CHAPTER 9
GENERAL DISCUSSION

Parkinson’s disease (PD) rates are rapidly increasing in almost every region of the world (Dorsey et al., 2020). In the Netherlands alone, there are around 57,000 people living with PD (ParkinsonNet, https://www.parkinsonnet.nl/parkinson/), which means that three out of 1,000 people have been diagnosed with the disease. Hypokinetic dysarthria is among the early manifestations of PD (Brabenec et al., 2017), and is the central theme of this dissertation.

In line with the majority of research on this topic, there is more than one way in which HD affects speech of people diagnosed with PD. Overall, all speech production difficulties that people with PD face lead to various troubles with communication. As such, conversational difficulties may have serious repercussions for many areas of life of PwPD, eventually leading to communication anxiety and the feeling of societal isolation (Miller et al., 2006; Miller, 2017). The main aim of this dissertation is to gain insights into listeners’ impressions of dysarthric speech and to uncover acoustic correlates of such impressions.

This final chapter discusses the findings and implications of the different studies comprising this dissertation from a higher vantage point. The studies in the dissertation provide a longitudinal and cross-sectional perspectives on changes in speech of PwPD and offer a reflection on how speakers with PD sound to other people. The presented findings demonstrate that listeners’ ability to recognize PD speech as unhealthy is rooted in the acoustic cues of such speech, not in its content. Moreover, such impressions depend on listeners’ training and familiarity with the speakers’ language. This discussion argues for a more rigorous approach in studies investigating subjective recognition of speech affected by HD.

This chapter unpacks the main research question of the dissertation which inquires into listeners’ recognition of speech of PwPD and the factors that may influence it. This dissertation approaches this question from three perspectives: comparing HD to other types of dysarthria, exploring effects of listeners’ training and language background, and monitoring speech changes longitudinally. The summary below (section 9.1) briefly describes the main results of the experimental studies comprising each of the three parts of the dissertation. Following the summary, sections 9.2, 9.3, and 9.4 discuss common topics in relation to the major findings of the dissertation, with section 9.5 addressing the implications for PwPD in a multilingual world. The chapter ends with a discussion of the limitations of the current dissertation together with an outline for future research (section 9.6) and concluding remarks (section 9.7).
9.1. Summary of Studies and Findings

The first part of this dissertation contributes to the existing evidence on acoustic characteristics of HD on segmental and prosodic levels. Chapter 2 describes the two pilot studies exploring acoustic correlates of vowel articulation in German PwPD with and without mild cognitive impairment (MCI), as well as acoustic correlates of vowel articulation and prosody in speech of Dutch PwPD. The results of both studies confirmed the sensitivity of the VAI measurement for detecting the vowel centralization in spontaneous speech of PwPD in Dutch and German. The first study also demonstrated the detrimental effect of MCI on vowel articulation leading to more pronounced vowel centralization effects in the male group of PwPD with MCI. The second study showed a monopitch trend and reduced speech rate in recordings of PwPD.

Moreover, the first part of the dissertation explores listeners’ recognition of linguistic prosody produced by PwPD. Building upon the classical description of deviant speech dimensions in HD by Darley et al. (1969b), chapter 3 systematically reviews the literature exploring how listeners with different expertise (trained and untrained in the field of speech and language therapy) assess aspects of dysarthric speech. The results suggested that similar acoustic cues are prominent for both trained and untrained listeners. Chapter 4 reports on the results of the experiment on recognition of linguistic prosody in dysarthric speech, providing insights into how well listeners recognize the intended linguistic prosody in speakers with HD compared to speakers with other types of dysarthria (ataxic and mixed). Among four different intonation recognition tasks, the sentence type recognition task proved to be the most sensitive for differentiating among dysarthria types. Both the recognition accuracy results and acoustic measurements of $f_0$ change within a phrase demonstrated that speakers with HD had the most prominent prosodic deviance among all three groups with dysarthria.

