[2]), 14 variables were selected. The assessment of these variables is described in section 2.2 and 2.3.

This section provides an overview of the hypotheses per variable. The additional I, II and/or III between brackets indicate whether a variable is used for the comparison of:

I. read speech between stable COPD patients and healthy controls;
II. read speech between COPD patients in exacerbation and stable COPD patients;
III. sustained vowels between COPD patients in exacerbation and stable COPD patients.

The additional a, b, c or d corresponds with the different hypotheses for each variable. These letters also correspond with those used in section 3.1.2, 3.2.2 and 3.3.2 to refer to a particular hypothesis.

1. *Formants (II, III)*. Formants are related to the length and shape of the vocal tract. Shrivastava, Tripathi and Singh (2018) found that F2 and F3 are lower in adults with respiratory issues compared to healthy controls, whereas F1 was higher for affected individuals (Shrivastava et al., 2018). There was no difference in F4 between both conditions. In addition, formant frequencies are associated with the amount of fluid or mucus in the lung parenchyma (Lotan et al., 2019). An increased F1 suggests an increased amount of mucus or fluid. Therefore, it is hypothesized that COPD patients show higher F1

[2] "Analyse-opties" is an internship report by Pasman (2018) which is available for download from a server that is not publicly accessible. The document can be consulted upon request.

values (a) and lower F2 and F3 values (b) during exacerbation compared to their stable condition, whereas the F4 values are expected to be similar (c).

2. *Frequency (I, II, III).* Shrivastava et al. (2018) found that respiratory problems are related to a lower pitch. In addition, Ayoub, Larrouy-Maestri and Morsomme (in press), and Vincent and Gilbert (2011) found that smokers present lower pitch values than nonsmokers. Therefore, it is hypothesized that COPD patients during exacerbation present a lower pitch value than COPD patients in stable condition (a) and that COPD patients in stable condition show a lower pitch than healthy controls (b).

3. *Intensity (I, II, III).* Shrivastava et al. (2018) found evidence for a higher intensity in individuals with respiratory issues compared to healthy controls, since the speech of affected adults shows a pattern of 'bursts' to produce small parts of the utterance. Therefore, it is expected that COPD patients show a higher intensity during exacerbation than in stable condition (a), whereas stable COPD patients show similar values in comparison with healthy controls (b).

4. *Duration (III).* Parsa and Jamieson (2001) found that healthy speakers outperform pathological speakers during sustained vowel tasks. Pathological speakers show lower duration values, depending on the etiology of the pathology. Speakers with respiratory issues tend to show the lowest duration values, because they are short of breath (Awan, 2011; Constantinescu et al., 2010). Therefore, it is hypothesized that COPD patients in exacerbation show lower duration values than COPD patients in stable condition.

5. *Shimmer (III).* Each shimmer parameter knows a different threshold value to determine whether a voice is pathological. The threshold values for shimmer (local) corresponds with 3.81%, whereas the threshold values for shimmer apq3 and shimmer apq5 are unclear (Teixeira, Oliveira, & Lopes, 2013). Low values are related to pathological voices, but also to smoking habits (Pinto, Crespo, & Mourão, 2014). It is expected that the values obtained for COPD speakers in exacerbation are lower than for COPD speakers in stable condition

(a) and it is hypothesized that both the COPD patients in exacerbation (b) and stable COPD patients score below the threshold for pathological speech (c).

6. *Jitter (III).* Each jitter parameter knows a different threshold value to distinguish between healthy and pathological speakers. The threshold values for Jitter (local), jitter rap and jitter ppq5 correspond with respectively 1.04%, .68% and .84% (Teixeira et al., 2013). High values are related to pathological voices, but also to smoking habits (Pinto et al., 2014). It is hypothesized that the values obtained for COPD speakers in exacerbation are higher than for COPD speakers in stable condition (a) and it is hypothesized that both the COPD patients in exacerbation (b) and stable COPD patients exceed the threshold for pathological speech (c).

7. *HNR (III).* A low HNR value is related to pathological voices and according to Teixeira et al. (2013), the threshold value is 7 dB. Values lower than 7 dB are a sign of pathology. On average, smokers present lower HNR values than nonsmokers, because they might struggle with the vibration or adduction of the vocal folds which results in more random noise. Moreover, low HNR values have been obtained for asthma patients (John et al., 2018). It is hypothesized that COPD patients during exacerbation show lower HNR values than stable COPD patients (a) and it is expected that both the COPD patients in exacerbation (b) and stable COPD patients score below the threshold level (c).

8. *Voice breaks (III).* Any non-zero value obtained for voice breaks indicates pathology of the voice (Boersma, 2004). Voice breaks occur more frequently in speech of smokers in comparison to nonsmokers (Simberg, Udd, & Santtila, 2015). Therefore, it is hypothesized that the degree of voice breaks is higher for COPD patients in exacerbation than for stable COPD patients (a) and that the threshold of zero for pathology is exceeded by both the COPD patients in exacerbation (b) and in stable condition (c).

9. *Syllables per breath group (I, II).* Patients with COPD suffer from respiratory problems, especially during exacerbation (GOLD, 2019). Therefore, it is expected that COPD patients produce fewer syllables per breath group during an exacerbation than in stable condition

(a) and that stable COPD patients produce fewer syllables per breath group than healthy controls (b).

10. *Speaking rate and articulation rate (I, II).* COPD patients experience inflammation and respiratory problems, which acutely worsen during exacerbation (GOLD, 2019). These issues (shortness of breath, coughing, mucus) result in a slow, irregular speech rate (Prelock & Hutchins, 2018). Therefore, it is hypothesized that the speaking rate of COPD patients in exacerbation is lower than the speaking rate of stable COPD patients (a), whereas stable COPD patients show a lower speaking rate than healthy controls (b). In addition, it is hypothesized that the articulation rate of COPD patients in exacerbation is lower than for stable COPD patients (c), whereas the articulation rate for stable COPD patients is lower than for healthy controls (d).

11. *Pitch variability (I, II).* Pathological speakers sometimes show increased laryngeal tension, which can result in limited pitch variability compared to healthy speakers (Harel, Cannizzaro, & Snyder, 2004). This results in a monotonous, flat, occasionally shrill voice (McGlone & Hollien, 1963). Therefore, it is hypothesized that COPD speakers during exacerbation show a more restricted pitch variability than stable COPD speakers (a), whereas stable COPD speakers show a similar pitch variability compared to healthy controls (b).

12. *Mean center of gravity (I, II).* The steepness of the spectral slope is determined by a measure of speech effort: the steepness of the glottal pulse (Van Son & Pols, 1996). A higher center of gravity (and a level slope) correlates with the perceived sentence accent. Van Son and Pols (1996) found that vowel de-accentuation and vowel reduction, as seen in many types of pathological speech, strongly correlate with a lower center of gravity. Pathological speakers or individuals with respiratory problems also tend to use less intonation than healthy speakers (Alcock, Passingham, Watkins, & Vargha-Khadem, 2000). Therefore, it is hypothesized that COPD patients during exacerbation present lower values for center of gravity than stable COPD patients (a), whereas it is expected that values for stable COPD patients and healthy controls are similar (b).

13. *(Non)-linguistic inhalations and exhalations (I, II).* Non-linguistic in- and exhalations refers to inhalations and exhalations in linguistically odd places, for example in the middle of a constituent. Due to the respiratory problems of COPD patients, it is expected that they in- and exhale more often in exacerbation than in stable condition (a) and that stable COPD patients in- and exhale more often than healthy controls (b). Moreover, it is hypothesized that COPD patients in exacerbation in- and exhale more often in linguistically odd places (c), whereas stable COPD patients in- and exhale more often in linguistically odd places than healthy controls (d).

14. *Ratio voiced/silence intervals (I, II).* The ratio of voiced to silence intervals refers to the ratio between the total duration of intervals in which a speaker is talking and the total duration of intervals in which a speaker pauses. It has been observed by Pasman (2018) that COPD patients in exacerbation generally pause for a long time between sentences. Moreover, previous research on pauses during conversational responses showed that the mean pause time is higher for asthmatic speakers compared to healthy speakers, even in stable condition (Wiechern, Liberty, & Pattemore, 2018). Therefore, it is hypothesized that the ratio between voiced and silence intervals is smaller for COPD patients in exacerbation than for stable COPD patients (a), whereas stable COPD patients show a smaller ratio compared to healthy controls (b).

### 1.4.4. Reading guide

The first section of the paper has provided the background information required for further understanding of the research. Chapter 2 is concerned with the methodology used for this study, including a description of the acoustic parameters described in section 1.4.3. The third chapter provides a justification for the statistical methods used, and it presents the findings of the research. These findings are then bundled and interpreted in chapter 4, followed by a conclusion in the fifth chapter.

# 2. Methods

## 2.1 Participants

The voices of 11 native speakers of Dutch (7 males, 4 females) were recorded twice in a treatment room of the lung department of the Radboud University Medical Centre in the period from August 2016 until April 2017. All participants had officially been diagnosed with COPD. The participants were hospitalized due to an exacerbation and they had to stay in the hospital for 2 to 23 days ($M$ = 8.82, $SD$ = 6.11). The participants were requested to participate in the experiment after first receiving urgent care for their exacerbation. Patients suffering from additional lung diseases were excluded from participation. Detailed participant information, such as the COPD stage, number of previous exacerbations, age and years since diagnosis, is missing due to the retrospective nature of the research. Table 1 shows the available information for this study.

Table 1

*Overview of patients' characteristics*

| **Participant number** | **Sex** | **Duration of hospitalization (in days)** |
|---|---|---|
| 1 | Male | 8 |
| 2 | Male | 13 |
| 3 | Male | 4 |
| 4 | Male | 9 |
| 5 | Male | 2 |
| 6 | Male | 2 |
| 7 | Male | 23 |
| 8 | Female | 5 |
| 9 | Female | 10 |
| 10 | Female | 13 |
| 11 | Female | 8 |

Furthermore, 5 recordings of 5 healthy, adult speakers (4 males, 1 female) from the Corpus Gesproken Nederlands (CGN) were selected in order to compare (characteristics of) their speech with the speech of the COPD patients. These 5 recordings were obtained between 1998 and 2004. Detailed participant information, such as the participant's age and the region of origin, is missing due to the retrospective nature of this study.

## 2.2 Materials

All COPD patients were recorded twice (see Figure 1). The first recordings took place within 24 hours after receiving the urgent care, while the second recordings were obtained just before the patients were discharged from the hospital. The condition in which the first recordings were made is referred to as 'in exacerbation', whereas the condition in which the second recordings were obtained is referred to as 'stable COPD'. The recordings were made using two Relitech microphones and each recording consisted of two parts: a sustained vowel and read speech.

1. The first part represented a sustained phonation of the vowel [a:]. The participants were free to choose a comfortable vocal intensity. These sustained vowels were recorded because they allow effective and efficient separation of normal from pathological voices (Boyanov & Hadjitodorov, 1997). The procedure for the sustained vowel recordings was the same for the first and second recording.

2. The second part consisted of a reading of the phonetically balanced story 'De Koning' (Bomans, 1946). This story contains existing, highly frequent words and the sentences differ in length (Haasnoot, 2012). The data collector selected part of the story to read. The parts differed slightly between subjects, because a few patients were unable to finish the whole part. The patients read different parts during their first and second recording to avoid a learning effect.

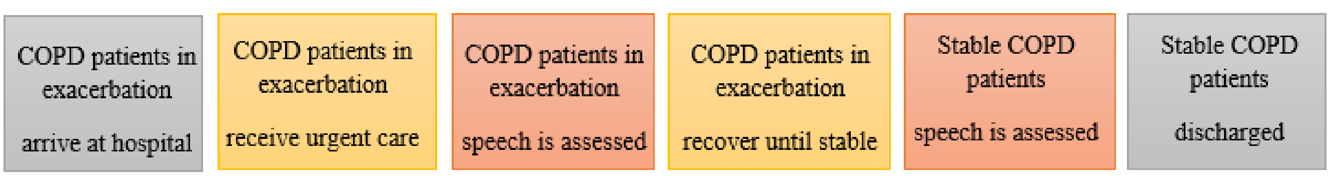


*Figure 1.* Schematic overview of data collection procedure

The recording of 1 patient reading aloud a story during stable COPD was damaged to such an extent that the speech was unintelligible. In addition, 1 patient was unable to execute the reading task during stable COPD due to reading difficulties. These 2 recordings were excluded. The database for the pathological speakers thus contained 11 recordings of a sustained vowel during an exacerbation (i), 11 recordings of a sustained vowel during stable COPD (ii), 11 recordings in which the patient read aloud a story during exacerbation (iii) and 9 recordings in which the patient read aloud a story during stable COPD (iv). The sampling frequency of all recordings was 16 kHz.

The recordings were compared with existing recordings from the healthy, adult control speakers from the CGN. The CGN was constructed between 1998 and 2004 and contains approximately 900 hours of Dutch speech, spoken by adults from Flanders and the Netherlands (Oostdijk, 2000). The corpus consists of different text types, such as stories, debates and conversations (Van Eerten, 2007). The control items were selected to match the speech of the COPD group. The speakers had to be healthy, adult, native speakers of Dutch as spoken in the Netherlands. Moreover, the controls had to read aloud (part of) a story. The 5 controls read a small part of the oronasal text ‘Papa en Marloes’ (Van de Weijer & Slis, 1991). They were allowed to read the text multiple times, but for this study, only the first recorded attempts were taken into account. The sampling frequency of all recordings was 16 kHz.

### 2.3 Procedure

The procedure of this study can be subdivided into 14 steps. Steps 1 to 4 were performed prior to the current research by a different data collector (see Figure 2), whereas steps 5 to 14 were executed during this research (see Figure 3 and 5).

1. All participants willing to participate received an oral explanation of the research by the data collector. After the description of the study, participants were allowed to ask questions. The data collector also explained that participants could stop at any given time without disclosing the reason.

2. The participants were then asked to sustain the vowel [a:] for as long as they could. The data collector marked the beginning of the task by clapping his hands. After the take, the patients could take a short break, but no patient took a break for longer than 30 seconds. There were no cases in which a retake was needed.

3. The participants started reading the assigned part of the story when they were ready. The data collector marked the end of the part by clapping his hands or by verbally telling them to stop. There were no cases in which a second take was needed.

4. This procedure was repeated on the day the patients were discharged from the hospital, but the participants then read a different part of the story ‘De Koning’ (Bomans, 1946) to avoid a learning effect.

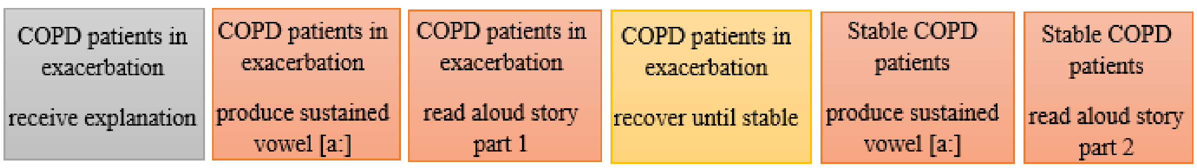


*Figure 2.* Schematic overview of step 1 – 4: recording procedure

5. Each data file was then split in two parts with the help of Praat (Version 6.0.54, Boersma & Weenink, 2019): one part containing the isolated sustained vowel and one part containing the isolated story. Both the sustained vowels and stories were checked for recording issues, during which the storytelling recordings of patient 4 and 5 in stable condition were excluded from analysis (as mentioned in section 2.2). Patient 4 had been unable to perform the task sufficiently, whereas the recording of patient 5 was severely damaged.

6. The files containing the story were automatically aligned on word and phoneme level for both the COPD patients in exacerbation (N = 11) and stable COPD patients (N = 9), using an existing forced aligner (Ganzeboom, 2014) and an adapted manual (see Appendix 5). This alignment was then checked manually in Praat. The forced alignment and manual check were repeated for the recordings of the healthy controls (N = 5).

7. Subsequently, the aligned files from both the COPD patients and the healthy controls were annotated according to the pre-designed protocol (see Appendix 1). The files each contained six tiers: interval tier for the story’s transcript (I), interval tier for word segmentations (II), interval tier for phonetic segmentations (III), interval tier for respiratory remarks (IV), point tier for speaker noises (V) and an interval tier for commentary (VI).

Figure 3 provides an example of the recording in Praat after the forced alignment, manual check and manual annotation.

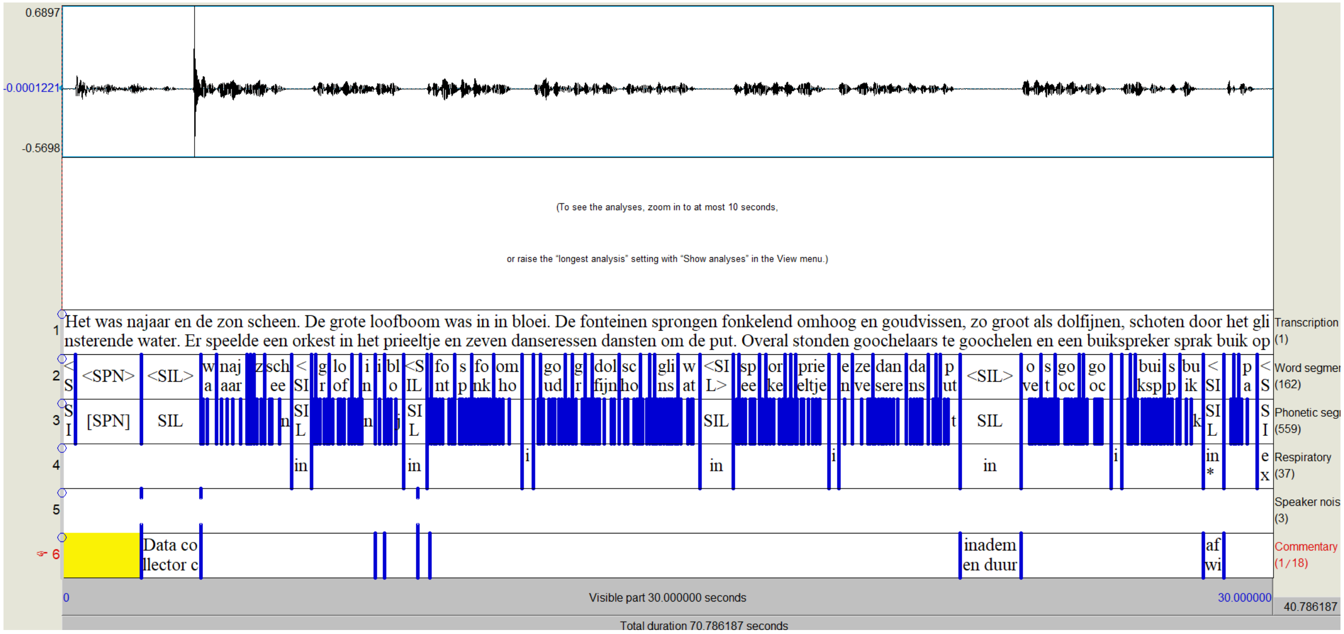


*Figure 3.* Example of recording with tiers after the forced alignment, manual check and manual annotation

8. The annotations were summarized in a pre-designed format and included in Appendix 2. Based on the annotations and summary per recording, several parameters and measures were manually calculated. Step 8.1 to 8.4 provides an explanation of the calculations.

   8.1 The number of syllables was manually calculated using the transcript (tier I). The syllables which were mistakenly produced (misread, slip of the tongue) were also taken into account. The number of syllables was not included as a measure for the analysis, but the number was used to calculate several other measures.

   8.2 The number of breath groups was manually calculated using the transcript and the annotations (tier IV). A breath group refers to the sequence of words that are articulated in one single exhalation, before the speaker pauses for breath. An overview of the breath groups per recording is included in Appendix 2. The number of breath groups was not included as a measure for the analysis, but the number was used to calculate the number of breath groups per syllable.

8.3 The number of in- and exhalations was manually calculated using the annotations (tier IV). An annotation was made each time the speaker in- or exhaled. These in- and exhalations were then added together (including the exhalations containing speech). The number of in- and exhalations was not included as a measure for the analysis, but the number was used to calculate the number of in- and exhalations per syllable.

8.4 The number of non-linguistic in- and exhalations was manually calculated using the annotations (tier IV) and the manual for linguistic pauses (Appendix 3). Each in- and exhalation was compared with an overview of linguistically acceptable in- and exhalations, based on a set of rules. The number of non-linguistic in- and exhalations consisted of those in- and exhalations that violated the rules specified in Appendix 3. The number was not included as a measure for the analysis, but the number was used to calculate the number of non-linguistic in- and exhalations per syllable.

9. Then, a script by Kerhoff (2015) was used to calculate the four formants, total duration, mean frequency, mean intensity, pitch variability, mean center of gravity, total duration of voiced intervals and the total duration of silence intervals for the recordings containing the story. The total duration of voiced intervals (all intervals containing speech) and the total duration of silence intervals (intervals containing breathing or pausing) were not included as measures for the final analysis, but they were used to calculate the ratio between voiced intervals and silence intervals.

10. The manually calculated measures (number of syllables, number of breath groups, number of in- and exhalations, number of non-linguistic in- and exhalations) were manually entered in SPSS. The results based on the script by Kerkhoff (2015) could not be imported into SPSS. Therefore, a script by Vos (2019) was used to extract the results from the result files. The results were finally imported into the existing SPSS file. The scripts by Kerkhoff

(2015) and Vos (2019) are included in Appendix 3. Figure 4 provides a schematic overview of the procedure for the files containing the storytelling.

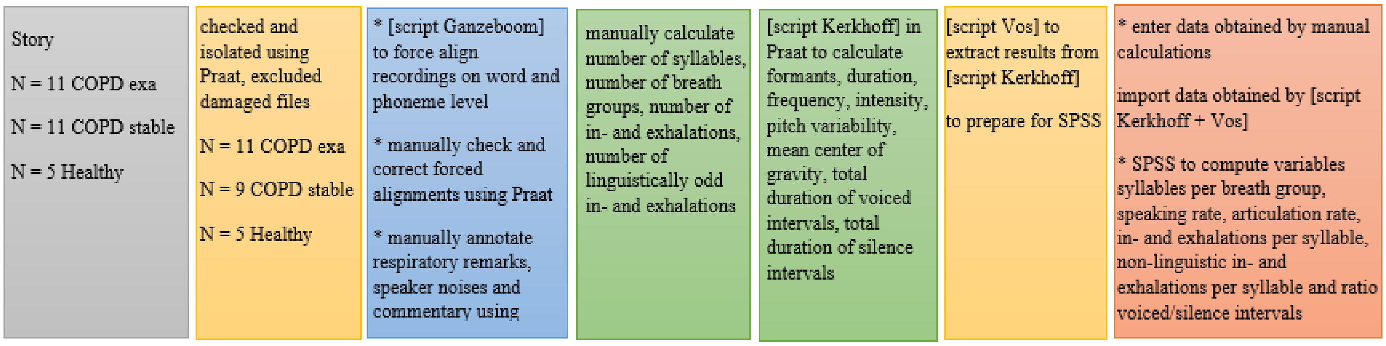


*Figure 4.* Schematic overview of the procedure for the storytelling recordings

11. The files containing the sustained vowels were analyzed using Praat's voice reports (Boersma & Weenink, 2019) to determine several acoustic parameters, namely shimmer (local), shimmer apq3, shimmer apq5, jitter (local), jitter ppq5, jitter rap, HNR and the degree of voice breaks. These measures were all used for the final analysis.

12. Subsequently, the files containing the sustained vowels were analyzed using the script by Kerkhoff (2015) to calculate the four formants, mean frequency, mean intensity and duration. These measures were all used for the final analysis.

13. The values obtained by the voice reports (shimmer (local), shimmer apq3, shimmer apq5, jitter (local), jitter ppq5, jitter rap, HNR, degree of voice breaks) were manually entered in SPSS. The results based on the script by Kerkhoff (2015) could not be imported into SPSS. Therefore, the script by Vos (2019) was also used to extract the results from these result files. The results for the sustained vowel were finally imported into the existing SPSS file. Figure 5 provides a schematic overview of this process for the sustained vowel.

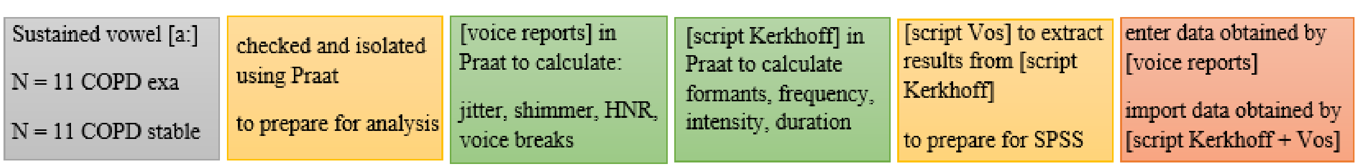


*Figure 5.* Schematic overview of the procedure for sustained vowels

14. Finally, the values in SPSS were used to compute additional variables:

    - The number of syllables per breath group was calculated dividing the number of produced syllables by the number of breath groups.

    - The speaking rate was calculated dividing the number of syllables by the total duration of the fragment.

    - The articulation rate was calculated dividing the adjusted number of syllables (speaking errors excluded) by the adjusted total duration of the voiced intervals (speaking errors excluded).

    - The number of in- and exhalations per syllable was calculated dividing the number of in- and exhalations by the number of syllables.

    - The number of non-linguistic in- and exhalations per syllable was calculated dividing the number of non-linguistic in- and exhalations by the number of syllables.

    - The ratio of voiced and silence intervals was calculated dividing the total duration of the voiced intervals by the total duration of the silence intervals.

## 2.4 Design and data analysis

This section provides an overview of the independent and dependent variables used for each analysis, followed by a short description of the type of analysis that was used for the comparison. The underlying assumptions for each type of analysis were assessed and described in section 3.1.1, 3.2.1 and 3.3.1.

### 2.4.1 Read speech: storytelling

Table 2 provides a list of variables used in the analysis of the storytelling recordings of stable COPD patients versus healthy controls (section 2.4.1.1) and COPD patients in exacerbation versus stable COPD patients (section 2.4.1.2).

Table 2

*Overview of (acoustic) parameters used for the within-subjects and between-subjects analysis of the storytelling recordings*

| **Parameter** | **Description** |
|---|---|
| Syllables per breath group | The number of syllables produced on one breath |
| Speaking rate | Speech tempo (in syllables per second) including pauses |
| Articulation rate | Speech tempo (in syllables per second) excluding pauses |
| Formants* | F1, F2, F3 and F4 in Hz |
| Mean frequency | Perceived pitch in Hz |
| Mean intensity | Perceived loudness in dB |
| Pitch variability | Range in variation in level and extent of pitch in Hz |
| Mean center of gravity | Weighted mean frequency in Hz |
| In- and exhalations | Inhalations and exhalations per syllable |
| Non-linguistic in- and exhalations | Inhalations and exhalations in non-linguistic places per syllable |
| Ratio voiced/silence intervals | Ratio between voiced intervals and silence intervals (-) |

Note: *Formants were included in the within-subjects analysis only, due to the possible influence of anatomical, physiological or social-linguistic speaker-related characteristics on formant frequencies in a between-subjects design (Adank, Smits, & Van Hout, 2004).

*2.4.1.1 Between COPD patients and healthy controls: story (I)*

The first comparison had a between-subjects design. The independent variable was ‘condition’ (stable COPD versus healthy) and the dependent variables corresponded with the 10 parameters (syllables per breath, ratio voiced/silence intervals, speaking rate, articulation rate, frequency, intensity, pitch variability, mean center of gravity, number of in- and exhalations per syllable, number of non-linguistic in- and exhalations per syllable). Group 1 (stable COPD) consisted of 9 participants, whereas group 2 (healthy controls) consisted of 5 participants.

A one-way multivariate analysis of variance (MANOVA) was conducted to investigate the effect of condition (stable COPD versus healthy controls) on the speech characteristics during the reading of a story. A MANOVA extends the possibilities of the traditional analysis of variance (ANOVA), because multiple dependent variables can be assessed simultaneously while the joint error rate does not increase (Grice & Iwasaki, 2007). Therefore, the risk of rejecting a true null hypothesis (type I error) is minimized. A MANOVA was preferred over ANOVA, because the current analysis included multiple dependent variables which were moderately correlated.

*2.4.1.2 Within COPD patients: story (II)*

The second comparison had a repeated measures design. The independent variable was 'condition' (in exacerbation versus stable COPD) and the dependent variables corresponded with the 14 parameters (syllables per breath group, ratio voiced/silence intervals, speaking rate, articulation rate, F1, F2, F3, F4, frequency, intensity, pitch variability, center of gravity, number of in- and exhalations per syllable, number of non-linguistic in- and exhalations per syllable).

A one-way repeated measures MANOVA was conducted to investigate the effect of condition (in exacerbation versus stable) on the speech characteristics of patients with COPD, because the design again included multiple, moderately correlated dependent variables. Section 2.2 and 2.3 described the exclusion of 2 recordings for the participants with stable COPD. A MANOVA is more reliable for a comparison between two or more groups with a similar size (Ito, 1980). Therefore, it was decided to also exclude the storytelling recordings of participant 4 and 5 in exacerbation. Group 1 (in exacerbation) and group 2 (stable COPD) thus consisted of the same 9 participants.

### 2.4.2 Sustained vowels

Table 3 provides the list of acoustic parameters for the within-subjects analysis of the sustained vowels.

Table 3

*Overview of (acoustic) parameters for the within-subjects analysis of the sustained vowels*

| **Parameter** | **Description** |
|---|---|
| Formants | F1, F2, F3 and F4 in Hz |
| Mean frequency | Perceived pitch in Hz |
| Mean intensity | Perceived loudness in dB |
| Duration | Duration (maximum phonation time) in seconds |
| Shimmer | Variability of the peak-to-peak amplitude in % |
| Shim apq3 | Three-point amplitude perturbation quotient in % |
| Shim apq5 | Five-point amplitude perturbation quotient in % |
| Jitter | Vocal fold frequency variability from cycle to cycle (frequency perturbation) in % |
| Jitter ppq5 | Five-point perturbation quotient in % |
| Jitter rap | Relative average perturbation in % |
| HNR | Proportion of the harmonic sound to noise in the voice (-) |
| Voice breaks | Fraction of pitch frames that are analyzed as unvoiced in % |

*2.4.2.1 Within COPD patients: sustained vowel (III)*

The third comparison had a repeated measures design. The independent variable was 'condition' (in exacerbation versus stable COPD) and the dependent variables corresponded with the 10 parameters (duration, F1, F2, F3, F4, frequency, intensity, jitter, jitter ppq5, jitter rap, shimmer, shimmer apq3, shimmer apq5, HNR and degree of voice breaks). Group 1 (in exacerbation) and group 2 (stable COPD) thus consisted of the same 11 participants. However, 1 participant had to be excluded from the analysis for both conditions due to the violation of an assumption. The reason for exclusion of this participant is disclosed in section 3.3.1. The final groups before analysis thus consisted of the same 10 participants.

A repeated measures MANOVA was conducted to investigate the effect of condition (in exacerbation versus stable COPD) on the speech characteristics of patients with COPD during the sustention of a vowel, because the design again included multiple, moderately correlated dependent variables.

# 3. Results

This chapter includes the results of the three MANOVAs announced in section 2.4.1 and 2.4.2. Each subsection starts with an examination of the data to rule out possible violations of the assumptions for MANOVAs. Then, the main effects and trends are described, whereas the non-significant results are exclusively mentioned in Tables 4 to 6. These tables show the results of all analyses, providing the mean values (with standard deviation between brackets), *F*-values, *p*-values and effect size ($\eta_p^2$). Each subsection ends with a short summary of the results.

## 3.1 Between COPD patients and healthy controls: story

### 3.1.1 Assumptions MANOVA I

Normality of data (i) was assessed using both statistical tests and visual inspection. Shapiro-Wilk tests did not indicate any violations of the assumption of normality ($p > .05$). Visual inspection of box plots and Q-Q plots for each variable confirmed these results, showing no univariate outliers. Multivariate outliers were examined calculating Mahalanobis distances. These values were compared with a chi-square distribution with the same degrees of freedom to calculate probabilities. These probabilities were higher than .001, indicating no multivariate outliers (ii).

Box's M test of equality of covariance indicated homogeneity of variance-covariance matrices (iii), whereas Levene's test of equality of error variances showed that the error variance of the dependent variables was equally distributed across groups ($p > 0.05$). Calculations of the Durbin-Watson statistic did not indicate a violation of the independence of observation assumption (iv). The values were not lower than 1.5 or greater than 2.5, and therefore, the data were not autocorrelated. Scatterplot matrices showed a linear or curvilinear relationship between the different pairs of dependent variables (v). Multicollinearity was assessed conducting correlations among the dependent variables. There were no correlations over .80 or VIF values exceeding 2.5 (vi).

### 3.1.2 Results MANOVA I

A one-way MANOVA was carried out to measure the effect of condition (stable COPD versus healthy controls) on the 10 dependent variables (see Table 4). The total number of in- and exhalations per syllable (1) and the number of non-linguistic in- and exhalations per syllable (2) were both higher for stable COPD patients ($M_1$ = .178, $SD_1$ = .038; $M_2$ = .030, $SD_2$ = .024) than for healthy controls ($M_1$ =.091, $SD_1$ = .028; $M_2$ = 0.000, $SD_2$ = .000). These differences of respectively .086 95% CI [.044 - .129] and .030 95% CI [.006 - .055] were both significant $F(1, 12) = 19.299$, $p = .001$, $\eta_p^2 = .617$; $F(1, 12) = 7.381$, $p = .019$, $\eta_p^2 = .381$. Both effects were medium to large sized. The ratio between voiced intervals and silenced intervals was higher for stable COPD patients ($M$ = 4.51, $SD$ = .66) than for healthy controls ($M$ = 2.07, $SD$ = .89). This difference of 2.439 95% CI [.032 – 4.847] was significant $F(1, 12) = 4.876$, $p = .047$, $\eta_p^2 = .289$. The effect of condition on ratio was small to medium sized. Finally, there was a trend for the effect of condition on the number of syllables per breath group. The number of syllables per breath group was lower for stable COPD patients ($M$ = 12.78, $SD$ = 3.18) than for healthy controls ($M$ = 17.43, $SD$ = 5.46). This difference of 4.648 95% CI [-.317 – 9.614] was not significant $F(1, 12) = 4.160$, $p = .064$, $\eta_p^2$ = .257. There was no effect of condition on the mean frequency, mean intensity, pitch variability and mean center of gravity, speaking rate and articulation rate ($p > .05$).

Table 4

*MANOVA I: Overview of F-values, p-values and $\eta_p^2$ for all variables*

| **Parameter** | **M (SD)** | | ***F*** | ***p*** | **$\eta_p^2$** |
|---|---|---|---|---|---|
| | **Stable (N = 9)** | **Healthy (N = 5)** | | | |
| Syllables per breath group | 12.78 (3.18) | 17.43 (5.46) | 4.160 | .064 | .257 |
| Speaking rate (syl/sec) | 3.67 (.19) | 3.51 (.25) | .284 | .604 | .023 |
| Articulation rate (syl/sec) | 4.64 (.21) | 5.28 (.28) | 3.418 | .089 | .222 |
| Inhalations & exhalations per syllable | .178 (.038) | .091 (.028) | 19.299 | .001** | .617 |
| Non-linguistic in- and exhalations per syllable | .030 (.024) | .000 (.000) | 7.381 | .019* | .381 |
| Ratio voiced/silence intervals (-) | 4.51 (.66) | 2.07 (.89) | 4.876 | .047* | .289 |
| | | | | | |
| Mean frequency (Hz) | 190.77 (91.69) | 128.90 (20.13) | 2.145 | .169 | .152 |
| Mean intensity (dB) | 63.32 (8.12) | 67.72 (5.31) | 1.161 | .302 | .088 |
| Pitch variability (Hz) | 710.27 (408.30) | 394.74 (116.01) | 2.768 | .122 | .187 |
| Mean center of gravity (Hz) | 504.34 (203.60) | 723.37 (255.97) | 3.117 | .103 | .206 |

Note: * = statistically significant with $p < .05$, ** = statically significant with $p < .01$, two-tailed testing.

### 3.1.3 Conclusion MANOVA I

The analysis showed a significant effect of condition (stable COPD versus healthy) on the number of in- and exhalations per syllable, the number of non-linguistic in- and exhalations per syllable and the ratio of voiced and silence intervals. The number of in- and exhalations per syllable and the number of non-linguistic in- and exhalations per syllable were higher for COPD patients than for healthy controls, which confirmed both hypotheses (13b, 13d). The ratio of voiced and silence intervals was higher for COPD patients compared to healthy controls, which was not in line with the hypothesis (14b). In addition, there was a trend for the effect of condition on the number of syllables per breath group. The number was higher for healthy controls than for stable COPD patients, which was in line with the hypothesis (9b). The remaining hypotheses (2b, 3b. 10b, 10d, 11b and 12b) could not be confirmed based on the results.

## 3.2 Within COPD patients: story

### 3.2.1 Assumptions MANOVA II

Normality of data (i) was assessed using both statistical tests and visual inspection. Shapiro-Wilk tests did not suggest any violations of the assumption of normality ($p > .05$). Visual inspection of box plots for each variable did not indicate any univariate outliers. Multivariate outliers (ii) were examined calculating Mahalanobis distances. These values were compared with a chi-square distribution with the same degrees of freedom to calculate probabilities. The probabilities were higher than .001, indicating no multivariate outliers.

Box's M test of equality of covariance (iii) showed that the data failed the assumption ($p < .05$). A series of Levene's $F$ tests showed that the error variance of the dependent variables was equally distributed across groups ($p > 0.05$), except for the number of syllables per breath group ($p < .05$). However, the homogeneity of variance assumption was considered satisfied, due to the robustness of MANOVA to violations of this assumption in the case of a similar sample size (Friendly & Sigal, 2018; Ito, 1980). Calculations of the Durbin-Watson statistic did not indicate a violation of the independence of observation assumption (iv). Multicollinearity (v) was assessed using VIF and tolerance values and by running correlations. Ideally, VIF values do not exceed 10 and bivariate correlations should not be higher than .9. The obtained values did not indicate a presence of multicollinearity, thus the assumption was not violated. The assumption of linearity among the pairs of dependent variables (vi) was violated. Most pairs showed a linear relationship, but few pairs (such as F1 – F2, speaking rate – articulation rate) tended towards a more curvilinear or less distinctive relationship. Section 4.4 provides a reflection on the compromised power of the MANOVA.

### 3.2.2 Results MANOVA II

A one-way repeated measures MANOVA was carried out to measure the effect of condition (in exacerbation versus stable COPD) on the dependent variables (see Table 5). Table 5 gives an overview of the means, standard deviations, *F*-value, *p*-value and $\eta_p^2$ for all variables. The analysis showed a medium sized effect of condition on the number of syllables per breath group $F(1, 16) = 7.413$, $p = .015$, $\eta_p^2 = .347$. The number of syllables per breath group was lower during exacerbation ($M = 9.50$, $SD = 1.70$) than in stable condition ($M = 12.78$, $SD = 3.18$). The difference of 3.28 95% CI [.725 – 5.827] was significant. The analysis also showed a small to medium sized effect of condition on the number of in- and exhalations per syllable ($F(1, 16) = 6.091$, $p = .025$, $\eta_p^2 = .276$). The number was higher during exacerbation ($M = .244$, $SD = .071$) than in stable condition ($M = .178$, $SD = .048$) and this difference of .066 95% CI [.009 - .123] reached significance. The results did not support an effect of condition on speaking rate, articulation rate, the number of non-linguistic in- and exhalations per syllable, the ratio of voiced and silence intervals, F1, F2, F3, F4, frequency, intensity, pitch variability and center of gravity ($p > .05$).

Table 5

*MANOVA II: Overview of the mean values (SD), F-values, p-values and $\eta_p^2$ for all variables*

| **Parameter** | **Mean (SD)** | | ***F*-value** | ***p*-value** | **$\eta_p^2$** |
|---|---|---|---|---|---|
| | **Exacerbation (N = 9)** | **Stable (N = 9)** | | | |
| Syllables per breath group | 9.50 (1.70) | 12.78 (3.18) | 7.413 | .015* | .317 |
| Speaking rate (syl/sec) | 3.55 (.53) | 3.67 (.57) | .234 | .635 | .014 |
| Articulation rate (syl/sec) | 5.01 (.51) | 4.64 (.61) | 1.908 | .186 | .107 |
| In- and exhalations per syllable | .244 (.071) | .178 (.048) | 6.091 | .025* | .276 |
| Non-linguistic in- and exhalations per syllable | .043 (.042) | .031 (.024) | .653 | .431 | .039 |
| Ratio voiced/silence intervals (-) | 2.92 (1.83) | 4.51 (2.39) | 2.505 | .133 | .135 |
| F1 (Hz) | 529.08 (54.08) | 555.49 (59.94) | .963 | .341 | .057 |
| F2 (Hz) | 1701.46 (112.70) | 1731.42 (91.99) | .382 | .545 | .023 |
| F3 (Hz) | 2859.84 (152.27) | 2871.88 (167.57) | .025 | .875 | .002 |
| F4 (Hz) | 3969.85 (201.77) | 3956.87 (179.95) | .021 | .887 | .001 |
| Mean frequency (Hz) | 154.05 (31.01) | 190.78 (91.69) | 1.296 | .272 | .075 |
| Mean intensity (dB) | 64.53 (6.73) | 63.33 (8.12) | .119 | .735 | .007 |
| Pitch variability (Hz) | 586.53 (238.99) | 710.27 (408.30) | .616 | .444 | .037 |
| Mean center of gravity (Hz) | 439.12 (152.32) | 504.34 (203.60) | .592 | .453 | .036 |

Note: * = statistically significant with $p < .05$, two-tailed testing.

### 3.2.3 Conclusion MANOVA II

The MANOVA showed an effect of condition (in exacerbation versus stable COPD) on the number of syllables per breath group and the number of in- and exhalations per syllable. The number of syllables per breath group was lower during exacerbation than in stable condition, which was in line with the hypothesis (9a). The number of in- and exhalations per syllable was higher for COPD speakers during exacerbation than for stable COPD speakers. This confirmed hypothesis (13a). The remaining hypotheses (1a – 1c, 2a, 3a, 10a, 10c, 11a, 12a, 13c, 14a) could not be confirmed based on the results.