The second part of the dissertation examines how accurately listeners can recognize speech of PwPD. It uncovers the acoustic cues that allow listeners to recognize speech of PwPD differently compared to the speech of control speakers. Chapter 5 deals with the questions of listeners’ expertise and training as well as familiarity with speakers’ language, while chapter 6 investigates the correlation between acoustic cues in dysarthric speech and listeners’ responses in a speech recognition task. By means of a series of cross-linguistic experiments with different listener groups, it appeared that training in either speech therapy or in phonetics together with familiarity with speakers’ language are significant factors and may have different effects on dysarthric speech recognition. While trained listeners outperformed untrained Dutch and non-Dutch listeners in recognizing the sentence type intonation, recognition of speech healthiness yielded contradictory results for the trained group. Independent of their training – either in speech and language therapy or in phonetics – trained listeners had lower accuracy for speech healthiness recognition compared to the untrained groups. At the same time, native language also played an important role in the recognition of speech of PwPD. Surprisingly, in the healthiness recognition task, it was the untrained non-Dutch listeners with native Germanic languages who performed the most accurate even amongst Dutch trained and untrained groups. The predictive modelling of listeners’ responses in the task of recognizing speech as healthy or unhealthy highlighted aspects of phonation and prosody that serve as promi-
nent markers of speech healthiness for listeners independent of what their first language or expertise are. Therefore, the second part of the dissertation reveals a close link between objectively quantified speech changes in HD and subjective listeners’ impressions of the speech affected by HD. Such subjective impressions may often bear a pejorative undertone, which has real-life implications for PwPD. Thus, knowing particular acoustic cues that trigger listeners’ impressions of “unhealthy” can provide specific therapeutic targets in addition to the existing dysarthria treatment in PwPD.

The third part of this dissertation follows one bilingual speaker with PD. The two studies included in part three report on the longitudinal changes as captured by both acoustic measurements and by experimental results reflecting listeners’ speech recognition. The speech healthiness assessment experiment in chapter 7 was conducted on the first year of the individual’s recordings, and the results demonstrated that trained and untrained listeners rated the later recordings as less healthy relative to the earlier ones. Moreover, listeners’ experience with speech disorders influenced the trend, which was more pronounced for the experienced listeners compared to the listeners with no prior experience with speech disorders. Once again, the findings demonstrated that in both cross-sectional and longitudinal case studies, the factor of listeners’ training influences the recognition of speech of PwPD. By tracking phonation and prosody related features over 3.6 years, including the pre-, during, and post-speech therapy periods, chapter 8 reveals that all measurements have similar trends in general, irrespective of the spoken language, while some of them demonstrate more pronounced and significant trends in one language and task type only (e.g., speech rate in second language reading). This exploratory analysis of acoustic change provided evidence of the presence of speech therapy effects in both languages, even though the therapy was administered only in the native language of the speaker - Dutch.

9.2. Does HD Affect Linguistic Prosody?

Prosody is an important part of human communication as it is integral to the expression of a variety of meanings from diverse linguistic functions such as intonation, stress and rhythm to expressing our emotions and state. The first part of this thesis brings to light the intersection of linguistic prosody and dysarthria.

Despite the varying designs of the studies concerned with perception and recognition of linguistic prosody produced by PwPD, throughout the literature the monopitch and monoloudness dimensions were reported to be important factors in the successful identification and discrimination of linguistic prosody by listeners (chapter 3). This highlights the fact that many studies, alongside the therapeutic management of dysarthria in PD, are mostly focused on pitch level and loudness (Pinto et al., 2017). However, HD-related changes in prosody, or in speech for that matter, are diverse and there is little evidence on how speech characteristics other than monopitch and monoloudness may contribute to listeners’ impressions of speech affected with HD. Moreover, research has demonstrated that there is still a gap to bridge between PwPD in need of speech therapy services and the availability and type of such services (Schalling et al., 2017).