## 3.3 Within COPD patients: sustained vowel

### 3.3.1 Assumptions MANOVA I

Normality of data (i) was assessed using both statistical tests and visual inspection. Shapiro-Wilk tests did not indicate any violations of the assumption of normality ($p > .05$). However, visual inspection of boxplots showed one univariate outlier for each variable. Extreme values were obtained for participant 6 (in stable condition). Moreover, Mahalanobis Distances were calculated to identify possible multivariate outliers (ii). After calculation the Mahalanobis Distances, these values were compared to a chi-square distribution with the same degrees of freedom. The resulting probabilities were greater than .001, except for patient 6 in stable condition ($p < .001$). Therefore, patient 6's recording of the sustained vowel in stable condition and in exacerbation were excluded from the analysis, resulting in two groups of 10 participants.

Box's M test of equality of covariance showed that the observed covariance matrices of the dependent variables (iii) were distributed equally across the groups ($F(28, 1393) = 34.901$, $p = .810$). Also, Levene's test of equality of error variances showed that the error variance of the dependent variables was equally distributed across groups ($p > 0.05$), except for the degree of voice breaks ($p = .02$) and the jitter and shimmer variables ($.01 < p < .05$). Nonetheless, these variables were included in the analysis, since MANOVA is robust to violations of the assumption of homogeneity, provided that the sample sizes are similar (Friendly & Sigal, 2018; Ito, 1980). The recording of participant 6 was excluded for both conditions, maintaining an equal sample size per group. Calculations of the Durbin-Watson statistic did not indicate a violation of the independence of observation assumption (iv). The obtained values did not indicate a presence of multicollinearity, thus the assumption was not violated. However, several variables (F1 – F2, F3 – F4) showed low correlations. The assumption of linearity among the pairs of dependent variables (vi) was violated. Most pairs showed a linear relationship, but few pairs (such as F1- F2, F3 – F4) tended towards a more curvilinear or less distinctive relationship. Section 4.4 provides a reflection on the compromised power of the MANOVA.

### 3.3.2 Results MANOVA III

A one-way repeated measures MANOVA was carried out to measure the effect of condition (in exacerbation versus stable COPD) on the acoustic parameters (see Table 6). Table 6 gives an overview of means and standard deviations per group, and the *F*-value, *p*-value and $\eta_p^2$ for all parameters. The mean duration of the sustained vowel was shorter during exacerbation (*M* = 5.96, *SD* = 2.98) than in stable condition (*M* = 8.70, *SD* = 2.52) and this difference of 2.745 95% CI [.055 – 5.435] did reach significance, $F(1, 18) = 4.636$, $p = .046$, $\eta_p^2 = .214$. The effect of condition on vowel duration was small to medium sized.

The analysis also revealed a small to medium sized effect of condition on shimmer ($F(1, 18) = 5.150$, $p = .037$, $\eta_p^2 = .232$), shimmer apq3 ($F(1, 18) = 4.462$, $p = .049$, $\eta_p^2 = .208$) and shimmer apq5 ($F(1, 18) = 5.693$, $p = .029$, $\eta_p^2 = .251$). Shimmer measurements were higher during exacerbation ($M_s = 8.57$, $SD_s = 5.64$; $M_{apq3} = 4.34$, $SD_{apq3} = 2.91$; $M_{apq5} = 5.23$, $SD_{apq5} = 3.36$) than in stable condition ($M_s = 18$, $SD_s = 1.41$; $M_{apq3} = 2.19$, $SD_{apq3} = .92$; $M_{apq5} = 2.48$, $SD_{apq5} = .81$). The differences for shimmer (4.397, 95% CI [.309 – 8.484]), shimmer apq3 (2.144, 95% CI [.002 – 4.285]) and shimmer apq5 (2.747, 95% CI [.318 – 5.177]) were significant ($p < .05$).

Moreover, the HNR was lower during exacerbation (*M* = 15.09, *SD* = 5.15) than in stable condition (*M* = 19.49, *SD* = 2.68) and this difference of 4.394 95% CI [.345 – 8.443] was significant, $F(1, 18) = 5.424$, $p = .035$, $\eta_p^2 = .236$. The effect of condition on HNR was small to medium sized. Finally, the analysis revealed a small to medium sized effect of condition on the degree of voice breaks. The degree of voice breaks was higher during exacerbation (*M* = 1.26, *SD* = 1.71) than in stable condition (*M* = .00, *SD* = .00). This difference of 1.258 95% CI [.052 – 2.464] was significant, $F(1, 18) = 4.843$), $p = .042$, $\eta_p^2 = .222$. The analysis did not show an effect of condition on intensity, frequency, formants or jitter measurements ($p > .05$).

Table 6

*MANOVA III: Overview of the means (SD), F-values, p-values and $\eta_p^2$ for all parameters*

| **Parameter** | **M (SD)** | | ***F*-value** | ***p*-value** | $\eta_p^2$ |
|---|---|---|---|---|---|
| | **Exacerbation (N = 10)** | **Stable (N = 10)** | | | |
| Intensity (dB) | 67.31 (8.99) | 68.77 (7.11) | .152 | .701 | .009 |
| Frequency (Hz) | 143.59 (29.44) | 150.61 (43.11) | .175 | .681 | .010 |
| Duration (s) | 5.96 (2.98) | 8.70 (2.52) | 4.636 | .046* | .214 |
| | | | | | |
| F1 (Hz) | 600.62 (114.48) | 613.39 (94.37) | .069 | .795 | .004 |
| F2 (Hz) | 1353.75 (181.26) | 1272.69 (167.82) | 1.016 | .328 | .056 |
| F3 (Hz) | 2828.83 (457.65) | 2874.41 (307.11) | .063 | .804 | .004 |
| F4 (Hz) | 4043.99 (178.72) | 4025.94 (217.49) | .039 | .845 | .002 |
| | | | | | |
| Jitter (%) | 1.45 (.96) | .81 (.45) | 3.310 | .086 | .163 |
| Jitter ppq5 (%) | .88 (.66) | .45 (.25) | 3.368 | .084 | .165 |
| Jitter rap (%) | .74 (.48) | .50 (.31) | 1.628 | .219 | .087 |
| Shimmer (%) | 8.57 (5.64) | 4.18 (1.41) | 5.150 | .037* | .232 |
| Shimmer apq3 (%) | 4.34 (2.91) | 2.19 (.92) | 4.462 | .049* | .208 |
| Shimmer apq5 (%) | 5.23 (3.36) | 2.48 (.81) | 5.693 | .029* | .251 |
| | | | | | |
| HNR (-) | 15.09 (5.15) | 19.49 (2.68) | 5.424 | .035* | .236 |
| Voice breaks (%) | 1.26 (1.71) | .00 (.00) | 4.843 | .042* | .222 |

Note: * = statistically significant with $p < .05$, two-tailed.

### 3.3.3 Conclusion MANOVA III

The results revealed a significant effect of condition (in exacerbation versus stable COPD) on duration, shimmer, shimmer apq3, shimmer apq5, HNR and the degree of voice breaks. The duration of the vowel was shorter, the shimmer outcomes were higher, the HNR was lower and the degree of voice breaks was higher for COPD patients during exacerbation compared to the stable condition. The shorter vowel duration and the lower HNR were in line with hypothesis 4 and 7a. However, both groups did not score below the threshold level of 7 dB, indicating no sign of pathology. This was not in line with hypothesis 7b and 7c. The shimmer outcomes were higher for patients in exacerbation compared to stable COPD patients, which was not in line with the hypothesis (5a). The shimmer outcomes did not exceed the threshold value in both conditions, indicating no sign of pathology, which was not in line with hypotheses 5b and 5c. The degree of voice breaks was higher during exacerbation than in stable condition, which confirmed hypothesis 8a. In addition, the outcomes for the degree of voice breaks during exacerbation did exceed the thresholds for pathological voices, unlike the outcomes for stable COPD speakers. This was in line with the hypothesis for COPD patients in exacerbation (8b), but not in line with the hypothesis for COPD patients in stable condition (8c). The remaining hypotheses (1a – 1c, 2a, 3 and 6a – 6c) could not be confirmed based on the results.

# 4. Discussion

The outcomes of this research have provided insight into the acoustic differences between the speech of COPD patients during exacerbation, COPD patients in stable condition and healthy controls. However, the results should be interpreted with caution due to the limitations of the current research. This chapter provides a reflection on the research process. The limitations and potential consequences of the design are discussed, as well as the implications for the interpretation of the results. The chapter ends with several recommendations for future research.

### 4.1 Participants

Many lung diseases are grouped under the umbrella term COPD. Most patients are diagnosed based on spirometry, but this method is unable to distinguish between different types of obstruction (Lopez et al., 2014). The diagnosis should be composed of spirometry combined with a clinical entity to rule out other lung diseases, such as asthma. The data were pre-recorded and there were no available records of participant characteristics or used methods to determine the diagnosis. Therefore, the patients could have suffered from comorbidity. The diagnosis should have been confirmed first to rule out any patients who suffer from similar lung diseases or other pathologies that could influence the speech characteristics. A larger, more homogeneous group of participants might have had resulted in less ambiguous and more reliable outcomes.

The limited information regarding the participants and the small sample sizes might have contributed to the large standard deviations and the lack of significant results in this research. It is possible that the sample consisted of different subgroups, based on etiology, age, years since diagnosis or the severity of the COPD (Gupta et al., 2016). Therefore, it is uncertain if the obtained results can be generalized to the COPD population. In addition, it was often assumed that the COPD patients were smokers. Many hypotheses based on this assumption, such as the hypotheses regarding pitch variability, formant frequencies and center of gravity, could not be confirmed based on the results of the current study. It is recommended to collect detailed participant information in order to formulate specific, substantiated hypotheses and collect additional evidence for the obtained trends.

Finally, the conditions were referred to as ‘in exacerbation’ and ‘stable’. The exacerbations were diagnosed by a medical professional. However, it is possible that not all patients were fully recovered on the day they were discharged. Patients might have preferred a recovery at home, leaving the hospital prematurely. There was great variation in the number of days patients stayed at the hospital (range: 2 – 23 days). This might have resulted in an insufficient contrast between the in exacerbation condition and the stable condition. It is recommended to assess the speech of patients for a second time once they are fully recovered instead of on the day they are discharged from the hospital. In addition, the COPD patients should be matched with healthy controls to minimize the risk of confounding factors, such as age, gender and level of education.

### 4.2 Materials

Tehrany, Barney and Bruton (2014) mentioned that spontaneous speech is more valuable and suitable for the detection of speech abnormalities in pathological speakers. The tasks used in this study consisted of reading aloud a phonetically balanced text and sustaining a vowel. These types of non-natural speech might have distorted the results. Healthy adults exhibit more and longer grammatically inappropriate breath group intervals during spontaneous speech compared to reading tasks (Wang, Green, Nip, Kent, & Kent, 2010). This suggests that grammatically inappropriate pauses might not be a useful marker in practice. On the other hands, respiratory markers have been proven useful to distinguish patients with Parkinson's disease from healthy controls based on spontaneous speech and unstructured tasks (Huber & Darling, 2011; Pakhomov et al., 2010). These results were also obtained for the detection of speakers with Friedreich's Ataxia (FRDA) during conversational speech (Rosen et al., 2010). It could be of importance to investigate if the difference in breathing pauses holds during a comparison of spontaneous speech.

The ratio of time that adult speakers spend pausing silently to the time spend speaking is influenced by the difficulty of the speech task (Thurber & Tager-Flusberg, 1993). Longer and syntactically more complex utterances are interrupted more frequently and these pauses tend to be longer (Grosjean & Collins, 1979). This means pauses are not only a sign of a distorted breathing pattern, but they could easily be the result of difficulty with reading (Huber, Darling, Francis, & Zhang, 2012). Moreover, previous research has shown that cognitive load can influence many acoustic parameters, such as formants and frequency (Boril, Sadjadi, Kleinschmidt, & Hansen, 2010; Yap, Epps, Ambikairajah, & Choi, 2015). It is possible that the reading task was more challenging for the COPD patients than for the healthy controls. This might have resulted in a distortion of the results regarding the breathing pattern (Yap et al., 2015).

In addition, the healthy speakers and the COPD patients read different stories. The healthy speakers read shorter sentences, which provided fewer options for non-linguistic inhalations, exhalations or pauses. Therefore, results regarding the difference in (non-)linguistic in- and exhalations and the ratio between voiced and silence intervals should be interpreted with caution. It is recommended to use the same story for patients and healthy controls to ensure a fair comparison in the case of read speech.

### 4.3 Procedure

The annotator was aware of the recordings' characteristics during the annotating process. This means that the annotator was aware of the speaker's identity (participant number) and the type of recording (during or after exacerbation). This might have influenced the annotator's judgments and limits the reliability of this study. Therefore, it is recommended to train multiple annotators and calculate reliability measures to determine the degree of agreement for future research. It is also advised to train a data collector before the recordings are obtained. The data collector often interrupted the patients as soon as they started to breathe heavily, whereas the patients were encouraged to read longer passages in stable condition. This might have distorted the results regarding the voiced-to-silence ratio, resulting in a higher ratio for patients in exacerbation compared to stable patients. In addition to this, the healthy speakers were recorded according to a different protocol. The original purpose of their recording was different and therefore, they might have received instructions to read slowly and carefully or to pause between sentences. These different instructions might have distorted the results, resulting in a higher voiced-to-silence ratio for stable COPD patients compared to healthy controls.

Roughness, hoarseness and breathiness indicate the presence of jitter. Even though these voice characteristics were observed by the medical specialists, the data collector and the annotator, the jitter outcomes did not differ significantly between COPD patients during exacerbation and in stable condition. These results were not in line with results from previous studies on jitter measures for pathological voices. The lack of difference might be attributed to differences in algorithms. Jitter algorithms assume periodicity, but commercial algorithms often generate jitter measures for signals that violate this assumption. Jitter values generated by commercial algorithms therefore show great variability, which might have caused the lack of difference between COPD patients during exacerbation and in stable condition (Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995).

Previous studies have shown the influence of recording systems on voice parameters, such as jitter and shimmer. Perry, Ingrisano and Blair (1996) found that analog recording systems introduce a significant amount of jitter and shimmer error. The details of the recording procedure for both the COPD patients and healthy controls were unclear. The results for jitter and shimmer might have been distorted due to the recording system (Perry et al., 1996). Moreover, varying distances to the microphone can cause differences in shimmer (Ferrand, 2002). It is possible that the patients were positioned differently during the recordings, which might have resulted in an overestimation of the shimmer values during exacerbation (Brockmann-Bauer, Bohlender, & Mehta, 2018).

Xue and Lower (2010) researched the impact of different bandwidths on acoustic measures, using low-cost internet technology. They collected voice samples of normal speech using Skype and VoiceEmotion software, while they obtained disordered speech samples through teaching CD samples. The fidelity analysis included average fundamental frequency, jitter percent, shimmer percent and HNR as acoustic parameters. The analyses showed an increase in (variability) of the average frequency for both the samples of normal and disordered speech, with a greater increase for the latter. Jitter, shimmer and HNR differed significantly pre- and post-transmission. This means that the acoustic measures normally obtained in person can be distorted by mode of transmission. It is uncertain if the acoustic measures have been distorted by the mode of transmission due to the lack of information regarding the recording and transmission procedure. The distortion of acoustic measures should be considered designing a monitoring application.

### 4.4 Design and data-analysis

Previous studies have suggested that qualitative voice characteristics can be measured objectively by identifying acoustic markers (Jones, Trabold, Plante, Cheetham, & Earis, 2008). Hoarseness seems to be an important voice characteristic in COPD patients due to the respiratory problems and coughing. Jones et al. (2008) have suggested objective, acoustic measures to measure hoarseness. In addition, the study by Vyshedskiy and Murphy (2016) showed that there are measurable differences between the lung sound patterns in COPD patients compared to age-matched controls. Statistical differences between COPD speakers and healthy controls were found for 11 parameters, including the inspiration-to-expiration ratio and the low-to-high-frequency ratio. These results confirm that there are additional relevant, acoustic parameters that should be taken into account while developing an application. Additional variables were not taken into account for this study due to the limitations of a MANOVA (the number of dependent variables should not exceed the number of participants) and the limitations of the script by Kerkhoff (2015).

The results were not always in line with the hypotheses. The shimmer measurements turned out to be lower for COPD patients in stable condition compared to COPD patients in exacerbation and both groups showed normal shimmer values. According to Rabinov et al. (1995) and Heman-Ackah et al. (2003), HNR and jitter and shimmer measures are unreliable in the quantification of perceptual voice characteristics. They suggested that measures related to the cepstral peak prominence are more reliable to assess voice changes and less sensitive to additional noise (Rabinov et al., 1995).

The dependent variables were moderately correlated in most cases. However, for the third MANOVA (within-subjects comparison of sustained vowel), several correlations were lower than .3. Moreover, the assumption of linearity amongst the dependent variables was violated for both repeated measures MANOVAs. Several pairs of dependent variables showed a very weak linear relationship or a curvilinear relationship. This has compromised the power of both MANOVAs. The lack of linearity and the low correlations might have been caused by the relatively small sample sizes ($5 < N < 10$). The use of multiple ANOVAs with a Bonferroni correction is suggested as a substitute for the MANOVA in various handbooks and articles (Cabin & Mitchell, 2000; Field, 2005a, 2005b; Huberty & Morris, 1989; O'Brien & Kaiser, 1985). However, this alternative is advised against if the remaining dependent variables are strongly correlated (Chatfield, 2018; Field 2005a, 2005b). This was the case for the current research.

### 4.5 Future research

The current study can be interpreted as a first step in the research on COPD speech characteristics. However, the results of this study had to be treated with caution due to the small sample size and the lack of details regarding the participants' characteristics and the data collection procedure. Future research could further examine the differences in speech characteristics between COPD patients in exacerbation, stable COPD patients and healthy controls, as well as contribute to a deeper understanding of the acoustic measurements suitable for e-health. Section 4.1 to 4.4 provided an overview of alternatives to improve the current research design. The following section includes additional suggestions for future research.

1. *Comparing automatic acoustic analysis with real time perceptual judgments*. Many studies, including systematic reviews, include preliminary results or show limited validity and reliability or low levels of evidence (Molini et al., 2015). Previous research has suggested the possibility of a gap between the benefits and efficacy mentioned in studies (with randomized controlled trials) and the effectiveness and user friendliness obtained in clinical reality (Granja, Janssen, & Johansen, 2018). In addition, a previous pilot study showed that the addition of remote monitoring did not improve the quality of life (Antoniades et al., 2012). Therefore, it could be of importance to examine the differences in sensitivity, specificity, effectiveness and quality of life between traditional diagnostic methods and traditional methods with the addition of remote monitoring.

2. *Comparing the costs of traditional methods with the costs of remote monitoring systems*. It is of importance to investigate the suggested cost benefits for e-health. Even though multiple reviews on e-health report cost benefits as an advantage of e-health, many papers do not confirm this assumption. Molini et al. (2015) mentioned that less than 13% of the studies they investigated proved to be cost effective.

# 5. Conclusion

The current research aimed to identify acoustic speech characteristics which mark the beginning of an exacerbation in COPD patients. The central questions for this research were:

1. *Which acoustic measures extracted from read speech differ for COPD speakers in stable condition and healthy speakers?*

2. *In which aspects does the speech of COPD patients during an exacerbation differ from speech of COPD patients during stable periods?*

All recordings containing the storytelling of ‘De Koning’ (Bomans, 1946) or ‘Papa en Marloes’ (Van de Weijer & Slis, 1991) were forced aligned using a script. Subsequently, all recordings were manually annotated in Praat to indicate respiratory actions, such as in- and exhaling. Various measures, such as the number of breath groups or the number of non-linguistic in- and exhalations, were manually calculated, whereas other measures, such as formants and pitch, were generated using a second script. The recordings containing the sustained vowel [a:] were analyzed using Praat. All measures were generated using the second script or the voice reports function in Praat.

The recordings of 9 stable COPD patients reading aloud part of De Koning were then compared with the recordings of 5 healthy controls reading aloud part of Papa en Marloes. The results showed a significant effect of condition on the number of in- and exhalations per syllable, the number of non-linguistic in- and exhalations per syllable and the ratio of voiced and silence intervals. The number of in- and exhalations per syllable and the number of non-linguistic in- and exhalations per syllable were higher for COPD patients than for healthy controls, which confirmed both hypotheses. However, the higher ratio of voiced and silence intervals for COPD patients compared to healthy controls was not in line with the hypothesis. This unpredicted result might have been caused by the different reading materials or recording procedures for both groups, or by a difference in reading skills. Moreover, there was a trend for the effect of condition on the number of syllables per breath group. The number of syllables per breath group was higher for healthy controls compared to COPD patients, which was in line with the hypothesis. There was no effect of condition on pitch, intensity, center of gravity, pitch variability, speaking rate and articulation rate.

Next, the recordings of 9 COPD patients reading aloud a story during exacerbation were compared with their 9 counterparts in stable condition to determine which acoustic measures extracted from read speech differed for both conditions. The results showed that there was a significant effect of condition on the number of syllables per breath group and the number of in- and exhalations per syllable. The lower number of syllables per breath group and the higher number of in- and exhalations per syllable for COPD patients in exacerbation compared to stable COPD patients were in line with the hypothesis. There was no effect of condition on pitch, intensity, formants, mean center of gravity, pitch variability, number of non-linguistic in- and exhalations per syllable, ratio between voiced and silence intervals, speaking rate and articulation rate.

Finally, the pre-recorded sustained vowels for 10 COPD patients during exacerbation were compared with their 10 counterparts in stable condition to determine which acoustic measures extracted from sustained vowels differed for both conditions. The results showed that there was a significant effect of condition on duration, shimmer measurements, HNR and the degree of voice breaks, whereas there was no effect of condition on pitch, intensity, formant frequencies and jitter measurements. The duration was shorter, the HNR was lower and the degree of voice breaks was higher for COPD patients in exacerbation compared to stable COPD patients, which confirmed all three hypotheses. However, the shimmer measurements were higher for COPD patients in exacerbation, which was not in line with the hypothesis. This unpredicted outcome might have been caused by differences in position during the recording or by the type of recording system.

This research has shown that the speech of COPD patients in exacerbation differs from the speech of COPD patients in stable condition. The HNR and duration of a sustained vowel might have potential for the detection of exacerbations, as well as the number of syllables per breath group and the number of in- and exhalations per syllable. However, sustained vowels rarely occur in spontaneous speech. Therefore, the last two outcome measures might have greater potential for the detection of beginning exacerbations. After the initial warning based on the number of syllables per breath group and the number of in- and exhalations per syllable, the patients could be presented with a quick check in the form of a sustained vowel. After production of the sustained vowel, HNR and duration could be calculated to reject or confirm the initial warning. Further research on the different outcome measures and their potential for the detection of exacerbations is needed due to the limitations of the current study.

# 6. References


Adank, P., Smits, R., & Van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America, 116*(5), 3099 – 3107.

Adolfsson, P., Parkin, C.G., Thomas, A., & Krinelke, L.G. (2018). Selecting the appropriate continuous glucose monitoring system – a practical approach. *European endocrinology, 14*(1), 24 – 29.

Alcock, K.J., Passingham, R.E., Watkins, K., & Vargha-Khadem, F. (2000). Pitch and timing abilities in inherited speech and language impairment. *Brain and Language, 75*(1), 34 – 46.

Alemami, Y., & Almazaydeh, L. (2018). Pathological voice signal analysis using machine learning based approaches. *Computer and Information Science,* 11(1), 8 – 13.

Allison, K.M., Annear, L., Policicchio, M., & Hustad, K.C. (2017). Range and precision of formant movement in paediatric dysarthria, *60*(7), 1864 – 1876.

American Speech-Language-Hearing Association (ASHA). (2014). Telepractice: overview. Retrieved June 2, 2019, from: https://www.asha.org/Practice-Portal/Professional-Issues/Telepractice/

Antoniades, N.C., Rochford, P.D., Pretto, J.J., Pierce, R.J., Gogler, J., Steinkrug, J., Sharpe, K., & McDonald, C.F. (2012). Pilot study of remote telemonitoring in COPD. *Telemedicine and e-health, 18*(8), 634 – 640.

Awan, S.N. (2011). The effect of smoking on the Dysphonia Severity Index in females. *Folia Phoniatrica et Logopaedica, 63,* 65 – 71.

Ayoub, M.R., Larrouy-Maestri, P., & Morsomme, D. (in press). The effect of smoking on the fundamental frequency of the speaking voice. *Journal of Voice.*

Bailly, G., & Gouvernayre, C. (2012). Pauses and respiratory markers of the structure of book reading. *13th Annual Conference of the International Speech Communication Association (Interspeech, 2012).* Portland, United States.

Barnes, P.J. (2004). Small airways in COPD. *New England Journal of Medicine, 350*(26), 2635 – 2637.

Boersma, P. (2004). Stemmen meten met Praat. *Stem- Spraak en Taalpathologie, 12*(4), 237 – 251.

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.54, retrieved June 1, 2019, from: http://www.praat.org/

Bolton, C.E., Waters, C.S., Peirce, S., & Elwyn, G. (2010). Insufficient evidence of benefit: a systematic review of home telemonitoring for COPD. *Journal of Evaluation in Clinical Practice, 17*(6), 1216 – 1222.

Bomans. G. (1946). *Sprookjes.* Amsterdam: Elsevier.

Boril, H., Sadjadi, O., Kleinschmidt, T., & Hansen, J. (2010). Analysis and detection of cognitive load and frustration in drivers' speech. *Proceedings of Interspeech*, 502 – 505.

Boyanov, B., & Hadjitodorov, S. (1997). Acoustic analysis of pathological voices. A voice analysis system for the screening of laryngeal diseases. *IEEE Engineering in Medicine and Biology, 16*(4), 74 – 82.

Brandt, C.J., Søgaard, G.I., Clemensen, J., Søndergaard, J., & Nielsen, J.B. (2018). Determinants of successful e-health coaching for consumer lifestyle changes: qualitative interview study among health care professionals. *Journal of Medical Internet Research, 20*(7), e237.

Brennan, D.M., Georgeadis, A.C., Baron, C.R., & Barker, L.M. (2004). The effect of videoconference-based telerehabilitation on story retelling performance by brain-injured subjects and its implications for remote speech-language therapy. *Telemedicine Journal and E-health, 10*(2), 147 – 154.

Brockmann-Bauer, M., Bohlender, J.E., & Mehta, D.D. (2018). Acoustic perturbation measures improve with increasing vocal intensity in individuals with and without voice disorders. *Journal of Voice, 32*(2), 162 - 168.

Buist, A.S., McBurnie, M.A., Vollmer, W.M., Gillespie, S., Burney, P., Mannino, D.M., … Nizankowska-Mogilnicka, E. (2007). International variation in the prevalence of COPD (The BOLD study): a population-based prevalence study. *Lancet, 370,* 741 – 750.

Cabin, R.J., & Mitchell, J. (2000). To Bonferroni or not to Bonferroni. When and how are the questions. *Bulletin of the Ecological Society of America, 81*(3), 246 – 248.

Celli, B.R., MacNee, A., Agusti, A., Anzueeto, B., Berg, A.S., Buist, P.M.A., … ZuWallack, R. (2004). Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *European Respiratory Journal, 23,* 932 – 946.

Chatfield, C. (2018). Procedures based on the multivariate normal distribution. In C. Chatfield (Ed.), *Introduction to multivariate analysis* (pp. 89 – 113). New York: Routledge.

Cho, M.H., Boutaoui, N., Klanderman, B.J., Sylvia, J.S., Ziniti, J.P., Hersh, C.P., … Silverman, E.K. (2010). Variants in FAM13A are associated with chronic obstructive pulmonary disease. *Nature Genetics*, *42*, 200 – 202.

Cho-Reyes, S., Mack, J.E., & Thompson, C.K. (2016). Grammatical encoding and learning in agrammatic aphasia: evidence from structural priming. *Journal of Memory and Language, 91,* 202 – 218.

Constantinescu, G., Theodoros, D., Russell, T., Ward, E., Wilson, S., & Wootton, R. (2010). Assessing disordered speech and voice in Parkinson's disease: a telerehabilitation application. *International Journal of Language & Communication Disorders, 45*(6), 630 – 644.

Cosio-Piqueras, M.G., & Cosio, M.G. (2001). Disease of the airways in chronic obstructive pulmonary disease. *European Respiratory Journal, 18,* 41 – 49.

Datta, A.K., Sengupta, R., Banerjee, K., & Ghosh, D. (2019). Jitter, shimmer and complexity perturbation analysis. In A.K. Dataa, R. Sengupta, K. Banerjee & D. Ghosh (Eds.), *Acoustical analysis of the tanpura. Signals and communication technology* (pp. 121 – 132). Singapore: Springer.

Davis, B.H. (2009). Examining pauses in Alzheimer's discourse. *American Journal of Alzheimer's Disease & Other Dementias, 24*(2), 141 – 154.

De Castro Corrêa, C., Ferrari, D.V., & Berretin-Felix, G. (2013). Quality, range and legibility in web sites related to orofacial functions. *International Archives of Otorhinolaryngology, 17*(4), 1809 – 1817.

Dell, G.S., & Burger, L.K. (1997). Language production and serial order: a functional analysis and a model. *Psychological Review, 104*(1), 123 – 147.

Dodd, B. (2005). Children with speech disorder: defining the problem. In B. Dodd (Ed.), *Differential Diagnosis and Treatment of Children with Speech Disorder* (pp. 7 – 30). West-Sussex: Whurr Publishers Ltd.

Edwards, M., Stredler-Brown, A., & Houston, K.T. (2012). Expanding use of telepractice in speech-language pathology and audiology. *Volta Review, 112,* 227 – 242.

El Elmary, I.M.M., Fezari, M., & Amara, F. (2014). Towards developing a voice pathologies detection system. *Journal of Communications Technology and Electronics, 59*(11), 1280 – 1288.

Eysenbach, G. (2001). What is e-health. *Journal of Medical Internet Research, 3*(2), e20.

Farrús, M., Hernando, J., & Ejarque, P. (2007). Jitter and shimmer measurements for speaker recognition. *INTERSPEECH 2007, 8th Annual Conference of the International Speech Communication Association*, 27 – 31.

Ferrand, C. (2002). Harmonics-to-Noise Ratio. *Journal of Voice, 16*(4), 480 – 487.

Ferrand, C. (2007). *Speech science: an integrated approach to theory and clinical practice.* Boston, MA: Pearson.

Field, A. (2005a). The beast of bias. In A. Field (Ed.), *Discovering Statistics Using IBM SPSS Statistics* (pp. 623 – 664). UK: Sussex.

Field, A. (2005b). Multivariate analysis of variance (MANOVA). In: A. Field (Ed.), *Discovering Statistics Using IBM SPSS Statistics* (pp. 163 – 212). UK: Sussex.

Friendly, M., & Sigal, M.J. (2018). Visualizing tests for equality of covariance matrices. *The American Statistician, 1*, 2 – 12.

Frühholz, S., Marchi, E., Schuller, B. (2016). The effect of narrow-band transmission on recognition of paralinguistic information from human vocalizations. *IEEE Access, 4,* 6059 – 6072.

Gayraud, F., Lee, H-R., & Barkat-Defradas, M. (2011). Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Clinical Linguistics & Phonetics, 25*(3), 198 – 209.

GOLD: Global Initiative for Chronic Obstructive Lung Disease (2019). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease [Report]. Accessed June 2, 2019, from: https://goldcopd.org/gold-reports/

Gómez, C.G., Jaeger, T.F., & Smyth, R. (2008). Incremental syntactic planning across clauses. *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 845 – 850).

Good-Fratturelli, M.D., Curlee, R.F., & Holle, J.F. (2000). Prevalence and nature of dysphagia in VA patients with COPD referred for videofluoroscopic swallow examination. *Journal of Communication Disorders, 33*(2), 93 – 110.

Granja, C., Janssen, W., & Johansen, M.A. (2018). Factors determining the success and failure of eHealth interventions: systematic review of literature. *Journal of Medical Internet Research, 20*(5), e10235.

Greenhalgh, T., Wherton, J., Sugarhood. P., Hinder, S., Procter, R., & Stones, R. (2013). What matters to older people with assisted living needs? A phenomenological analysis of the use and non-use of telehealth and telecare. *Social Science & Medicine, 93*, 86 – 94.

Grice, J.W., & Iwasaki, M. (2007). A truly multivariate approach to MANOVA. *Applied Multivariate Research, 12*(3), 199 – 226.

Grosjean, F., & Collins, M. (1979). Breathing, pausing and reading. *Phonetica*, *36*, 98 – 114.

Gupta, N., Pinto, L., Benedetti, A., Li, P.Z., Tan, W.C., Aaron, S.D., … Bourbeau, J. (2016). The COPD Assessment Test: can it discriminate across COPD subpopulations? *Chest*, *150*(5), 1069 – 1079.

Haasnoot, R. (2012). De ontwikkeling van de fonetisch uitgebalanceerde standaardtekst De Koning [Master thesis]. Radboud University. Retrieved June 2, 2019, from: http://haasnoot.net/sitebuildercontent/sitebuilderfiles/dekoningrhaasnootdefinitief.pdf

Harel, B., Cannizzaro, M., & Snyder, P.J. (2004). Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: a longitudinal case study. *Brain and Cognition*, *56*(1), 24 – 29.

Hawkins, P.R. (1971). The syntactic location of hesitation pauses. *Language and Speech*, *14*(3), 277 – 288.

Heman-Ackah, Y.D., Heuer, R.J., Michael, D.D., Ostrowski, R., Horman, M., Baroody, M.M., Hillenbrand, J., & Salatoff, R.T. (2003). Cepstral peak prominence: a more reliable measure of dysphonia. *Annals of Otology, Rhinology, and Laryngology*, *112*(4), 324 – 333.

Hill, A.J., Theodoros, D., Russell, T., & Ward, E. (2009). Using telerehabilitation to assess apraxia of speech in adults. *International Journal of Language & Communication disorders*, *44*(5), 731 – 747.

Huber, J.E. (2008). Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory Physiology & Neurobiology*, *164*(3), 323 – 330.

Huber, J.E., & Darling, M. (2011). Effect of Parkinson's disease on the production of structured and unstructured speaking tasks: respiratory, physiologic and linguistic considerations. *Journal of Speech, Language, and Hearing Research*, *54*(1), 33 – 46.

Huber, J.E., Darling, M., Francis, E.J., & Zhang, D. (2012). Impact of typical aging and Parkinson's disease on the relationship among breath pausing, syntax, and punctuation. *American Journal of Speech-Language Pathology*, *21*(4), 368 – 379.

Huberty, C.J., & Morris, J.D. (1989). Multivariate analysis versus multiple univariate analysis. *Psychological Bulletin, 105*(2), 302 – 308.

Hunter, E.J., & Titze, I.R. (2016). Progression of voice breaks in a non-pathological voice as an indicator of aerodigestive health. *Journal of the American Geriatrics Society, 64*(10), e.93 – e.94.

Hurst, J.R., Vestbo, J., Anzueto, A., Locantore, N., Müllerova, H., Tal-Singer, R., … Wedzicha, J.A. (2010). Susceptibility to exacerbation in chronic obstructive pulmonary disease. *The New England journal of medicine, 363*(12), 1128 – 1138.

Ito, P.K. (1980). 7 Robustness of ANOVA and MANOVA test procedures. *Handbook of Statistics, 1*, 199 – 236.

Jennett, P.A., & Andruchuk, K. (2001). Telehealth: 'real life' implementation issues. *Computer Methods and Programs in Biomedicine, 64*(3), 169 – 174.

Jiang, J.J., Zhang, Y., MacCallum, J., Sprecher, A., & Zhou, L. (2009). Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. *Folia Phoniatrica et Logopaedica, 61*(6), 342 – 349.

John, K., Balagpangu, S., Jeromy, A.K., Albert, D.A., Christopher, N., & Godfred, B.A. (2018). Speech signal analysis as an alternative to spirometry in asthma diagnosis. *Journal of Biomedical Engineering and Medical Devices, 3*(3), 1 – 6.

Jones, T.M., Trabold, M., Plante, F., Cheetham, B.M.G., Earis, J.F. (2008). Objective assessment of hoarseness by measuring jitter. *Clinical Otolaryngology, 2*(1), 29 – 32.

Kaplan, B., & Litewka, S. (2008). Ethical challenges of telemedicine and telehealth. *Cambridge Quarterly of Healthcare Ethics, 17*(4), 401 – 416.

Kim, J., Kim, S.W., Jeon, H.R., Woo M.R., & Kim, H.I. (2017). Speech and linguistic features of children with articulation disorder. *Annals or Rehabilitation Medicine, 41*(5), 836 – 842.

Klersy, C., De Silverstri, A., Gabutti, G., Raisaro, A., Curti, M., Regoli, F., & Auricchio, A. (2014). Economic impact of remote patient monitoring: an integrated economic model derived from a meta-analysis of randomized controlled trials in hearth failure. *European Journal of Hearth Failure, 13*(4), 450 – 459.

Kummer, A.W. (2018). A pediatrician's guide to communication disorders secondary to cleft lip/palate. *Pediatric clinics, 65*(1), 31 – 46.

Lobanov, B.M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America, 49*(2), 606 – 608.

Lee, J-W., Kang, H-G., C., J-Y., Son, Y-I. (2013). An investigation of vocal tract characteristics for acoustic discrimination of pathological voices. *BioMed Research International*, e758731.

Levelt, W.J.M. (2002). Phonological encoding in speech production: Comments on Jurafsky et al., Schiller et al., and Heuven and Haan. In C. Gussenhoven, N. Warner (Eds.), *Laboratory Phonology VII* (pp. 87 – 99). Berlin: Mouton de Gruyter.

López, C.C., Serdá, G.J., Lacalzada, C.C., Medina, A.M., Blanco, J.A.G., Bello, M.A.G., & Navarro, P.C. (2014). Prevalence of Chronic Obstructive Pulmonary Disease in the Canary Islands. *Archivos de Bronconeumologia, 50*(7), 272 – 277.

Lotan, C., Kremer-Tal, S., Lotan, A., Schlik, Z., Gemer, A., Snir, Y., … Sheinkerman, M. (2019). Assessment of a pulmonary condition by speech analysis [patent application]. *US Patent Application Publication No. US 2019/0080803 A1*.

Maryn, Y., Corthals, P., De Bodt, M., Van Cauwenberge, P., & Deliyski, D. (2009). Perturbation measures of voice: a comparative study between multi-dimensional voice program and Praat. *Folia Phoniatrica et Logopaedica, 61*(4), 217 – 226.

Maryn. Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (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.

Maryn, Y., De Bodt, M., & Roy, N. (2010). The acoustic voice quality index: toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders, 43*(3), 161 – 174.

Maryn, Y., & Roy, N. (2012). Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. Jornal da Sociedade Brasileira de Fonoaudiologia, *24*(2), 107 – 112.

Maslan, J., Leng, X., Rees, C., Blalock, D., & Butler, S.G. (2011). Maximum phonation time in healthy older adults. *J Voice*, *25*(6), 709 – 713.

Mashima, P., & Doarn, C. (2008). Overview of telehealth activities in speech-language pathology. *Telemed Journal*, *14*(10), 1101 – 1117.

McBride, S., Gradyon, J., Sidani, S., & Hall, L. (1999). The therapeutic use of music for dyspnoea and anxiety in patients with COPD who live at home. *Journal of Holistic Nursing, 17*(3), 750 – 758.

McGlone, R.E., & Hollien, H. (1963). Vocal pitch characteristics of aged women. *Journal of Speech and Hearing Research,* 164 – 170.

McGarvey, L.P., John, M., Anderson, J.A., Zvarich. M., & Wise, R.A. (2007). Ascertainment of cause-specific mortality in COPD: operations of the TORCH Clinical Endpoint Committee. *Thorax, 62*(5), 411 – 415.

Mefferd, A.S., & Green, J.R. Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *Journal of Speech, Language and Hearing Research, 53*(5), 1206 – 1219.

Mokhlesi, B. (2003). Clinical implications of gastroesophageal reflux disease and swallowing dysfunction in COPD. *American Journal of Respiratory Medicine, 2*(2), 117 – 121.

Mokhlesi, B., Logemann, J.A., Rademaker, A.W., Stangl, C.A., & Corbridge, T.C. (2002). Ortopharyngeal deglutition in stable COPD patients. *Chest, 121*(2), 315 – 317.

Molini-Avejonas, D.R., Rondon-Melo, S., De La Higuera Amato, C.A., & Samelli, A.G. (2015). A systematic review of the use of telehealth in speech, language and hearing sciences. *Journal of Telemedicine and Telecare, 21*(7), 367 – 376.

Monoson, P.K., & Fox, A.W. (1987). Preliminary observation of speech disorder in obstructive and mixed sleep apnea. *Chest, 92*(4), 670 – 675.

Nathan, V., Vatanparvar, K., Rahman, M., Nemati, E., & Kuang, J. (2019). Assessment of chronic pulmonary disease patients using biomarkers from natural speech recorded by mobile devices. *2019 IEEE 16th International Conference on wearable and implantable body sensor networks (BSN).*

O'Brien, R.G., & Kaiser, M.K. (1985). MANOVA. Method for analyzing repeated measures designs. An extensive primer. *Psychological Bulletin, 97*(2), 316 – 333.

O'Connell, D.C., & Kowal, S. (1983). Pausology. In W.A. Sedelow & S.Y. Sedelow (Eds.), *Computers in Language Research 2* (pp. 221 – 301). Berlin: Mouton Publishers.

Oh, H., Rizo, C., Enkin, M., & Jadad, A. (2005). What is eHealth: a systematic review of published definitions. *Journal of Medical Internet Research, 7*(1), e1.

Oliveria, M. (2002). The role of pause occurrence and pause duration in the signalling of narrative structure. In E. Ranchhod & N.J. Mamede (Eds.), *Advances in Natural Language Processing* (pp. 43 – 51). Berlin: Springer.

Omori, K. (2011). Diagnosis of voice disorders. *Japan Medical Association Journal, 54*(4), 248 – 253.

Ong, M.K., Romano, P.S., Edgington, S., Aronow, H.U., Auerback, A.D., Black, J.T., … Fonarrow, G.C. (2016). Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure. *JAMA Internal Medicine*, *176*(3), 310 – 318.

Oostdijk, N. (2000). Digitiaal. Het Corpus Gesproken Nederlands. *Nederlandse Taalkunde*, *5*(3), 280 – 285.

Palsbo, S.E. (2007). Equivalence of functional communication assessment in speech pathology using videoconferencing. *Journal of telemedicine and telecare, 13*(1), 40 – 43.

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

Parsa, V., & Jamieson, D.G. (2001). Acoustic discrimination of pathological voices: sustained vowels versus continuous speech. *Journal of Speech, Language and Hearing Research, 44*(2), 327 – 339.

Pasman, I. (2018). Analyse-opties (intership report). Nijmegen: Radboud University.