HD is known to affect both linguistic and emotional prosody (chapter 3), which are crucial aspects of communication. The presence of prosodic deviances holds true for
other types of dysarthria as well, and many studies have indicated that different dysarthria types have different prosodic deficit profiles (Darley et al., 1969b; Liss et al., 2009, 2010). The comparative study described in chapter 4 explores this topic by focusing on the ability of speakers with different types of dysarthria to convey different linguistic functions of Dutch prosody (Rietveld and Van Heuven, 2009; Martens et al., 2011). From the accuracy results of prosody recognition, it appears that HD stands out from other dysarthria types: intonation of the speakers with HD was the least recognizable. It also becomes clear that among the four prosody recognition tasks (lexical stress, boundary marking, focus, sentence type), only the sentence type recognition task was sensitive enough for differentiating dysarthric and control speech, as it showed the lowest percentage of unspecified (i.e., "I don't know") answers. Also, it was the only prosody recognition task that demonstrated correlation with acoustic measurements of $f_0$ trajectory estimation. This finding that sentence type recognition appears to be the least confusing prosody recognition task for listeners is not surprising. It is likely that in comparison to the other recognition tasks (see chapter 3 for details), deciding whether they heard question or statement was the least ambiguous task for the listeners. These findings fueled the studies described in chapters 5 and 6. The sentence type intonation recognition task was included in the experimental design to test the effect of language background and expertise of the listeners, as this task also tests the intelligibility of sentence type intonations produced by PwPD. The reported results of the studies in chapters 5 and 6 indicate that irrespective of their language background or expertise, listeners were able to recognize question and statement intonations of PwPD less accurately than the same intonations produced by the control speakers. This evidence not only adds to the existing knowledge on the intelligibility of speech of PwPD, but also indirectly demonstrates that the presence of HD would hinder any communication, mono- or multilingual. In the light of “misperception that PwPD are depressed, disinterested, or tired, when they are not, and the negative evaluation of PwPD by listeners” (Miller, 2017, pp. 267-268), our cross-linguistic findings of lower intelligibility of sentence type intonation support a number of studies indicating the importance and benefits of speech therapy to target prosodic deviancy in HD (Peppé, 2009; Anand and Stepp, 2015; Miller, 2017).

**9.3. WHAT INFLUENCES SPEECH RECOGNITION?**

Research has demonstrated that many factors related to both, speech itself and to listener’s experiences can affect how people assess and recognize dysarthric speech. For example, in the study by Carvalho et al. (2020), the authors summarized listener-related factors that affect the intelligibility results for PD speech: listener’s experience and training, listener’s familiarity with dysarthric speech, the type of stimuli (such as differences in content or in dysarthria severity), and listener’s biases (such as fatigue or attention lapses). Chapters 5 and 7 of this dissertation contribute to the understanding of how the different task types (recognition of speech healthiness and of sentence type) as well as listener-related factors (training and language background) affect the recognition of speech of PwPD. While chapter 5 reports on two studies with cross-sectional designs, chapter 7 provides a longitudinal perspective on the recognition of speech of PwPD by different groups of listeners.
9.3.1. Expertise and experience

As discussed in chapter 5, research provides us with somewhat conflicting evidence regarding the role of listeners’ training and expertise in the assessment of dysarthric speech. Some authors (Walshe et al., 2008; Smith et al., 2019) have demonstrated that trained and untrained listeners judge intelligibility of dysarthric speech with no significant differences, while others reported more pronounced trends (Verkhodanova et al. (2019), see chapter 7) or higher intelligibility values (Carvalho et al., 2020) for the trained listener groups. Moreover, Carvalho et al. (2020) demonstrated that healthcare professionals working specifically with PD understand PwPD better than people from other listener groups, including speech language therapists and neurologically healthy listeners familiar with PD. Such results are supported by findings in this dissertation.

In chapter 5, both listeners trained in the field of speech and language therapy and listeners with expertise in phonetics and speech sciences demonstrated better accuracy results than untrained listener groups when recognizing sentence type intonation produced by PwPD. However, the results were strikingly different in the task of recognizing speech as healthy or unhealthy. This surprising discrepancy in two listening tasks aligns in part with the suggestion put forward by Kreiman et al. (1990), that clinical training and experience affect speech perception causing listeners to differ more in their assessments. It appears to hold true for training in phonetics and speech sciences, as discussed in chapter 5. In their study, Kreiman et al. (1990) explained the differences in patterns of how trained and untrained listeners judged speech dimensions, as well as the diverging strategies in their trained group, by the larger range of information about speech and voice that was available to the trained listener group due to their training.