Pasman, I. (2018). Ideeën luisteren (internship report). Nijmegen: Radboud University.

Perera, W.R., Hurst, J.R., Wilkinson, T.M.A., Sapsford, R.J., Müllerova, H., Donaldson, G.C., & Wedzicha, J.A. (2007). Inflammatory changes, recovery and recurrence at COPD exacerbation. *European Respiratory Journal*, *29*, 527 - 534.

Perry, C.K., Ingrisano, D.R-S., & Blair, W.B. (1996). The influence of recording systems on jitter and shimmer estimates. *American Journal of Speech-Language pathology, 5*(2), 86 – 90.

Pinto, A.G.L., Crespo, A.N., & Mourão, L.F. (2014). Influence of smoking isolated and associated to multifactorial aspects in vocal acoustic parameters. *Brazilian Journal of Otorhinolaryngology*, *80*(1), 60 – 67.

Postma, D.S., Bush, A., & Van den Berge, M. (2015). Risk factors and early origins of chronic obstructive pulmonary disease. *Lancet*, *385*(9971), 7 – 13.

Prelock, P.A., & Hutchins, T.L. (2018). Children with speech disorders. In P.A. Prelock & T.L. Hutchins, *Clincial guide to assessment and treatment of communication disorders. Best practices in child and adolescent behavioural health care* (pp. 75 – 87) . Cham, Switzerland: Springer.

Rabinov, C.R., Kreiman, J., Gerratt, B., & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of speech and hearing research, 38*(1), 26 – 32.

Ramanarayanan, V., & Bresch, E. (2009). Analysis of pausing behaviour in spontaneous speech using real-time magnetic resonance imaging of articulation. *The Journal of the Acoustical Society of America, 126,* EL160.

Reich, S.S. (1980). Significance of pauses for speech perception. *Journal of Psycholinguistic Research, 9*(4), 379 – 389.

Reynolds, A.L., Vick, J., & Haak, N.J. (2009). Telehealth applications in speech-language pathology: a modified narrative review. *Journal of Telemedicine an Telecare, 16*(6), 310 – 316.

Robbins, J.A., Coyle, J.L., Rosenbek, J.C., Roecker, E.B., & Wood, J.L. (1999). Differentiation of normal and abnormal airway protection during swallowing using the penetration-aspiration scale. *Dysphagia, 14*(4), 228 – 232.

Robinson, D.J., Jerrard-Dunne, P., Greene, Z., Lawson, S., Lane, S., & O'Neill, D.O. (2011). Oropharyngeal dysphagia in exacerbations of chronic obstructive pulmonary disease. *European Geriatric Medicine, 2*(4), 201 – 203.

Rong, P., Loucks, T., Kim, H., Hasegawa-Johnson, M. (2012). Relationship between kinematics, F2 slope and speech intelligibility in dysarthria due to cerebral palsy, *Clinical Linguist Phonology, 26*(9), 806 – 822.

Rosen, K., Murdoch, B., Folker, J., Vogel, A., Cahill, L., Delatycki, M., & Corben, L. (2010). Automatic method of pause measurement for normal and dysarthric speech. *Clinical Linguistics & Phonetics, 24*(2), 141 – 154.

Rosenbek, J.C., Robbins, J.A., Roecker, E.B., Coyle, J.L., & Wood, J.L. (1996). A penetration-aspiration scale. *Dysphagia, 11*(2), 93 – 98.

Rutschmann, O.T., Cornuz, J., Poletti, P-A., Bridevaux, P-A., Hugli, O.W., Qanadli, S.D., & Perrier, A. (2007). Should pulmonary embolism be suspected in exacerbation of chronic obstructive pulmonary disease? *Thorax, 62*(2), 121 – 125.

Sáenz-Lechón, N., Osma-Ruiz, V., Godino-Llorente, J.I., Blanco-Velasco, M., Cruz-Roldán, F., & Arias-Londoño, J.D. (2008). Effects of audio compression in automatic detection of voice pathologies. *IEEE transactions on bio-medical engineering, 55*(12), 2831 – 2835.

Sataloff, R.T., & Hawkshaw, M.J. (2014). Common medical diagnoses and treatments in patients with voice disorders. In J.S. Rubin, R.T. Sataloff, & G.W. Korovin (Eds.), *Diagnosis and Treatment of Voice Disorders* (pp. 371 – 388). San Diego, USA: Plural Publishing.

Shastry, A., Balasubramanium, R.K., & Acharya, P.R. (2014). Voice analysis in individuals with chronic obstructive pulmonary disease. *International Journal of Phonosurgery & Laryngology, 4*(2), 45 – 49.

Shrivastava, P., Tripathi, N., Singh, B.K. (2018). Assessment of respiratory disorders using speech parameters. *International Journal on Future Revolution in Computer Science & Communication Engineering, 4*(2), 461 – 464.

Silva, D.G., Oliveira, L.C., & Andrea, M. (2009). Jitter estimation algorithms for detection of pathological voices. *EURASIP Journal on Advances in Signal Processing,* e567875.

Simberg, S., Udd, H., & Santtila, P. (2015). Gender differences in the prevalence of vocal symptoms in smokers. *Journal of Voice, 29*(5), 588 – 591.

Soriano, J.B., Zielinksi, J., & Price, D. (2009). Screening for and early detection of chronic obstructive pulmonary disease. *Lancet, 374,* 721 – 732.

Tehrany, R., Barney, A., & Bruton, A. Differentiating between health and respiratory disease using speech breathing pattern analysis. *European Respiratory Journal, 23*(44), 58.

Teixeira, J.P., & Gonçalves, A. (2016). Algorithm for jitter and shimmer measurement in pathological voices. *Procedia Computer Science, 100,* 271 – 279.

Teixeira, J.P., Oliveira, C., & Lopes, C. (2013). Vocal Acoustic Analysis - jitter, shimmer and HNR parameters. *Procedia Technology, 9,* 1112 – 1122.

Theodoros, D., Russell, T.G., Hill, A., Cahill, L., & Clark, K. (2003). Assessment of motor speech disorders online: a pilot study. *Journal of Telemedicine and Telecare, 9*(2), 66 – 68.

Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders, 23*(2), 309 – 322.

Trappenburg, J.C., Monninkhof, E.M., Bourbeau, J., Troosters, T., Schrijvers, A.J., Verheij, T.J., & Lammers, J.W. (2011). Effect of an action plan with ongoing support by a case manager on exacerbation-related outcome in patient with COPD: a multicentre randomised controlled trial. *Thorax*, *66*(11), 977 – 984.

Van de Weijer, J.C., & Slis, I.H. (1991). Nasaliteitsmeting met de nasometer. *Logopedie en Foniatrie*, *63*, 97 – 101.

Van Eerten, L. (2007). Over het Corpus Gesproken Nederlands. *Nederlandse Taalkunde, 12*(3), 194 – 215.

Van Son, R.J.J.H., & Pols. L.W. (1996). An acoustic profile of consonant reduction. *Conference: the 4$^{th}$ International Conference on Spoken Language Processing.* Philadelphia, PA.

Vestal, L., Smith-Olinde, L., Hicks, G., Hutton, T., & Hart, J. (2006). Efficacy of language assessment in Alzheimer's disease: comparing in-person examination and telemedicine. *Clinical Interventions in Aging, 1*(4), 467 – 471.

Vieira, M.N. (2002). On the influence of laryngeal pathologies on acoustic and electroglottographic jitter measures. *The Journal of the Acoustic Society of America*, *111*(2), e1045.

Vincent, I., & Gilbert, H.R. (2011). The effects of cigarette smoking on the female voice. *Logopedics Phoniatrics Vocology*, *37*(1), 22 – 32.

Vyshedskiy, A., & Murphy, R. (2016). Acoustic biomarkers of chronic obstructive lung disease. *Research Ideas and Outcomes, 2*, e9173.

Wang, Y-T., Green, J.R., Nip, I.S.B., Kent, R.D., & Kent, J.F. (2010). Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatrica et Logopaedica, 62*, 297 – 302.

Wang, J., Green, J.R., Samal, A., & Yunusova, Y. (2013). Articulatory distinctiveness of vowels and consonants: a data-driven approach. *Journal of Speech Language and Hearing Research*, *56*(5), 1539 – 1551.

Wiechern, K.A., Liberty, P., & Pattemore, E. (2018). The effects of asthma on breathing during reading aloud. *Speech, Language and Hearing*, *21*(1), 30 – 40.

Wormald, R.N., Moran, R.J., Reilly, R.B., & Lacy, P.D. (2008). Performance of an automated, remote system to detect vocal fold paralysis. *Annals of Otology, Rhinology & Laryngology, 117*(11), 834 – 838.

Woubie, A., Luque, J., & Hernando, J. (2014). Jitter and shimmer measurements for speaker diarization. *IberSPEECH 2014,* 19 – 21.

Xue, S.A., & Lower, A. (2010). Acoustic fidelity of internet bandwidths for measures used in speech and voice disorders. *The Journal of the Acoustical Society of America, 128*(3), e1366.

Yap, T.F., Ambikairajah, E., Epps, J., & Choi, E.H.C. (2015). Voice source under cognitive load. Effects and classification. *Speech Communication, 72*, 74 – 95.

Young, L.B., Foster, L., Silander, A., & Wakefield, B.J. (2011). Home telehealth: patient satisfaction, program functions, and challenges for the care coordinator. *Journal of Gerontological Nursing*, *37*(11), 38 – 46.

Yunusova, Y., Graham, N.L., Shellikeri, S., Phuong, K., Kulkarni, M. Rochon, E., … Green, J.R. (2016). Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). *PLoS ONE, 11*(1), e0147573.

Zealouk, O., Satori, H., Mohamed, H., & Laaidi, N. (2018). Vocal parameters analysis of smoker using Amazigh language. *International Journal of Speech Technology*, *21*(1), 85 – 91.

## Appendix 1: Protocol for preparation and annotation

The materials for this study can be subdivided into five different types:

i. Sustained vowel [a:] by COPD speakers during exacerbation;
ii. Sustained vowel [a:] by COPD speakers during stable COPD;
iii. Storytelling of ‘De Koning’ by COPD speakers during exacerbation;
iv. Storytelling of ‘De Koning’ by COPD speakers in stable condition;
v. Storytelling of ‘Papa en Marloes’ by healthy speakers.

**Type i: Sustained vowels**
The sustained vowels were recorded shortly before or after reading aloud ‘De Koning’. The vowels were isolated and saved as a separate recording, but this type of speech did not require further annotation before analysis.

**Type ii, iii, iv and v: Storytelling of COPD speakers and healthy speakers**
Speech type ii, iii, iv and v were prepared and annotated following the same steps:

1. Each file was clipped in order to remove the instructions preceding the reading task and the commentary after the reading task.
2. The files were then converted using Praat, changing the sample size of 44 kHz to 16 kHz, to prepare the files for forced alignment.
3. The transcript of each file was then added in the first interval tier ‘transcript’ to prepare the files for forced alignment.
4. Each file was processed using the forced aligner according to its manual (Xue, De Reus, & Merkus, 2019; see Appendix 4), creating two additional interval tiers ‘wordsegmentation’ and ‘phoneticsegmentation’ containing the segmentations of the transcripts on word and phoneme level (in SAMPA).
5. These segmentations on both levels were then manually checked and corrected.
6. After finalizing the segmentation process, three more tiers were manually added to each file for the annotation process. The first (interval) tier ‘respiratory’ was added for respiratory-related annotations, such as ‘inhalation’, ‘exhalation’, ‘pause’. This way, information about the respiratory pattern could be obtained. The second (point) tier ‘speakernoise’ was added for voice quality-related annotations, such as ‘cough’, ‘creaky’, ‘slimy’, to obtain information regarding the voice quality of the speaker. A point tier was preferred over an interval tier, because the study focused on the number of occurrences. The third (interval) tier ‘commentary’ was added to clarify the speaker noises or to keep a record of additional observations. The respiratory tier and the speakernoise tier were annotated separately.
7. The first tier to be annotated was the respiratory tier. Each silence interval was studied to determine if the speaker was inhaling (in), exhaling (ex) or if there was an audible pause (p). This way, the transcript could be subdivided into different ‘breathing groups’. The commentary tier was used to specify the annotations.
8. The second tier to be annotated was the speakernoise tier. The fragment was replayed to obtain information about the condition of the voice. Audible disturbances, such as

hoarseness, creakiness and sliminess of the voice, were addressed using the speakernoise tier to indicate the occurrence and the commentary tier to describe the nature of the disturbance.

9. The fragment was then replayed a final time to check the annotations.

## Appendix 2: Analyses per audio fragment

This appendix contains the transcripts of the read text fragments per speaker. Patient 1 to 11 refer to the 11 COPD patients, whereas healthy 1 to 5 refer to the 5 healthy control speakers. The audio files were annotated according to the protocol in Appendix 1. Based on these annotations, the number of breath groups was determined. Each analysis shows the transcript of the text fragment per breath group, the number of syllables, the number of breath groups, the number of in- and exhalations, the duration and the number of non-linguistic in- and exhalations. In addition, each analysis shows the number of syllables per breath group, the number of in- and exhalations per syllable, the number of non-linguistic in- and exhalations per syllable and the speaking rate. The transcripts contain several abbreviations, namely [p] for 'pause', [in] for 'inhale' and [ex] for 'exhale'.

### Patient 1

**Exacerbation**

| **Transcript of text fragment per breath group** | **Manual calculations** |
| --- | --- |
| 1. [in] Het was najaar en de zon scheen<br>2. [in] De grote loofboom was in bloei<br>3. [in] De fonteinen sprongen fonkelend omhoog<br>4. [in] en goudvissen zo groot als dolfijnen schoten door het glinsterende water | Number of breath groups: 18 |
| 5. [in] Er speelde een orkest in het prieeltje<br>6. [in] en [p] zeven danseressen [p] dansten om de put<br>7. [in] Overal stonden goochelaars te goochelen<br>8. [in] en een buikspreker sprak buik | Number of inhalations: 18<br>Number of exhalations: 21 (incl. speech) |
| 9. [in] op het pad [ex]<br>10. [in] Gebeurde er [p] maar wat [p] zei de koning [p] Ik word Ik word gek [p] van de stilte in de kop | Number of non-linguistic inhalations: 1 (9)<br>Number of non-linguistic exhalations: 2 (12, 15) |
| 11. [in] Toen riep de koning zijn minister van Plezier en Vermaak<br>12. [in] Hij zei [ex] bedenk eens iets nieuws alsjeblieft<br>13. [in] Wat vindt u van vuurwerk vroeg de minister<br>14. [in] Even later schoten de raketten pijlen bollen omhoog<br>15. [in] en tot slot vloog het [ex] hele paleis de de de in de lucht<br>16. [in] Vervelend zei de koning toen het afgelopen was.<br>17. [in] Kun je iets grappigs bedenken<br>18. [in] En rap | Total duration: 59.37 seconds<br>Number of syllables: 206 |

## Patient 1

**Stable**

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] De minister had nog één tip ik zeg<br>2. [in] Nee Ik zelf<br>3. [in] schaterde hij<br>4. [in] terwijl hij in het [ex] riet dook. Hij trok zijn jas uit<br>5. [in] ging op zijn hoofd staan en zoemde een lied<br>6. [in] Dat is evengoed flauw [ex] zei de koning [p] toen het voorbij was<br>7. [in] ik wil een huis van glas met kozijnen van diamant<br>8. [in] en een gouden sieraad<br>9. [in] En als het af is zal ik het stukslaan met een stok<br>10. [in] Alles [ex] de tafels de stoelen de kussens de bedden<br>11. [in] de kolen in de kelder en zelfs het tafelkleed waren van glas<br>12. [in] en de ramen blonken de [p] als diamant<br>13. [in] [ex] En toen het af was liep de koning erin rond [p] en sloeg het stuk [p] met een stok<br>14. [in] Toen alles kapot was [ex] en hij iets nieuws wilde bedenken<br>15. [in] zei de minister Halt Het geld is op [ex]<br>16. [in] Kunnen we niet iets [p] met het paleis [ex]<br>17. [in] dacht de koning [ex] Ik heb het bevel<br>18. [in] Verkoop het paleis [p] Het paleis<br>19. [in] is in de lucht [p] gevlogen koning [ex]<br>20. [in] Dan wil ik de bromtol nog even zie [p] eens zien [ex]<br>21. [in] Want het was het beste maar de minister weigerde</td><td>Number of breath groups: 21</td></tr>
<tr><td>Number of inhalations: 21<br>Number of exhalations: 31 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 8 (3, 4, 5, 8, 11, 15, 17, 19)<br>Number of non-linguistic exhalations: 5 (4, 6, 10, 14, 20)</td></tr>
<tr><td>Total duration: 64.68 seconds<br>Number of syllables: 234</td></tr>
</table>

## Patient 2

**Exacerbation**

| Transcript of text fragment per breath group | Manual calculations |
| --- | --- |
| 1. [in] Het was najaar de zon scheen<br>2. [in] De grote loofboom was in bloei<br>3. [in] de fonteinen sprongen<br>4. [in] fonkelend omhoog<br>5. [in] en goudvissen zo groot als dolfijnen schoten<br>6. [in] door het glinsterende water<br>7. [in] Er speelt een orkest [ex] in het prieeltje<br>8. [in] en zeven danseressen dansten [p] om de put<br>9. [in] Overal stonden goochelaars te goochelen<br>10. [in] en een buikspreker sprak buik<br>11. [in] op het pad<br>12. [in] Gebeurde er wat zei de koning<br>13. [in] Ik word gek van al de stilte in mijn kop<br>14. [in] Toen riep de koning [p] zijn minister<br>15. [in] van Plezier en Vermaak. Hij zei<br>16. [in] Bedenk eens iets nieuws [p] alsjeblieft<br>17. [in] Wat vindt u van [ex] vuurwerk<br>18. [in] vroeg de minister<br>19. [in] Even later schoten de raketten en [p] pijlen en bollen omhoog<br>20. [in] en tot [ex] slot vloog<br>21. [in] het hele paleis in de lucht<br>22. [in] Vervelend zei de koning<br>23. [in] Toen [ex] het afgelopen was<br>24. [in] Kun je nog iets grappigs bedenken [p] en rap<br>25. [in] Een cursus koning. [p] Lijkt u dat wat | Number of breath groups: 25<br><br>Number of inhalations: 25<br>Number of exhalations: 29 (incl. speech)<br><br>Number of non-linguistic inhalations: 7 (4, 6, 11, 15, 18, 21, 23)<br>Number of non-linguistic exhalations: 4 (7, 17, 20, 23)<br><br>Total duration: 67.17 seconds<br>Number of syllables: 209 |

## Patient 2

**Stable**

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] Alles [p] de tafels de stoelen [p] de kussens de bedden<br>2. [in] de kolen in de kelder [p] en zelfs tafelkleed wa-<br>3. [in] waren van glas<br>4. [in] en de ramen blonken [p] als diamant<br>5. [in] En toen was het af [p] was<br>6. [in] liep de koning erin<br>7. [in] rond en sloeg [p] het stuk<br>8. [in] met een stok<br>9. [in] Toen alles kapot was [ex] en hij iets [p] iets niets wilde bedenken [ex] zei de minister Halt<br>10. [in] Het geld is op<br>11. [in] Kunnen we niet iets anders het paleis<br>12. [in] dacht de koning [p] Hij gaf het bevel<br>13. [in] Verkoop het paleis [ex] Het paleis [p] het paleis is in [p] het paleis [p] is in de vlucht gevlogen<br>14. [in] koning<br>15. [in] Dan wil ik [ex] de bromtol nog eens zien [p] want dat was het beste<br>16. [in] maar de minister weigerde<br>17. [in] Toen begreep de koning [ex] dat hij [ex] arm was [p] en hij slok<br>18. [in] Uit wanhoop verliet hij [ex] zijn eigen rijk [p] onze [p] en onze koning<br>19. [in] kwam menig [p] man tegen<br>20. [in] De mensen hadden [ex] hun schouders [p] de mensen halen hun schouders op en zeiden<br>21. [in] Wat zou hij hebben doorstaan [p] Acht stink</td><td>Number of breath groups: 21</td></tr>
<tr><td>Number of inhalations: 21<br>Number of exhalations: 25 (incl.speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 9 (2, 3, 6, 7, 8, 10, 12, 14, 19)<br>Number of non-linguistic exhalations: 7 (9, 9, 13, 15, 17, 17, 18, 20)</td></tr>
<tr><td>Total duration: 76.014 seconds<br>Number of syllables: 218</td></tr>
</table>

## Patient 3

**Exacerbation**

| Transcript of text fragment per breath group | Manual Calculations |
|---|---|
| 1. [in] Het was najaar<br>2. [in] en de zon scheen<br>3. [in] De grote loofboom was in bloei [p] De fonteinen sprongen fonkelend omhoog<br>4. [in] en goudvissen zo groot als dolfijnen<br>5. [in] schoten door het glinsterende water<br>6. [in] Er speelde een orkest in het prieeltje en zeven danseressen<br>7. [in] dansten om de put<br>8. [in] Overal stonden goochelaars te goochelen<br>9. [in] en een buikspreker sprak [p] buik op het pad Gebeurde er maar wat zei de koning<br>10. [in] Ik word gek van die [p] van [p] die stilte in m’n kop Toen riep de koning zijn minister van Plezier en Vermaak<br>11. [in] Hij zei [p] bedenk eens iets nieuws alstublieft<br>12. [in] Wat vindt u van<br>13. [in] vuurwerk vroeg de minister<br>14. [in] Even later schoten de raketten pijlen boven omhoog<br>15. [in] en tot slot vloog het hele paleis in de lucht<br>16. [in] Vervelend zei de koning toen het afgelopen was<br>17. [in] Kun je niet iets grappigs bedenken en rap<br>18. [in] Een<br>19. [in] circus koning lijkt u dat wat | Number of breath groups: 19<br><br>Number of inhalations: 19<br>Number of exhalations: 19 (incl. speech)<br><br>Number of non-linguistic inhalations: 5 (2, 5, 7, 13, 19)<br>Number of non-linguistic exhalations: 0<br><br>Total duration: 59.296 seconds<br>Number of syllables: 212 |