Therefore, while in a sentence type recognition task the idea of understanding the intended intonation is very similar to the common intelligibility tasks, the healthiness recognition task is different and requires listeners to conceptualize what constitutes "healthiness" of speech for them. Similar to the task-related discrepancy of the comprehension and intelligibility experiments described in Hustad (2008), it is likely that in the speech healthiness recognition task, listeners were focused on constructing a coherent picture of what constitutes a healthy voice and speech, actively drawing upon their world knowledge and experience. While in the sentence type recognition task, it is likely that listeners were focused solely on the intonation contrasts.

As the concept of healthiness was not defined in the healthiness recognition task, this task was meant to be an approximation of those recognition-based decisions that people make every day. Listeners were free to employ whatever concept of healthiness they had constructed and refer to whatever speech and voice characteristics seemed relevant for each given stimulus. Therefore, in the absence of well-structured rating scales and benchmarks that SLTs are used to in their clinical work, and the lack of phonetically explicated phenomena corresponding to "healthiness" which phoneticians could possibly use as reference (chapter 5), the groups of trained listeners performed less accurately in this task compared to the groups of untrained listeners. Their attentive efforts of using specific criteria imposed by their training may have led to the lower accuracy in their recognition of dysarthric speech as unhealthy. Therefore, the proposed idea of distraction in (chapter 5) and (chapter 6) reflects the interpretations suggested in earlier studies (Kreiman et al., 1990; Walshe et al., 2008). That is, the influence of the trained
9.3.2. LANGUAGE BACKGROUND

A similar distraction paradigm explains the differences seen in the responses to recognition tasks given by listeners with diverse language backgrounds. Objectively, it is impossible to investigate one isolated factor that contributes to the certain interpretation of the speech signal. When it comes to phonetics, the recognition process “requires that listeners take into account not only the acoustic cues present in the speech signal, but also their own knowledge of the patterns of their language, in order to interpret what they hear” (Crystal, 2011, p. 356). However, the studies described in part two of this dissertation attempted to focus on the acoustics of the dysarthric speech and to deduce the effect of content influence by including non-Dutch listeners with different native languages.

Alongside expertise, chapters 5 and 6 discuss recognition of untrained non-Dutch listeners including larger groups of speakers with Slavic and Germanic native languages. This cross-linguistic comparison yielded surprising results, as the Germanic group performed best among other groups in the healthiness recognition task, lending support to the idea of distraction on a different level, namely, the influences of the native language systems. Where the Dutch group could have been distracted by the content of the message, the non-Dutch group, when listening to the unfamiliar language intuitively filtered the unfamiliar sounds through, in the words of Nikolai Trubetskoy, “the familiar ‘phonological sieve’ of their mother tongue to analyze what has been said” (1939, p. 51). That is, they were distracted by mismatches with the phonological system of their native language. This explanation could partly contribute to the distributions of the scores in the healthiness recognition task. While we witnessed lower scores for trained and untrained Dutch listeners, Germanic non-Dutch listeners with no working knowledge of Dutch but with relatively close phonetic and prosodic systems in their native languages, outperformed every other group. The lowest scores of listeners with Slavic language background may also be explained by the presence of distraction, namely the Slavic listeners’ unfamiliarity with the phonetic and prosodic inventories of Dutch.

The effect was different for the sentence type recognition task, where Dutch listeners outperformed the non-Dutch ones, and the Germanic group outperformed the Slavic group. Here, the clear concept behind the task – intelligibility of dysarthric prosody – together with controlled stimuli may have provided both fewer distractions and alternatives for interpretation. The results from the sentence type recognition task (see chapter 5) suggest that the language-specific nature of PwPD’s compensatory strategies for prosody, contrary to expectations, do not seem to translate as universal to listeners of other languages.

9.3.3. PREDICTING RECOGNITION OF SPEECH HEALTHINESS

The idea of employing acoustic measurements as an additional help or even as a substitute to the clinical diagnosis of dysarthria in PD inspires many researchers (Brabenec et al., 2017). However, while there are many studies that focus on acoustic correlates of
intelligibility of speech produced by PwPD (Anand and Stepp, 2015; Weismer et al., 2001; Feenaughty et al., 2014), less is known about which acoustic cues in dysarthric speech trigger complex and more "global" impressions in listeners. While previously discussed results of the sentence type recognition task demonstrate a clear pattern of question and statement intonation intelligibility depending on listeners’ language background and expertise, results of a healthiness recognition task paint a more complex picture (see chapter 5). Therefore, the question of mapping between acoustic features and listeners’ responses in a speech healthiness recognition task was of particular interest (see chapter 6).

Conventional acoustic features used in the experiments from chapter 6 appeared to be sufficiently predictive of listeners’ responses in the recognition task: the classification accuracy of the model trained on acoustic features was 84% at the lowest. Even though the decision was not to perform any model optimizations to avoid researcher bias, it is likely that the reported results of the average predicting accuracy of 85-86% is already (close to) the optimal value with a given set of predictors. Moreover, the results of variable importance ranking revealed the apparent universality of certain acoustic features contributing to the prediction of speech healthiness judgements of listeners with different language backgrounds or expertise. Interestingly, independent of the listener group, the highest ranked predictors were the same and corresponded to the domains of prosody, phonation and articulation, despite the differences in the phonetic systems of the native languages of the Dutch and non-Dutch listeners. Everyone appears to be sensitive to acoustically measurable vowel centralization, changes in breathing, voice quality and prosody.

Chapter 6 provides insights into the predictive power of conventional acoustic features and explores whether demographic features can shed some light on recognition strategies of different listener groups. Interestingly, in contrast to expectations, the presence or absence of a speaker’s self-reported dialect was always ranked low on the list of variables contributing to prediction accuracy. On the other hand, speakers’ age and gender appeared to be more important predictors for the responses of the Slavic listeners than of Dutch or Germanic listeners. Such evidence suggests that the impressions of healthy and unhealthy speech, while predictable from acoustic measurements, measurements, depend to a great degree on the listener’s world view, knowledge and experience. It is possible to speculate that the higher rank of speaker’s age and gender predictors for the Slavic listener group is most likely rooted in the demographic situation in the former Soviet countries. For example, consider how in Russia, women comprise 62% of the total population of 65-69 years and 70% in the age range of 70+ (Russian Federal Statistics Service, https://rosstat.gov.ru/). Therefore, with the higher mortality rate of ageing men in Russia, Russian listeners may be biased when assessing elderly male voices, since being an elderly man may already be suggestive of the presence of some sort of ailment.
9.4. Does time matter?

While many have highlighted that HD manifestations may often have an individual pattern of expression (Skodda et al., 2012, 2013; Arias-Vergara et al., 2018), surprisingly little is known about dysarthric developments in different speech domains in the course of the disease. Part three of this dissertation addresses the gap present in the literature of longitudinal studies with dense observations over time by following one bilingual individual with PD over 3.6 years. The evidence from the first year of the Dutch recordings is discussed in chapter 7, this included both recognition experiments targeting healthiness assessment as well as acoustic measurements related to voice quality and prosody. The decline in some of the acoustic parameters throughout the first year also appeared to be prominent for listeners as they rated the later recordings as less healthy.

Interestingly, the observed trends of the healthiness assessment for just one speaker demonstrated benefits of the experience with speech and language disorders which is in contrast to the group studies discussed in chapters 5 and 6. Such a discrepancy may be explained by a number of differences in study designs. First, the scaled assessment of healthiness (7-point Likert scale vs dichotomous question with extra "I don’t know" option) is closer to the structured and scaled assessments used in SLT work. Second, the "trained" groups were comprised of different populations: while all trained listeners in the group studies had completed SLT training and most of them had experience working as SLTs, the trained group included in the single PD case study was made up from both trained SLTs and master students of a neurolinguistics programme that provides students with knowledge about many speech and language disorders including HD. Third, and most importantly, the speech material analysed in the group studies included many speakers, while in the single case study there was only one voice. The design of the experiment in the single PD case study may have provided a lack of distraction for the trained group compared to the group studies. It had a more familiar structured rating scale together with just one voice to assess. The results demonstrated benefit of training that allowed listeners to catch the specific changes in speech and voice without drawing upon the wider range of their knowledge to analyse a more global picture of speech healthiness for every new voice. However, confirming or rejecting such hypothesis calls for additional research into listeners’ impressions and assessment of speech changes, specifically longitudinal speech changes, in PwPD. Moreover, as discussed in the introduction to this dissertation, aligning the listeners’ impressions of longitudinal speech changes in PwPD to the measurable acoustic correlates of such changes would enable researchers to get a detailed perspective of HD development rooted in the context of communication. This perspective is currently lacking in the literature. Chapter 8 focuses on acoustic measurements completed over the whole 3.6 year period of recordings in both languages of the speaker - Dutch and English. The evidence discussed in that chapter demonstrates similar trends of acoustic change as well as the presence of a speech therapy effect for both languages of the speaker. Measurements of jitter and \( f_0 \) variability in Dutch monologues showed significant trends in the first year but did not retain trend significance over the period of 3.6 years. This may be an effect of speech therapy that the participant started right after the first year of recordings and/or of potential non-linearity of such changes that may reflect many unknown factors including the amount of conversations the participant had on the day of
recording, physical fatigue and many others that were not controlled for in these studies. However, even though both studies in chapter 7 and chapter 8 are of an exploratory nature, the evidence indicates a clear benefit of tracking different speech parameters in PwPD. Such tracking may help with monitoring the progression of speech changes and objectively evaluating long-term effects of speech therapy (see chapter 8), as well as helping in the early detection of dysarthria in PD (see chapter 7). Following the arguments put forward in chapter 8, a more detailed picture of the speech disorder in progression would equip speech therapists with information about the ‘process’ rather than the ‘product’ of speech changes or therapy effects.