## Patient 3

**Stable**

| **Transcript of text fragment per breath group** | **Manual calculations** |
|---|---|
| 1. [in] De volgende dag was er een dromedaris [ex] die muziek speelde op zijn fluit en er waren zes<br>2. [in] loslopende olifanten die iets deden [p] wat andere olifanten [p] nooit doen | Number of breath groups: 12 |
| 3. [in] en het [p] circus was compleet met het [p] gespikkelde beest die een brulaap aan zijn staart had hangen<br>4. [in] De koning zat op de voorste rij en vond het eentonig | Number of inhalations: 12<br>Number of exhalations: 14 (incl. speech) |
| 5. [in] Dat hebben we ook weer achter de rug zei hij toen het uit was<br>6. [in] kun je niet iets bedenken waarvan je zegt is dat nou zo vrolijk | Number of non-linguistic inhalations: 1 (2)<br>Number of non-linguistic exhalations: 2 (1, 11) |
| 7. [in] De minis had de minister had nog één tip.<br>8. [in] Ik zelf schaterde hij terwijl [p] hij in het riet dook<br>9. [in] Hij trok zijn jas uit, ging op zijn hoofd staan en zoemde een lied<br>10. [in] Dat is evengoed flauw zei de koning toen het voorbij was<br>11. [in] ik wil een [p] huis [ex] van glas met kozijnen van diamant en een gouden sieraad<br>12. [in] En als het af is zal ik het [p] stukslaan met een stok | Total duration: 52.623 seconds<br>Number of syllables: 209 |

## Patient 4

**Exacerbation**

| **Transcript of text fragment per breath group** | **Manual calculations** |
|---|---|
| 1. [in] Het was najaar [ex]<br>2. [in] De [p] en de zon scheen de grote<br>3. [in] loofboom was [p] in bloei<br>4. [in] de [p] de fonteinen sprongen<br>5. [in] fonkelend omhoog [ex] de de goudvissen zo groot als een dolfijn<br>6. [in] schoten door<br>7. [in] hè het glinsterende water<br>8. [in] Er speelde een orkest in het<br>9. [in] prieeltje en zevende danseressen<br>10. [in] dansten [p] o [p] op de put [ex]<br>11. [in] Overal stonden goochelaars<br>12. [in] met te goochelen en<br>13. [in] de buik [p] spreker sprak [p] buik [p] op het pad [ex]<br>14. [in] Gebeurde er maar wat zei koning<br>15. [in] Ik word gek van al die stilte [ex]<br>16. [in] in mijn kop<br>17. [in] Toen riep de koning<br>18. [in] Ze zijn mister [ex]<br>19. [in] van plezier [p] en vermaak<br>20. [in] Hij zei bedankt [p] eens<br>21. [in] iets nieuws of bedenk ooit iets nieuws alsjeblieft<br>22. [in] Wat vindt u van vuurwerk vroeg de minister<br>23. [in] Even later schoten de raket [p] pijlen [p] en en bollen omhoog<br>24. [in] tot slot vloog<br>25. [in] het hele paleis in de lucht<br>26. [in] Vervelend zei de koning [p] toen het afgelopen was<br>27. [in] Kunt u niet [p] iets grappigs bedenken | Number of breath groups: 27<br><br>Number of inhalations: 27<br>Number of exhalations: 32 (incl. speech)<br><br>Number of non-linguistic inhalations: 14 (2, 3, 5, 6, 7, 9, 10, 12, 13, 16, 18, 19, 21, 25)<br>Number of non-linguistic exhalations: 4 (1, 5, 15, 18)<br><br>Total duration: 67.658 seconds<br>Number of syllables: 212 |

## Patient 4

**Stable**

**Audio file was not damaged, but patient was unable to execute task.**

## Patient 5

**Exacerbation**

| Transcript of text fragment per breath group | Manual calculations |
|---|---|
| 1. [in] Het was najaar en de zon scheen<br>2. [in] De grote loofboom was in bloei<br>3. [in] De fonteinen sprongen fonkelend omhoog<br>4. [in] en goudvissen [p] zo groot als dolfijnen | Number of breath groups: 15 |
| 5. [in] schoten door het glinsterende water<br>6. [in] Er speelde een [ex] orkest in een prieeltje<br>7. [in] en zeven danseressen dansten om de put<br>8. [in] Overal stonden goochelaars te goochelen | Number of inhalations: 15<br>Number of exhalations: 17 (incl. speech) |
| 9. [in] en een buikspreker sprak op<br>10. [in] sprak buik op het pad<br>11. [in] Gebeurde er maar wat zei de koning Ik word er gek van | Number of non-linguistic inhalations: 4 (5, 10, 12, 13)<br>Number of non-linguistic exhalations: 2 (6, 15) |
| 12. [in] en<br>13. [in] ik word er gek van die stilte in m'n kop<br>14. [in] Toen riep de koning zijn minister van plezier en vermaak<br>15. [in] Hij zei [ex] bedenk eens iets nieuws alsjeblieft | Total duration: 43.925 seconds<br>Number of syllables: 147 |

## Patient 5

**Stable**

**Audio file was damaged.**

# Patient 6

## Exacerbation

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] Het was najaar en de zon scheen<br>2. [in] De grote loofboom was in bloei<br>3. [in] De fonteinen sprongen fonkelend omhoog<br>4. [in] en goudvissen zo groot als dolfijnen<br>5. [in] schoten door het glinsterende water<br>6. [in] Er speelde een orkest in het prieeltje [ex] en zeven danseressen dansten om de put<br>7. [in] Overal stonden goochelaars te goochelen en een buikspreker sprak<br>8. [in] buik op het pad<br>9. [in] Gebeurde er maar wat zei de koning<br>10. [in] Ik word gek van de stilte in m’n kop<br>11. [in] Toen riep de koning zijn ministers<br>12. [in] van plezier en vermaak. Hij zei<br>13. [in] bedenk eens iets nieuws alsjeblieft<br>14. [in] Wat vindt u van vuurwerk vroeg de minister<br>15. [in] Even later schoten de raketten pijlen, bollen omhoog<br>16. [in] en tot slot<br>17. [in] vloog het hele paleis in de lucht<br>18. [in] Vervelend zei de koning<br>19. [in] toen het afgelopen was<br>20. [in] Kun je niet iets grappigs bedenken [p] en rap<br>21. [in] Een circus koning [p] lijkt u dat wat</td><td>Number of breath groups: 21</td></tr>
<tr><td>Number of inhalations: 21<br>Number of exhalations: 22 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 5 (5, 8, 12, 17, 19)<br>Number of non-linguistic exhalations: 1 (6)</td></tr>
<tr><td>Total duration: 46.046 seconds<br>Number of syllables: 210</td></tr>
</table>

## Patient 6

**Stable**

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] De minister had nog een tip [p] Ik zelf schaterde hij terwijl hij<br>2. [in] in het riet dook<br>3. [in] Hij trok zijn jas uit ging op zijn hoofd staan [p] en hij zoemde een lied<br>4. [in] Dat is even goed flauw zei de koning<br>5. [in] toen het voorbij was [p] Ik wil een huis van glas [p] met kozijnen van diamant en een gouden sieraad<br>6. [in] En als het af is [p] zal ik het stukslaan [p] met een stok<br>7. [in] Alles de tafels de stoelen de kussens de bedden de kolen in de kelder<br>8. [in] en zelfs het tafelkleed waren van glas<br>9. [in] en de ramen blonken als diamant<br>10. [in] En toen was het [ex] toen het af was<br>11. [in] liep de koning erin rond en sloeg stuk met een stok<br>12. [in] Toen alles kapot was [p] en hij niets meer wilde bedenken [p] zei de minister [p] halt het geld is op<br>13. [in] Kunnen we niets iets [p] met het paleis dacht de koning [p] Ik had het bevel [ex] verkoop het paleis<br>14. [in] Het paleis is in de lucht gevlogen koning<br>15. [in] Dat [p] wil [p] de bromtol nog eens zien<br>16. [in] want het is het beste<br>17. [in] maar de minister weigerde [p] Toen begreep de koning dat hij arm was en hij schrok [ex] Uit wanhoop verliet hij zijn eigen rijk<br>18. [in] En onze koning kwam menig man tegen<br>19. [in] De mensen haalden hun schouders op en zeiden<br>20. [in] wat zou je hebben doorstaan Ach stik.<br>21. [in] Nu gebeurde het [p] dat hij moe en angstig in het weiland een man zag Hij vroeg<br>22. [in] wat ben je aan het doen De man antwoordde met<br>23. [in] niets [p] alle dagen helemaal niets<br>24. [in] De man stond op en begon het koren te maaien<br>25. [in] De hongerige koning maaide eveneens mee met de man</td><td>Number of breath groups: 25</td></tr>
<tr><td>Number of inhalations: 25<br>Number of exhalations: 28 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 4 (2, 5, 8, 11)<br>Number of non-linguistic exhalations: 1 (10)</td></tr>
<tr><td>Total duration: 82.070 seconds<br>Number of syllables: 365</td></tr>
</table>

## Patient 7

**Exacerbation**

| Transcript of text fragment per breath group | Manual calculations |
|---|---|
| 1. [in] Het was najaar en de zon scheen<br>2. [in] De grote loofboom was in bloei<br>3. [in] De fonteinen sprongen fonkelend omhoog<br>4. [in] en goudvissen zo groot als dolfijnen<br>5. [in] schoten door het glinsterende water<br>6. [in] Er speelde een orkest in het prieeltje<br>7. [in] en zeven danseressen dansten om de put<br>8. [in] Overal stonden goochelaars te goochelen<br>9. [in] en een buikspreker sprak buik<br>10. [in] - [ex]<br>11. [in] op het pad<br>12. [in] - [ex]<br>13. [in] - [ex]<br>14. [in] Gebeurde er maar wat zei de koning<br>15. [in] Ik word gek van al die stilte in mijn kop<br>16. [in] Toen riep de koning zijn minister van Plezier<br>17. [in] en Vermaak<br>18. [in] Hij zei bedenk eens iets nieuws alsjeblieft<br>19. [in] Wat vindt u van vuurwerk vroeg de minister<br>20. [in] Even later schoten de raketten<br>21. [in] pijlen en bollen | Number of breath groups: 21<br><br>Number of inhalations: 21<br>Number of exhalations: 24 (incl. speech)<br><br>Number of non-linguistic inhalations: 5 (5, 11, 13, 17, 21)<br>Number of non-linguistic exhalations: 2 (10, 13)<br><br>Total duration: 60.406 seconds<br>Number of syllables: 164 |

## Patient 7

**Stable**

| Transcript of text fragment per breath group | Manual calculations |
| --- | --- |
| 1. [in] Een circus koning [p] Lijkt u dat wat<br>2. [in] De volgende dag was er een dromedaris<br>3. [in] die muziek speelde op zijn fluit<br>4. [in] en er waren zes loslopende olifanten<br>5. [in] die iets deden wat andere olifanten nooit doen<br>6. [in] en het circus was compleet<br>7. [in] met het gespikkelde beest<br>8. [in] die een brulaap aan zijn staart had hangen<br>9. [in] De koning zat op de voorste rij en vond het eentonig<br>10. [in] Dat hebben we ook weer achter de rug zei hij [p] toen het uit was<br>11. [in] kun je niet iets [p] bedenken [p] waarvan je zegt<br>12. [in] das nou écht vrolijk<br>13. [in] De minister had nog [p] één tip. [p] Ik zelf<br>14. [in] schaterde hij terwijl hij [p] het<br>15. [in] riet in ’t riet dook<br>16. [in] Hij trok zijn jas uit<br>17. [in] ging op zijn hoofd staan en zoemde een lied [ex]<br>18. [in] Das evengoed flauw zei de koning [p] toen het voorbij was<br>19. [in] ik wil een huis van glas<br>20. [in] met kozijnen van diamant<br>21. [in] en een gouden sieraad<br>22. [in] En als het af is zal ik [p] het stukslaan<br>23. [in] met een stok | Number of breath groups: 23<br><br>Number of inhalations: 21<br>Number of exhalations: 22 (incl. speech)<br><br>Number of non-linguistic inhalations: 10 (15, 17, 19, 20, 26, 27, 29, 32, 33, 35)<br>Number of non-linguistic exhalations: 0<br><br>Total duration: 64.241 seconds<br>Number of syllables: 212 |

## Patient 8

**Exacerbation**

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] Het was [ex] najaar en de zon scheen<br>2. [in] de grote loofboom was in bloei<br>3. [in] de fonteinen s [p] prongen fonkelend omhoog<br>4. [in] en goudvissen [p] zo groot als dolfijnen<br>5. [in] schoten door het glinsterende water<br>6. [in] Er speelde een orkest in het prieeltje<br>7. [in] en zeven danseressen dansten om de put<br>8. [in] Overal stonden goochelaars te goochelen<br>9. [in] en een buiksp [p] spreker sprak [p] buik op het pad<br>10. [in] Gebeurde er maar wat zei de koning<br>11. [in] Ik word gek van al die stilte in mijn kop<br>12. [in] Toen riep de koning zijn minister van Plezier en Vermaak<br>13. [in] Hij zei<br>14. [in] Ik Bedenk eens iets nieuws alsjeblieft<br>15. [in] Wat vindt u van vuurwerk vroeg de minister<br>16. [in] Even later schoten de raketten<br>17. [in] pijlen [p] en bollen omhoog<br>18. [in] en tot slot vloog het hele paleis in de lucht</td><td>Number of breath groups: 18</td></tr>
<tr><td>Number of inhalations: 18<br>Number of exhalations: 19 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 2 (5, 17)<br>Number of non-linguistic exhalations: 1 (1)</td></tr>
<tr><td>Total duration: 52.786 seconds<br>Number of syllables: 180</td></tr>
</table>

## Patient 8

**Stable**

| **Transcript of text fragment per breath group** | **Manual calculations** |
|---|---|
| 1. [in] De volgende dag was er een dromedaris<br>2. [in] die muziek speelde op zijn fluit [p] en er waren zes loslopende olifanten<br>3. [in] die iets deden wat andere olifanten nooit doen<br>4. [in] en het circus was compleet<br>5. [in] met het gespikkelde beest<br>6. [in] die een brulaap aan zijn staart had hangen<br>7. [in] De koning zat op de eer voorste rij en vond het<br>8. [in] Eentonig [ex]<br>9. [in] Dat hebben we ook weer achter de rug zei hij toen het uit was<br>10. [in] kun je niet iets bedenken waarvan je zegt<br>11. [in] dat is nou [p] écht vrolijk<br>12. [in] De minister had nog één tip<br>13. [in] Ik zelf! [p] schaterde hij<br>14. [in] terwijl [p] hij in [p] het riet dook<br>15. [in] Hij trok zijn jas uit ging op zijn hoofd staan en zoem<br>16. [in] de een lied | Number of breath groups: 16<br><br>Number of inhalations: 16<br>Number of exhalations: 17 (incl. speech)<br><br>Number of non-linguistic inhalations: 6 (2, 3, 5, 8, 14, 16)<br>Number of non-linguistic exhalations: 0<br><br>Total duration: 41.996 seconds<br>Number of syllables: 156 |

## Patient 9

**Exacerbation**

| **Transcript of text fragment per breath group** | **Manual calculations** |
|---|---|
| 1. [in] Het was najaar en de zon scheen [ex]<br>2. [in] - [ex]<br>3. [in] de grote loofboom was in bloei [ex]<br>4. [in] - [ex]<br>5. [in] de fonteinen sprongen fonkelend omhoog [ex]<br>6. [in] - [ex]<br>7. [in] En goudvissen [ex]<br>8. [in] - [ex]<br>9. [in] zo groot als dolfijnen schoten door het glinsterende water<br>10. [in] - [ex]<br>11. [in] Er speelde een orkest in het prieeltje [p] en zeven danseressen dansten om de put<br>12. [in] Overal stonden goochelaars te goochelen<br>13. [in] en een buikspreker sprak buik op het pad [ex]<br>14. [in] - [ex]<br>15. [in] Gebeurde er maar wat [ex] zei de koning<br>16. [in] Ik word gek van al die stilte in mijn kop<br>17. [in] - [ex]<br>18. [in] Toen riep de koning zijn minister van Plezier en Vermaak<br>19. [in] – [ex]<br>20. [in] Hij zei [ex] Bedenk eens iets nieuws alsjeblieft [ex] | Number of breath groups: 20<br><br>Number of inhalations: 20<br>Number of exhalations: 32 (incl. speech)<br><br>Number of non-linguistic inhalations: 9 (2, 3, 6, 8, 9, 10, 14, 17, 19)<br>Number of non-linguistic exhalations: 11 (2, 4, 6, 7, 8, 10, 14, 15, 17, 19, 20)<br><br>Total duration: 46.178 seconds<br>Number of syllables: 138 |

## Patient 9

**Stable**

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] Alles [p] de tafels de stoelen de kussens de bedden<br>2. [in] de kolen in de kelder [p] en zelfs het tafelkleed waren van glas<br>3. [in] en de ramen blonken als ‘n diamant<br>4. [in] En toen het af was liep de koning erin rond en sloeg het stuk met een stok<br>5. [in] Toen alles kapot was en hij iets nieuws wilde bedenken zei de minister<br>6. [in] Halt [ex] Het geld is op<br>7. [in] Kunnen we niet iets met het paleis [ex] dacht de koning<br>8. [in] Hij gaf het bevel [ex] Verkoop het paleis<br>9. [in] Het paleis is in de lucht gevlogen koning<br>10. [in] Dan wil ik de bromtol nog eens zien want dat was het beste [p] maar de minister weigerde<br>11. [in] Toen begreep de koning dat hij arm was en hij schrok<br>12. [in] Uit wanhoop verliet hij zijn eigen rijk<br>13. [in] En onze koning kwam menig man tegen</td><td>Number of breath groups: 13</td></tr>
<tr><td>Number of inhalations: 13<br>Number of exhalations: 16 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 1 (2)<br>Number of non-linguistic exhalations: 1 (6)</td></tr>
<tr><td>Total duration: 48.537 seconds<br>Number of syllables: 176</td></tr>
</table>

## Patient 10

**Exacerbation**

| **Transcript of text fragment per breath group** | **Manual calculations** |
|---|---|
| 1. [in] Het was najaar<br>2. [in] en de zon scheen [ex]<br>3. [in] de grote loofboom was in bloei [ex]<br>4. [in] de fonteinen sprongen fonkelend omhoog [ex]<br>5. [in] en goudvissen<br>6. [in] zo groot als dolfijnen<br>7. [in] schoten door het glinsterende water<br>8. [in] Er speelde een orkest in het prieeltje [ex]<br>9. [in] en zeven danseressen dansten om de put [ex]<br>10. [in] Overal stonden goochelaars te goochelen [ex]<br>11. [in] en een buikspreker [p] sprak buik op het pad [ex]<br>12. [in] Gebeurde er maar wat zei de koning [ex]<br>13. [in] Ik word gek van al die stilte in mijn kop [ex]<br>14. [in] Toen riep de koning zijn minister van Plezier en<br>15. [in] Vermaak [ex]<br>16. [in] Hij zei [ex]<br>17. [in] Bedenk eens iets nieuws alsjeblieft [ex]<br>18. [in] Wat vindt u van vuurwerk? [ex]<br>19. [in] vroeg de minister [ex] | Number of breath groups: 19<br><br>Number of inhalations: 19<br>Number of exhalations: 33 (incl. speech)<br><br>Number of non-linguistic inhalations: 5 (2, 5, 6, 15, 18)<br>Number of non-linguistic exhalations: 5 (4, 8, 10, 16, 18)<br><br>Total duration: 53.369 seconds<br>Number of syllables: 159 |

## Patient 10

**Stable**

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] Alles de tafels de stoelen<br>2. [in] de kussens [p] de bedden<br>3. [in] de kolen in de kelder [p] en zelfs het tafelkleed waren van glas<br>4. [in] en de ramen blonken als diamant<br>5. [in] En toen het af was [p] liep de koning erin rond [p] en sloeg het stuk met een stok<br>6. [in] Toen alles kapot was [p] en hij iets nieuws wilde bedenken<br>7. [in] zei de minister [p] Halt [p] Het geld is op<br>8. [in] Kunnen we niet iets met het paleis [p] dacht de koning [ex]<br>9. [in] Hij gaf het bevel<br>10. [in] Verkoop het paleis<br>11. [in] Het paleis is [p] in de lucht gevlogen, koning [ex]<br>12. [in] Dan wil ik de bromtol nog eens zien<br>13. [in] Want dat was het beste [p] maar de minister weigerde<br>14. [in] Toen begreep de koning<br>15. [in] dat hij arm was en hij schrok<br>16. [in] Uit wanhoop verliet hij zijn eigen rijk<br>17. [in] En onze koning [p] kwam menig man tegen De mensen haalden hun schouders op en zeiden<br>18. [in] Wat zou hij hebben doorstaan? [ex] Ach stik Nu gebeurde het dat hij moe en angstig<br>19. [in] In een weiland een man zag</td><td>Number of breath groups: 19</td></tr>
<tr><td>Number of inhalations: 19<br>Number of exhalations: 22 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 4 (2, 3, 15, 19)<br>Number of non-linguistic exhalations: 0</td></tr>
<tr><td>Total duration: 71.164 seconds<br>Number of syllables: 216</td></tr>
</table>

# Patient 11

### Exacerbation

| Transcript of text fragment per breath group | Mean calculations |
|---|---|
| 1. [in] Het was najaar en de zon scheen<br>2. [in] De grote loofboom was in bloei<br>3. [in] de fonteinen sprongen fonkelend omhoog en goudvissen zo groot als dolfijnen<br>4. [in] schoten door het glinsterende water<br>5. [in] Er speelde een orkest in het prieeltje [p] en zeven danseressen dansten om de put<br>6. [in] Overal stonden **de** goochelaars te goochelen<br>7. [in] en een buikspreker sprak buik op het pad<br>8. [in] Gebeurde er maar wat zei de koning [ex]<br>9. [in] Ik word gek van al die stilte in mijn kop<br>10. [in] Toen riep de koning zijn minister van Plezier en Vermaak<br>11. [in] Hij zei Bedenk eens iets nieuws alsjeblieft<br>12. [in] Wat vindt u van vuurwerk vroeg de minister<br>13. [in] Even later schoten de raketten pijlen en bollen omhoog<br>14. [in] en tot slot vloog het hele paleis in de lucht<br>15. [in] Vervelend zei de koning toen het afgelopen was<br>16. [in] Kun je niet iets grappigs bedenken<br>17. [in] En rap<br>18. [in] Een circus koning | Number of breath groups: 18<br><br>Number of inhalations: 18<br>Number of exhalations: 19 (incl. speech)<br><br>Number of non-linguistic inhalations: 1 (4)<br>Number of non-linguistic exhalations: 0<br><br>Total duration: 54.332 seconds<br>Number of syllables: 207 |

## Patient 11

**Stable**

| Transcript of text fragment per breath group | Manual calculations |
|---|---|
| 1. [in] De minister had nog één tip [p] Ik zelf schaterde hij terwijl hij in het riet dook<br>2. [in] Hij trok zijn jas uit ging op zijn hoofd staan en zoemde een lied<br>3. [in] Dat is evengoed flauw zei de koning toen het voorbij was<br>4. [in] ik wil een huis van glas met kozijnen van diamen van diamant en een gouden sieraad<br>5. [in] En als het af is zal ik het stukslaan met een stok<br>6. [in] Alles De tafels de stoelen de kussens de bedden de kolen in de kelder en zelfs het tafelkleed waren van glas<br>7. [in] en de ramen blonken als diamant<br>8. [in] En toen het af was liep de koning erin rond en sloeg het stuk met een stok<br>9. [in] Toen alles kapot was en hij iets nieuws wilde bedenken zei de minister<br>10. [in] Halt [p] Het geld is op<br>11. [in] Kunnen we niet iets met het paleis dacht de koning [p] Hij gaf het bevel Verkoop het paleis<br>12. [in] Het paleis [p] is in de lucht gevlogen koning [ex]<br>13. [in] Dan wil ik de bromtol nog eens zien want dat was het beste [p] maar de minister weigerde | Number of breath groups: 13<br><br>Number of inhalations: 14<br>Number of exhalations: 15 (incl. speech)<br><br>Number of non-linguistic inhalations: 0<br>Number of non-linguistic exhalations: 0<br><br>Total duration: 50.812 seconds<br>Number of syllables: 229 |

## Healthy 1

| Transcript of text fragment per breath group | Manual calculations |
|---|---|
| 1. [in] Papa en Marloes staan op het station<br>2. [in] Ze wachten op de trein<br>3. [in] Eerst hebben ze een kaartje gekocht maar er stond een hele lange rij dus dat duurde wel even<br>4. [in] Nu wachten ze tot de trein eraan komt<br>5. [in] Het is al vijf over drie dus 't duurt nog vier minuten<br>6. [in] Er staan nog veel meer mensen te wachten Marloes kijkt naar links In de verte ziet ze de trein al aankomen | Number of breath groups: 6 |
| | Number of inhalations: 6<br>Number of exhalations: 6 (incl. speech) |
| | Number of non-linguistic inhalations: 0<br>Number of non-linguistic exhalations: 0 |
| | Total duration: 26.111 seconds<br>Number of syllables: 93 |

## Healthy 2

<table>
<tr><th>Transcript of text fragment per breath group</th><th>Manual calculations</th></tr>
<tr><td rowspan="4">1. [in] Papa en Marloes staan op het station. Ze wachten op de trein<br>2. [in] Eerst hebben ze een kaartje gekocht maar er stond een hele lange rij dus dat duurde wel even<br>3. [in] Nu wachten ze tot de trein eraan komt. Het is al vijf over drie dus 't duurt nog vier minuten<br>4. [in] Er staan nog veel meer mensen te wachten Marloes kijkt naar links In de verte ziet ze de trein al aankomen</td><td>Number of breath groups: 4</td></tr>
<tr><td>Number of syllables: 93<br>Number of inhalations: 3<br>Number of exhalations: 3 (incl. speech)</td></tr>
<tr><td>Number of non-linguistic inhalations: 0<br>Number of non-linguistic exhalations: 0</td></tr>
<tr><td>Total duration: 21.63 seconds<br>Number of syllables: 93</td></tr>
</table>

## Healthy 3

| **Transcript of text fragment per breath group** | **Manual calculations** |
|---|---|
| 1. [in] Papa en Marloes staan op het station<br>2. [in] Ze wachten op de trein<br>3. [in] Eerst hebben ze een kaartje gekocht<br>4. [in] Er stond een hele lange rij dus dat duurde wel even<br>5. [in] Nu wachten ze tot de trein eraan komt<br>6. [in] Het is al vijf over drie dus 't duurt nog vier minuten<br>7. [in] Er staan nog veel meer mensen te wachten Marloes kijkt naar links In de verte ziet ze de trein al aankomen | Number of breath groups: 7 |
| | Number of syllables: 92<br>Number of inhalations: 7 |
| | Number of non-linguistic inhalations: 0<br>Number of non-linguistic exhalations: 0 |
| | Total duration: 25.04 seconds<br>Number of syllables: 92 |

## Healthy 4

| Transcript of text fragment per breath group | Manual calculations |
|---|---|
| 1. [in] Papa en Marloes staan op het station Ze wachten op de trein<br>2. [in] Eerst hebben ze een kaartje gekocht maar er stond een hele lange rij dus dat duurde wel even<br>3. [in] Nu wachten ze tot de trein eraan komt Het is al vijf over drie dus 't duurt nog vier minuten<br>4. [in] Er staan nog veel meer mensen te wachten Marloes kijkt naar links In de verte ziet ze de trein al aankomen | Number of breath groups: 4 |
| | Number of syllables: 93<br>Number of inhalations: 4<br>Number of exhalations: 4 (incl. speech) |
| | Number of non-linguistic inhalations: 0<br>Number of non-linguistic exhalations: 0 |
| | Total duration: 26.111 seconds<br>Number of syllables: 93 |

## Healthy 5

| Transcript of text fragment per breath group | Manual calculations |
|---|---|
| 1. [in] Papa en Marloes staan op het station Ze wachten op de trein. Eerst hebben ze een kaartje gekocht<br>2. [in] Er stond een hele lange rij dus dat duurde wel even<br>3. [in] Nu wachten ze tot de trein eraan komt<br>4. [in] Het is al vijf over drie dus 't duurt nog vier minuten Er staan nog veel meer mensen | Number of breath groups: 6 |
| | Number of inhalations: 4<br>Number of exhalations: 4 (incl. speech) |
| | Number of non-linguistic inhalations: 0<br>Number of non-linguistic exhalations: 0 |
| | Total duration: 24.97 seconds<br>Number of syllables: 72 |

# Appendix 3: Scripts

**Praat script (Kerkhoff, 2015)**

```
 FUNCTION: SCRIPT Calculation F0(min,max,mean,variability), Intensity
               Pitch variability, Formants, Centre of gravity
  AUTHOR: J. Kerkhoff

  DATE: maart 2015

  COPYRIGHT:  University of Nijmegen
            Department of Language and Speech
            Erasmusplein 1
            P.O. Box 9103, 6500 HD  Nijmegen
            The Netherlands.

form Analysis of labeled Intervals
  comment Analysis files:
    sentence input_directory C:\
    sentence input_files *.wav
    sentence result_file results.txt
  comment Which segments should be analyzed?
    positive tier_number 1
    sentence match_label_(interval) # (#:all intervals;*=not empty intervals)
    sentence begin_end_labels_(point) # - # (# - # = segments between all labels)
  comment Segment analysis options:
    boolean Duration yes
    boolean Pitch_variability no
            boolean Pitch_Min_Max_Mean no
    boolean Intensity_Min_Max_Mean no
    boolean Formants no
    boolean Centre_of_gravity no
  comment Settings Pitch Analysis
    real Time_step 0.0
    positive Pitch_floor_(Hz) 75
    positive Pitch_ceiling_(Hz) 600
  comment Settings Formant Analysis
    integer Max_number_of_formants 5
    positive Maximum_formant_(Hz) 5500 (= adult female)
    positive Window_length_(s) 0.025
    positive Preemphasis_from_(Hz) 50
endform

########### Preparations #############
semitones = 0
kill_octave_jumps = 0
smooth_pitch = 0
```

```
interpolate_pitch = 0
# determine directory:
if (right$(input_directory$,1) = "\")
  directory$ = input_directory$
else
  directory$ = input_directory$ + "\"
endif

# remove existing result file
resultFile$ = directory$ + result_file$
if fileReadable(resultFile$)
   pause Overwrite result file: continue
   filedelete 'resultFile$'
endif

# interval match label
interval$ = extractWord$("'match_label$'","")
startLabel$ = extractWord$("'begin_end_labels$'","")
endLabel$ = extractWord$("'begin_end_labels$'","-")

########### End Preparations #############

# clear info screen
clearinfo

##### Write settings to message screen #####
printline
printline Settings:
printline Input files 'tab$' 'directory$''input_files$'
printline Result file 'tab$' 'resultFile$'
printline Tier number 'tab$' 'tier_number'
printline label interval 'tab$' 'interval$'
printline labels 'tab$' 'startLabel$' - 'endLabel$'
printline
printline Results:
##### Write settings to result file #####
fileappend "'resultFile$'" 'directory$''input_files$', Tier number 'tier_number''newline$'
#fileappend "'resultFile$'" Results:'newline$'
##### End writing settings #################

########### Make list of audio files #############
Create Strings as file list... list 'directory$''input_files$'
numberOfFiles = Get number of strings

##### Start reading and labeling for all files in the list #####
for ifile to numberOfFiles
```

```
    select Strings list
    fileName$ = Get string... ifile
    Read from file... 'directory$''fileName$'

    ### search last sound/grid file
    select all
    soundname$ = selected$ ("Sound",-1)

    Read from file... 'directory$''soundname$'_checked.tg
    select TextGrid 'soundname$'_checked
    plus Sound 'soundname$'


    ##### place or move existing labels #####
    #Edit
    #if 'pause'
    #  pause Modify labels
    #endif

    ##### save TextGrid file #####
    #select TextGrid 'soundname$'_checked
    #Write to text file... 'directory$''soundname$'.TextGrid

    ### Write info ###
    printline
    printline File: 'soundname$'
    fileappend "'resultFile$'" 'newline$''soundname$'

    ###### Analyze audio file, calculate praat objects ############
    call AnalyzeAudio
    call ComputeAnalysis 'tier_number'

    ### remove used files ###
    select all
    minus Strings list
    Remove
endfor

##### End for all files in the list #####

### Write info screen ###
filedelete 'directory$'analysis.info
fappendinfo 'directory$'analysis.info

### remove rest
select Strings list
Remove
```

```
#---------------Procedures------------------#
##############################################
procedure AnalyzeAudio

   # Initilaize counters
    if (duration)
      interval_counter = 0
      interval_durations = 0
    endif
    if (pitch_variability or pitch_Min_Max_Mean)
      select Sound 'soundname$'
      To Pitch... 'time_step' 'pitch_floor' 'pitch_ceiling'
   # for calculating curve-mean
      Down to PitchTier
   # Convert to semitones
      if (semitones)
        Formula... hertzToSemitones(self)
      endif
   # Remove octave jumps
      if (kill_octave_jumps)
        Kill octave jumps
        Rename... dummy
        select Pitch 'soundname$'
        Remove
        select Pitch dummy
        Rename... 'soundname$'
      endif
   # Smooth Pitch contour
      if (smooth_pitch)
        Smooth: 10
        Rename... dummy
        select Pitch 'soundname$'
        Remove
        select Pitch dummy
        Rename... 'soundname$'
      endif
   # Interpolate Pitch contour
      if (interpolate_pitch)
        Interpolate
        Rename... dummy
        select Pitch 'soundname$'
        Remove
        select Pitch dummy
        Rename... 'soundname$'
      endif
```

```
  # Calculate voided and unvoiced parts
#     To PointProcess
#     To TextGrid (vuv): 0.02, 0.01

   endif
   if intensity_Min_Max_Mean
      select Sound 'soundname$'
      To Intensity... 'pitch_floor' 0 yes
   endif
   if formants
      select Sound 'soundname$'
      To Formant (burg)... time_step max_number_of_formants maximum_formant window_length
preemphasis_from
   endif
endproc


################################################################
##### Calculate Pitch/Intensity/formants from interval tier
################################################################

procedure ComputeAnalysis tierNumber

  ##### check tier type #####
  select TextGrid 'soundname$'_checked
  tier = Is interval tier... 'tier_number'
   if (tier)
     tierType$ = "interval"
   else
     tierType$ = "point"
   endif

  ##### Analyze interval tier #####
  select TextGrid 'soundname$'_checked

  ##### if INTERVAL tier #####
   if (tierType$ = "interval") or (tierType$ = "Interval")
      numberOfIntervals = Get number of intervals... 'tierNumber'
      for interval from 2 to (numberOfIntervals-1)
         select TextGrid 'soundname$'_checked
         label$ = Get label of interval... 'tierNumber' interval
         labelSegment$ = label$
         index = rindex(label$,interval$)
         if (index>0) or (interval$ = "#") or ((interval$ = "*") and (label$ <> ""))
            timeBegin = Get starting point... 'tierNumber' interval
                    timeEnd = Get end point... 'tierNumber' interval
```

```
      call AnalysisFunctions timeBegin timeEnd
    endif
  endfor
endif

##### if POINT tier #####
if (tierType$ = "point") or (tierType$ = "Point")
  numpoints = Get number of points... 'tierNumber'

  # analyze all labeled segments
  if (startLabel$ = "#") and (endLabel$ = "#")
    for point from 1 to (numpoints-1)
      select TextGrid 'soundname$'_checked
      timeBegin = Get time of point... 'tierNumber' point
      timeEnd = Get time of point... 'tierNumber' (point+1)
      label1$ = Get label of point... 'tierNumber' (point)
      label2$ = Get label of point... 'tierNumber' (point+1)
      labelSegment$ = label1$ + "-" + label2$
      if labelSegment$ = ""
        labelSegment$ = "-"
      endif
      call AnalysisFunctions timeBegin timeEnd
    endfor
  endif

  # analyze all matched labeled segments
  if (startLabel$ <> "#") and (endLabel$ <> "#")
    point = 1
    while point < (numpoints)
      firstFound = 0
      secondFound = 0
      select TextGrid 'soundname$'_checked
      label$ = Get label of point... 'tierNumber' point
      index = rindex(label$,startLabel$)
      while (index=0) and (point < numpoints)
        point = point + 1
        label$ = Get label of point... 'tierNumber' point
        index = rindex(label$,startLabel$)
      endwhile
      if (index>0)
       firstFound = 1
       labelSegment$ = label$
       timeBegin = Get time of point... 'tierNumber' point
      endif
      while (point < numpoints)
        point = point + 1
        label$ = Get label of point... 'tierNumber' point
```

```
          index = rindex(label$,endLabel$)
          while (index=0) and (point < numpoints)
            point = point + 1
            label$ = Get label of point... 'tierNumber' point
            index = rindex(label$,endLabel$)
          endwhile
          if (index=0) and (point = numpoints)
            printline Endlabel -'endLabel$'- not found......, endlabel = last label
          endif
          if (index>0)
           secondFound = 1
           labelSegment$ = labelSegment$ + "-" + label$
           timeEnd = Get time of point... 'tierNumber' point
          endif
        endwhile

        if firstFound and secondFound
          if labelSegment$ = ""
            labelSegment$ = "-"
          endif
          call AnalysisFunctions timeBegin timeEnd
        endif
        point = point + 1
      endwhile
    endif
  endif
  if (duration)
   ##### write results to console (in ms) #####
   printline Total intervals 'interval_counter:0' Total duration 'interval_durations:0'
   ##### write results to file (in ms/Hz) #####
   fileappend "'resultFile$'"  tot_int 'interval_counter:0' tot_dur 'interval_durations:0'
   ##### End writing #####
  endif
endproc


##################################################################
##### Check analysis settings
##################################################################
procedure AnalysisFunctions timeBegin timeEnd

#  ##### write label to console #####
#  printline Segment: 'labelSegment$'
#  ##### write label to file #####
#  fileappend "'resultFile$'"  'labelSegment$'
#  ##### End writing #####
```

```
  if duration
    call Durationvalues timeBegin timeEnd
  endif
  if pitch_Min_Max_Mean
    call PitchMinMaxMean timeBegin timeEnd
  endif
  if pitch_variability
    call PitchVariability timeBegin timeEnd
  endif
  if intensity_Min_Max_Mean
    call IntensityMinMaxMean timeBegin timeEnd
  endif
  if formants
     call Formantvalues timeBegin timeEnd
  endif
  if centre_of_gravity
    call Gravityvalue timeBegin timeEnd
  endif
endproc

#################################################################
##### Calculate Duration of an interval
#################################################################
procedure Durationvalues timeBegin timeEnd

   intervalDuration = (timeEnd - timeBegin)*1000
   interval_counter = interval_counter + 1
   interval_durations = interval_durations + intervalDuration
   ##### write results to console (in ms) #####
   printline Duration 'labelSegment$' 'intervalDuration:0'
   ##### write results to file (in ms/Hz) #####
   fileappend "'resultFile$'"  'labelSegment$' 'intervalDuration:0'
   ##### End writing #####
endproc

#################################################################
##### Calculate Pitch of all labels
#################################################################
procedure PitchAllLabels

   select TextGrid 'soundname$'
   ##### if INTERVAL tier #####
   if (tierType$ = "interval") or (tierType$ = "Interval")
      numberOfIntervals = Get number of intervals... 'tierNumber'
      for interval from 2 to (numberOfIntervals-1)
         select TextGrid 'soundname$'_checked
         label$ = Get label of interval... 'tierNumber' interval
```

```
        timeBegin = Get starting point... 'tierNumber' interval
        timeEnd = Get end point... 'tierNumber' interval
        select Pitch 'soundname$'
        pitchBegin = Get value at time... 'timeBegin' Hertz linear
        if pitchBegin = undefined
          pitchBegin = 0
        endif
        pitchEnd = Get value at time... 'timeEnd' Hertz linear
        if pitchEnd = undefined
          pitchEnd = 0
        endif
        timeBegin = timeBegin * 1000
        timeEnd = timeEnd * 1000
        ##### write results to console (in ms) #####
        printline 'label$': 'timeBegin:0' 'pitchBegin:2' 'timeEnd:0' 'pitchEnd:2'
        ##### write results to file (in ms/Hz) #####
        fileappend "'resultFile$'"  'label$' 'timeBegin:0' 'pitchBegin:2' 'timeEnd:0' 'pitchEnd:2'
        ##### End writing #####
    endfor
   endif

  ##### if POINT tier #####
  if (tierType$ = "point") or (tierType$ = "Point")
    numberOfSegments = Get number of points... 'tierNumber'
    for point from 1 to numberOfSegments
        select TextGrid 'soundname$'_checked
        label$ = Get label of point... 'tierNumber' point
        timeValue =  Get time of point... 'tierNumber' point
        select Pitch 'soundname$'
        pitchValue = Get value at time... 'timeValue' Hertz linear
        if pitchValue = undefined
          pitchValue = 0
        endif
        timeValue = timeValue * 1000
        ##### write results to console (in ms) #####
        printline 'label$': 'timeValue:0' 'pitchValue:2'
        ##### write results to file (in ms/Hz) #####
        fileappend "'resultFile$'"  'label$' 'timeValue:0' 'pitchValue:2'
        ##### End writing #####
    endfor
  endif
endproc

##############################################
procedure PitchVariability timeBegin timeEnd

   select Sound 'soundname$'
```

```
    Extract part: 'timeBegin', 'timeEnd', "rectangular", 1, "no"
    To Pitch... 0 'pitch_floor' 'pitch_ceiling'
    meanSlope = Get mean absolute slope: "Hertz"

    select Sound 'soundname$'_part
    plus Pitch 'soundname$'_part
    Remove
   ##### write results to console (in Hz/sec) #####
   printline Pitch variability 'labelSegment$' 'meanSlope:2'
   ##### write results to file (in Hz/sec) #####
   fileappend "'resultFile$'"  'labelSegment$' 'meanSlope:2'
   ##### End writing #####

endproc

################################################################
##### Calculate Pitch Min-Max-Mean of an interval
################################################################
procedure PitchMinMaxMean timeBegin timeEnd

   select Pitch 'soundname$'
   minPitch = Get minimum... timeBegin timeEnd Hertz Parabolic
   if minPitch = undefined
     minPitch = 0
   endif
   maxPitch = Get maximum... timeBegin timeEnd Hertz Parabolic
   if maxPitch = undefined
     maxPitch = 0
   endif
   meanPitch = Get mean... timeBegin timeEnd Hertz
   if meanPitch = undefined
     meanPitch = 0
   endif
   stdDevPitch = Get standard deviation: timeBegin, timeEnd, "Hertz"
   select PitchTier 'soundname$'
   meanPitchCurve = Get mean (curve)... timeBegin timeEnd

   ##### write results to console (in Hz) #####
#      printline Pitch (min, max, mean(points), mean(curves)) 'labelSegment$' 'minPitch:0' 'maxPitch:0' 'meanPitch:0' 'meanPitchCurve:0'
   printline Pitch (min, max, mean, stddev) 'labelSegment$' 'minPitch:2' 'maxPitch:2' 'meanPitch:2' 'stdDevPitch:2'
   ##### write results to file (in ms/Hz) #####
#      fileappend "'resultFile$'"    'labelSegment$'   'minPitch:0'   'maxPitch:0'   'meanPitch:0' 'meanPitchCurve:0'
   fileappend "'resultFile$'"  'labelSegment$' 'minPitch:2' 'maxPitch:2' 'meanPitch:2' 'stdDevPitch:2'
   ##### End writing #####
```

```
endproc


##############################################
procedure IntensityMinMaxMean timeBegin timeEnd

   select Intensity 'soundname$'
   minIntens =  Get minimum: timeBegin, timeEnd, "Parabolic"
   maxIntens = Get maximum: timeBegin, timeEnd, "Parabolic"
   meanIntens = Get mean: timeBegin, timeEnd, "energy"
   stDevIntens = Get standard deviation: timeBegin, timeEnd

   ##### write results to console (in Hz) #####
#      printline Pitch (min, max, mean(points), mean(curves)) 'labelSegment$' 'minPitch:0' 'maxPitch:0' 'meanPitch:0' 'meanPitchCurve:0'
   printline Intensity (min, max, mean, stdev) 'labelSegment$' 'minIntens:2' 'maxIntens:2' 'meanIntens:2' 'stDevIntens:2'
   ##### write results to file (in ms/Hz) #####
#      fileappend "'resultFile$'"  'labelSegment$' 'minPitch:0' 'maxPitch:0' 'meanPitch:0' 'meanPitchCurve:0'
   fileappend "'resultFile$'"  'labelSegment$' 'minIntens:2' 'maxIntens:2' 'meanIntens:2' 'stDevIntens:2'
   ##### End writing #####

endproc

##############################################
procedure Formantvalues timeBegin timeEnd

   ### calculation time point, middle of segment
   timePoint = (timeEnd + timeBegin)/2

   ### calculation formant values
   select Formant 'soundname$'
   formant1 = Get value at time... 1 'timePoint' Hertz Linear
   formant2 = Get value at time... 2 'timePoint' Hertz Linear
   formant3 = Get value at time... 3 'timePoint' Hertz Linear
   formant4 = Get value at time... 4 'timePoint' Hertz Linear

   ##### write results to console (in ms/Hz) #####
   timePoint = timePoint * 1000
   printline Formants (F1, F2, F3, F4) 'labelSegment$' 'formant1:0' 'formant2:0' 'formant3:0' 'formant4:0'
   ##### write results to file (in ms) #####
   fileappend "'resultFile$'"  'labelSegment$' 'formant1:0' 'formant2:0' 'formant3:0' 'formant4:0'
   ##### End writing #####
```

```
endproc

###############################################
procedure Gravityvalue timeBegin timeEnd

   select Sound 'soundname$'
   Extract part: 'timeBegin', 'timeEnd', "Hanning", 1, "no"
   To Spectrum... yes

   gravityValue = Get centre of gravity... 2
   select Sound 'soundname$'_part
   plus Spectrum 'soundname$'_part
   Remove


   ##### write results to console (Hz)#####
   printline Centre of gravity 'labelSegment$' 'gravityValue:0'
   ##### write results to file #####
   fileappend "'resultFile$'"  'labelSegment$' 'gravityValue:0'
   ##### End writing #####

endproc
###############################################
```