9.5. IMPLICATIONS FOR PARKINSON’S DISEASE IN A MULTILINGUAL WORLD

As mentioned in the introduction to this dissertation, an increasingly multilingual society and migration pose additional challenges for PwPD for whom their main language differs from the dominant language of their environment.

At mid-2020, older migrants comprised an estimated 34.3 million people or 12.2% of the international migrant stock according to the 2020 report of United Nations Department of Economics and Social Affairs (Migration Data Portal, based on UN DESA: https://migrationdataportal.org/themes/older-persons-and-migration). The same data demonstrates that in 2020 in Europe, the percentage of 65+ year old migrants was 14.4% and is projected to grow. Therefore, the growing numbers of older migrants and of people developing PD highlight the complicated issues of the speech and language related challenges that a migrant person with PD may face when relocating to a different country.

Many studies on ageing and migration have observed a "language barrier" for older migrants when they have to access healthcare and communicate in a clinical setting (Lai and Chau, 2007; Dias et al., 2008; Kaitelidou et al., 2020; Pot et al., 2018a). Such language problems may affect the life of older migrants on many levels. Research has demonstrated that a limited proficiency in migrants’ second language may impact their well-being, fuel feelings of second language insecurity, and eventually have serious repercussions for their life quality (Pot et al., 2018a,b). Yet, it becomes even more complicated with the well-known life quality related consequences of HD for migrant PwPD, and in the light of the difficulties that language barriers pose for older migrant people, older migrant PwPD are even more vulnerable.

Alongside the difficulties any older migrant person may experience with regards to healthcare related communication, migrant PwPD face additional troubles. Migrant PwPD with dysarthria are often in need of speech therapy services. The difficulties of the assessment of people with dysarthria whose native language is different from the native language of an SLT have been addressed by a Swedish team (Näsström and Schalling, 2020). The authors highlighted the increasing demand on SLTs to treat speakers with dysarthria with different native languages. Näsström and Schalling (2020) describe a novel protocol for dysarthria assessment with the help of an interpreter that appears to be solid and a reliable proof-of-concept. However, it has yet to be tested on a larger scale.

In general, there are very few cross-linguistic studies focused on speech of PwPD (see
Pinto et al. (2017) and Näsström and Schalling (2020) for an overview), and apart from the studies reported in this dissertation no one has explored cross-linguistic recognition of speech of PwPD by untrained listeners. Thus, chapters 5, 6 and partly 7 provide the first insights into multilingual aspect of speech changes in PwPD and precarious efficiency of multi- and cross-lingual communication of PwPD. Therefore, to fully understand the multifold effects of HD on communication in a changing multilingual world, it is necessary to enhance our knowledge of dysarthria in multilingual contexts. There is a need for additional research into speech changes in different languages of bilingual and multilingual speakers with HD, universal and language-specific compensatory strategies of PwPD, cross-linguistic recognition of HD speech and listeners’ attitudes regarding nonnative or migrant speakers with HD. However, taking into account the language insecurities of older migrants (Pot et al., 2018b) and the existing communication problems in PwPD that often lead to depression and social isolation (Miller et al., 2006), it becomes clear that elderly migrant people with PD are severely at risk of poor well-being.