**Python script (Vos, 2019)**

```
# Python script to read in different result files and output one excel file containing all the relevant information
# Created by Maarten Vos s4385527
from __future__ import print_function
import glob
import pandas as pd
import re


def create_dataframe():
    """
    Create a dataframe in which all the data is stored
    :return: The dataframe that is created
    """
    columns = ['type', 'file_name', "duur_tot", 'tot_int', 'expression', 'nr_undefined', 'dur_mean', 'pitch_min_mean',
               'pitch_max_mean', 'pitch_mean_mean', 'pitch_std_mean', 'pitch_var_mean',
               'intensity_min_mean', 'intensity_max_mean', 'intensity_mean_mean', 'intensity_std_mean', 'f0_mean',
               'f1_mean', 'f2_mean', 'f3_mean', 'grav_center_mean']
    return pd.DataFrame(columns=columns)


def extract_from_results(file_name):
    """
    Extract the relevant information from the result file
    :param nr: The name of the current file
    :return: A list containing the extracted features from the file
    """
    # Read in the lines of the current file
    with open(file_name, "r") as f:
        lines = f.readlines()

        # Determine the type of the recording (stable, exacerbation, reference)
        type = file_name.split('_')[1][:-5]

        # Create a list containing for each element in the file a dictionary
        dict_list = [{} for _ in range(len(lines[2:]))]

        for i, line in enumerate(lines[2:]):

            # Split the line into an easy to use list
            line_list = line.split(" ")

            # Use the row dict in which all the information for the current row is stored
```

```
row_dict = dict_list[i]

# Add the type to the list
row_dict['type'] = type

# Add the name of the file from which the information is extracted
row_dict['file_name'] = line_list[0]

# Add the total duration
row_dict["duur_tot"] = int(line_list[-1])

# Add the number of intervals
row_dict['tot_int'] = int(line_list[-3])


# Extract the expression that is said and how many of the measurements are not defined
nr_und = 0
means = [0 for _ in range(15)]  # duration, pitch min max mean std var,
                               # intensity min max mean std, f0, f1, f2, f3, grav
nr_means = means.copy()
expression = ['dummy']
line_list = list(filter(None, line_list))
for d, x in enumerate(line_list[1:-4]):
    if not re.match("[\d]+", x):
        if x == "--undefined--":
            nr_und += 1
        else:
            if not expression[-1] == x:
                expression.append(x)

                # store the mean values
                ind = 0
                for b in line_list[d+1:d+22]:
                    if b != line_list[d+1] and 'tot_dur' not in line_list[d+1:d+22]:
                        if b != "--undefined--" and float(b) > 0:
                            nr_means[ind] += 1
                            means[ind] += float(b)
                        ind += 1
    elif float(x) <= 0:
        nr_und += 1

for nr, item in enumerate(['dur_mean', 'pitch_min_mean', 'pitch_max_mean',
'pitch_mean_mean',
                          'pitch_std_mean', 'pitch_var_mean', 'intensity_min_mean',
'intensity_max_mean',
                          'intensity_mean_mean', 'intensity_std_mean', 'f0_mean', 'f1_mean',
'f2_mean',
```

```
                        'f3_mean', 'grav_center_mean']):
            row_dict[item] = means[nr]/nr_means[nr]

        row_dict['expression'] = " ".join(expression[1:])
        row_dict['nr_undefined'] = nr_und

        # add the row back to the list
        dict_list[i] = row_dict
    return dict_list

    # de 66 de 171.91 184.76 178.68 5.33 de 798.78 de 54.55 70.65 67.13 6.19 de 372 1636 2791 4545 de 282
    # duration -> pitch min max mean std -> pitch variability -> intensity min max mean std -> formants -> center of gravity


def run():
    """
    Run the program to extract the relevant information from different text files and store in one excel file
    """

    # Create a new dataframe to store all information in
    data = create_dataframe()

    # Loop through all the result files in the current directory
    files_total = glob.glob("results_[a-z0-9]*.txt")
    for i, f in enumerate(files_total):

        # Extract the information from the corresponding files and store in the dataframe
        result = extract_from_results(f)
        for d in result:
            data = data.append(d, ignore_index=True)

        # Print the progress
        print("File " + str(i + 1) + " of " + str(len(files_total)) + " done")

    # Save the dataframe
    output = "extracted_info.xlsx"
    print(str(len(files_total)) + " files processed. Storing the extracted data into " + output)
    data.to_excel(output, index=False)


run()
```

## Appendix 4: Linguistic pauses

Pauses are considered to be "periods of silence in the speech of a person" (O'Connell & Kowal, 1983, p. 221) and their occurrence is determined by factors such as emphasis, breathing and syntactic complexity (Oliveira, 2002). The pause is one of the most efficient prosodic feature for signaling structure in a text (Ramanarayanan & Bresch, 2009). In order to subdivide the transcript into standard linguistic groups, the following definition was adopted: "A pause is a period of vocal inactivity of a certain duration embedded in the stream of speech, which does not include the occurrence of filled pauses." (Oliveria, 2002, p. 44). Grammatical pauses are choreographed and occur between phrases, while nongrammatical pauses occur within phrases (Reich, 1980; Thurber & Tager-Flusberg, 1993). A phrase or clause is a group of words that contains a subject and predicate. A phrase can either be independently used as a complete sentence or dependently used as a subordinate clause (Hawkins, 1971). A phrase would be perfectly read if there do not occur pauses within the phrase. The persistence of a steady breathing (or pausing) pattern becomes more difficult for speakers with a speech-related pathology (Davis, 2009; Gayraud, Lee, & Barkat-Defradas, 2011; Yunusova et al., 2016).

For this paper, the grammatical boundaries were defined after annotating the sound files to minimize the risk of bias. The boundaries were determined according to the following set of criteria (Bailly & Gouvernayre, 2012):

i. Sentence-final pauses were judged to be grammatical;
ii. Sentence-internal pauses were judged grammatical if they were the result of a major punctuation, such as a colon;
iii. Sentence-internal pauses were judged grammatical if they were located at major syntactic boundaries, such as a dependent clause starting with a verb, preposition or coordination.

Pauses which did not meet these requirements were judged ungrammatical. Syntactic pauses produced with no inhalation were not taken into account, because they are more likely to reflect difficulties with reading instead of a distortion of the breathing pattern (Yunusova et al., 2016).

Therefore, audible inhalation or exhalation during a pause was required to be taken into account. Appendix 5 gives an overview of the linguistic groups for De Koning.

**Linguistic pauses in ‘De Koning’ (Bomans, 1946)**

1. Het was najaar en de zon scheen.
2. De grote loofboom was in bloei.
3. De fonteinen sprongen fonkelend omhoog
4. en goudvissen, zo groot als dolfijnen, schoten door het glinsterende water.
5. Er speelde een orkest in het prieeltje
6. en zeven danseressen dansten om de put.
7. Overal stonden goochelaars te goochelen
8. en een buikspreker sprak buik op het pad.
9. “Gebeurde er maar wat”, zei de koning:
10. “Ik word gek van al die stilte in mijn kop.”
11. Toen riep de koning zijn minister van Plezier en Vermaak.
12. Hij zei:
13. “Bedenk eens iets nieuws, alsjeblieft.”
14. “Wat vindt u van vuurwerk?” vroeg de minister.
15. Even later schoten de raketten, pijlen en bollen omhoog,
16. en tot slot vloog het hele paleis in de lucht.
17. “Vervelend”, zei de koning toen het afgelopen was.
18. “Kun je niet iets grappigs bedenken?
19. En rap!”
20. “Een circus, koning!
21. Lijkt dat u wat?”
22. De volgende dag was er een dromedaris die muziek speelde op zijn fluit
23. en er waren zes loslopende olifanten die iets deden wat andere olifanten nooit doen.
24. Het circus was compleet met het gespikkelde beest die een brulaap aan zijn staart had hangen.
25. De koning zat op de voorste rij en vond het eentonig.
26. “Dat hebben we ook weer achter de rug”, zei hij toen het uit was.

27. “Kun je niet iets bedenken, waarvan je zegt:
28. ‘Dat is nou écht vrolijk?’”
29. De minister had nog één tip:
30. “Ik zelf!” schaterde hij terwijl hij op het riet dook.
31. Hij trok zijn jas uit, ging op zijn hoofd staan en zoemde een lied.
32. “Dat is evengoed flauw”, zei de koning toen het voorbij was.
33. “Ik wil een huis van glas, met kozijnen van diamant en een gouden sieraad.
34. Als het af is, zal ik het stukslaan met een stok.”
35. De tafels, de stoelen, de kussens, de bedden, de kolen in de kelder en zelfs het tafelkleed waren van glas.
36. En de ramen blonken als diamant.
37. En toen het af was, liep de koning erin rond en sloeg het stuk met een stok.
38. Toen alles kapot was en hij iets nieuws wilde bedenken, zei de minister:
39. “Halt. Het geld is op.”
40. “Kunnen we niet iets met het paleis?” dacht de koning.
41. Hij gaf het bevel:
42. “Verkoop het paleis.”
43. “Het paleis is in de lucht gevlogen, koning.”
44. “Dan wil ik de bromtol nog eens zien
45. want dat was het beste.”
46. Maar de minister weigerde.
47. Toen begreep de koning dat hij arm was en hij schrok.
48. Uit wanhoop verliet hij zijn eigen rijk.
49. En onze koning kwam menig man tegen.
50. De mensen haalden hun schouders op en zeiden:
51. “Wat zou hij hebben doorstaan?”
52. Ach, stik!
53. Nu gebeurde het dat hij moe en angstig in een weiland een man zag.
54. Hij vroeg:
55. “Wat ben je aan het doen?”
56. De man antwoordde met:

57. “Niets, al dagen helemaal niets.”
58. De man stond op en begon het koren te maaien.
59. De hongerige koning maaide eveneens mee met de man.
60. Toen het avond was geworden legden zij zich neer in het gras en keken samen naar de sterren.
61. Zo deden zij vele weken.
62. Toen keerde de koning terug naar zijn koninkrijk met een verjongd gemoed.

**Linguistic pauses in ‘Papa en Marloes’ (Van de Weijer & Slik, 1991)**

1. Papa en Marloes staan op het station
2. Ze wachten op de trein
3. Eerst hebben ze een kaartje gekocht,
4. maar er stond een hele lange rij dus dat duurde wel even
5. Nu wachten ze tot de trein eraan komt
6. Het is al vijf over drie dus 't duurt nog vier minuten
7. Er staan nog veel meer mensen te wachten
8. Marloes kijkt naar links
9. In de verte ziet ze de trein al aankomen

# Appendix 5: Manual forced aligner

**STEP 1: CREATE A FOLDER**:

1. Open PuTTY and login:
   Login as: **username**
   Password: **password**

   *Username: first letter of first name + family name. Lowercase letters.*
   *e.g.: Piet van Dijk → pvandijk*

2. Create a folder under: /vol/tensusers2/**username.**

   To create a folder, use command **mkdir**, e.g.
   [pvandijk@applejack:/vol/tensusers2/] $ **mkdir** pvandijk]
   - pvandijk@applejack means you are logged in on Ponyland on applejack [computer cluster name] with account pvandijk.
   - /vol/tensusers2/ means you are under this path.
   - **mkdir** is the command to make a directory. To use this, follow it by a space and then the name you want to give to the folder [in this example *pvandijk*].

3. Go to the new folder and copy the data and the forced aligner to the new folder.
   To go to the folder (this will change the path): use the command **cd**.
   Example: [**username**@applejack:/vol/tensusers2/] $ **cd username**
   [**username**@applejack:/vol/tensusers2/username/]
   Current location: /vol/tensusers2/**username**

4. To copy the data: use command **cp**.
   Example (copying forced aligner): [username@applejack:/vol/tensusers2/**username**/] $ **cp** –r /vol/tensusers/mganzeboom/clst-asr_forced-aligner .
   - **cp** = the command to copy things
   - **-r** = the option for **cp** command. Here it means that you are copying a folder.
   - /vol/tensusers/mganzeboom/clst-asr_forced-aligner = the thing you want to copy.
   - . (little dot) = you want to copy something to the current position/path which in this example is under /vol/tensusers2/**username**/.

   Current situation: the data and the forced aligner are in your folder. You have two folders under your own folder username: clst-asr_forced-aligner and data

**STEP 2: RUN THE FORCED ALIGNER**

General instructions, without errors: https://www.cls.ru.nl/clst-asr/doku.php?id=forced-aligner
Be sure: wave files + textgrid files have the same name without extensions.
*e.g. pp1_reading.wav and pp1_reading.Textgrid are compatible, but the combination pp1_reading.wav and pp1_readingstory.Textgrid is not.*

Steps to run the forced aligner (copied from website cls.ru.nl):

1. Login to one of the ponies (do not use (the old) applejack because of an older CUDA version).

2. Run from your home directory:
   /vol/tensusers/mganzeboom/clst-asr_forced-aligner/run.sh
   <absolute-path-to-directory-with-recordings>.
   The script will create a directory in your home directory and copy the default config and lexicon files.

3. Open ~/clst-asr-fa/align_config.rc with your favourite editor and change the configuration settings to your liking (the defaults are fine on average).

4. Run step 2 once again and a job will be added to the Slurm queue manager starting the forced alignment of the provided directory. The logs of this job can be found in ~/clst-asr-fa/slurm-logs/slurm-<job-id>.out. Provide multiple input directories at the command line to queue multiple jobs at once.

5. The force alignment logs can be found at <absolute-path-to-directory-with-recordings>/logs when all Slurm jobs have completed.
   -aligner.tgscriptions are not in the lexicon provided with the acoustic models. Run the script /vol/tensusers/mganzeboom/clst-asr_forced-aligner/list-missing-words.sh <absolute-path-to-directory-with-recordings> to print a list of these words and their corresponding transcript files to the command line. You could then add a phonemic transcription of these words to your custom lexicon file in ~/clst-asr-fa/lexicon.txt. It is recommended to base these new transcripts on parts of already existing ones. Afterwards, rerun the script from step 4.

Current situation: stage 3 (of 7) is finished. Under audio folder you see a subfolder 'log'. In this 'log' folder, another subfolder is created called 'splits'.

To check in which stage you are: open the script called 'run_forced_alignment.sh'. Use command **cat** (no editing in the script) or **nano** (editing and save, be careful using this) which is under path e.g. /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/
E.g. cat: [**username**@applejack:/vol/tensusers2/**username**/] $ **cat** clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/run_forced_alignment.sh
E.g. **nano**:

- You want to open the script called "run_forced_alignment.sh" as shown in the command the light blue part run_forced_alignment.sh.
- But to open this, since it's not in your current path, you need to specify the path to that script which is the purple part in the command clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/. Here you use the relative path since it's in a subfolder in where you currently is. You could also use absolute path, it always works. In this example, the absolute path: /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/
- You can see, the absolute path is combined with your current path (dark blue part) and the subfolder path (purple part).

**STEP 2: RUN FORCED ALIGNER MANUALLY**

7. When there is no final ali in the folder (check this with notepad), manually run the forced aligner:

/vol/tensusers2/**username/**clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/run_forced_alignment.sh --config ~/clst-asr-fa/align_config.rc <absolute-path-to-directory-with-recordings>
/vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/run_forced_alignment.sh --config /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/align_config.rc <absolute-path-to-directory-with-recordings>
Current situation: stage 4 is finished. You can see the output in the window.

8. Manually run the praat script: praat --run: /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/align2praat/createtextgrid.praat <absolute-path-to-splits-directory-under-recordings-folder> < absolute-path-to-directory-with-recordings> <phone_tier_name>
   Current situation: stage 5 is finished.

9. Manually run another praat script: praat --run: /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/align2praat/createWordTextGrids.praat <absolute-path-to-log-directory-with-recordings> <absolute-path-to-directory-with-recordings> <absolute-path-to-splits-directory-under-recordings-folder> <word_tier_name>
   Current situation: stage 6 is finished.

10. Change stage to 7 (stage=7) in your config file in your home directory. Re-run: /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/run_forced_alignment.sh –config ~/clst-asr-fa/align_config.rc <absolute-path-to-directory-with-recordings>

11. Finally, manually run another praat script: praat --run /vol/tensusers2/**username**/clst-asr_forced-aligner/kaldi/egs/clst-asr_forced-aligner/s5/align2praat/stackTextGrids.praat <absolute-path-to-splits-directory-under-recordings-folder> <absolute-path-to-directory-with-recordings>

Output: text grid file named: -aligner.tg with word level & phoneme level alignments. You did it! Congratulations!

**Tips:**

**Map Ponyland on your own laptop**

If you have problems with running praat scripts in the command line, here is another option. First, map the Ponyland to your own laptop. This is easier than filezilla and looks like a windows thing.

1. Open your file explorer.
2. Right click “This PC”, then click “Map network drive…”
3. In the new window, give the drive name, default is X:. Give Folder name, here is e.g. \\derpy.science.ru.nl\tensusers2**username**
   You probably need to login with your username and password. This could also work outside the campus once you connect the VPN.

Result: you see everything in that folder and editing things as you normally do in Windows