Therefore, the insufficient availability of the SLT services for PwPD (Schalling et al., 2017), the increasing and unmet need for multilingual assessment of speakers with dysarthria (Näsström and Schalling, 2020), and the lack of (longitudinal) studies on bilingual and multilingual PwPD demonstrates that research and therapy are running behind in a rapidly changing world.

9.6. STUDY LIMITATIONS AND CONSIDERATIONS FOR FUTURE RESEARCH

The ideas and evidence presented in this dissertation should prompt further research focusing on unraveling factors contributing to the communication difficulties of PwPD and on exploring the cross-linguistic bi- and multi-lingual sides of such difficulties. Following the evidence and arguments in this dissertation regarding the importance of listener-related factors for speech recognition and of dense longitudinal speech measurements of the bilingual speaker, the current section offers suggestions and pointers to the ways research in this dissertation can be improved and/or extended and continued.

The understanding of what constitutes relevant training and expertise was rather broad. Chapter 5 describes studies with trained groups constituted from speech language therapists and phoneticians, while chapter 7 reports on the students of a neurolinguistics master programme. However, comparing an SLT group which specialized in HD could have provided insights into the importance of specialized knowledge for dysarthric speech recognition. The participant recruitment for the studies in this dissertation did not yield a sufficient number of SLTs with working PD experience to allow for group comparisons. Based on results on speech intelligibility ratings among different listener groups (Carvalho et al., 2020), one may speculate that in the sentence type recognition task, SLTs with expertise in HD would be more successful than their colleagues who lack such specialized expertise. Nonetheless, it is difficult to hypothesize whether experience with PD would be beneficial or distracting in the speech healthiness recognition task. This topic warrants additional research.

Furthermore, expanding the list of listener-related factors that could influence the
dysarthric speech recognition is a worthwhile direction for future investigation. Little is known about dysarthric speech recognition by different groups of listeners. Exploring listeners' age, diverse linguistic backgrounds, and specifics of occupation or familiarity with speech disorders could help distill speech changes that are universally prominent for the different groups of listeners. Another question that could be answered by exploring further the listener-related factors and correlating listener responses with acoustic measurements is the ways in which different acoustic cues influence the impressions of different listener groups. Such research would be beneficial with regards to a number of issues; unraveling the mechanisms of successful communication of PwPD, exploring the formation of negative attitudes towards dysarthric speakers, and developing a feature set for building a speech recognition system based on the subjective knowledge of listeners.

Another question that has not been addressed by the studies in this dissertation is the presence of adjustment or learning effect throughout the speech recognition task. Previous research has suggested that while assessing dysarthric speech recordings the trained group demonstrated a shift in their responses throughout the task (Walshe et al., 2008). Therefore, exploring the possible adjustments and changes in the response patterns within the task might provide insights into recognition strategies of both trained and untrained listeners. Similar issues should be raised regarding cross-linguistic recognition: is there a certain period of adjustment that could alleviate the effect of surprising sounds of another language for listeners with a language background different from that of the speakers?

The topic of language-specific and language-universal factors of speech of PwPD remains under-researched. While this dissertation provides some insights into cross-linguistic recognition of speech of PwPD (see chapters 5 and 6), the evidence presented is limited to Dutch speakers with PD and few listener groups. However, the major argument about listeners being distracted by an unfamiliar sound system suggests that further research into cross-linguistic experiments with dysarthric speech recognition would provide valuable data to gain insight into how and to what extent listeners with different native phonetic systems recognize speech (and specifically prosody) of PwPD speaking in a language unfamiliar to the listeners.

Further investigation of the relationship between objective acoustic cues and subjective listeners' impressions would be insightful for creating a feature set reflecting communication troubles in PwPD. One possible direction for future research is to study intensity influences on listeners’ responses in different recognition tasks. The technical setting of the recording procedure for all studies included in this dissertation did not allow for the reliable measurements of speech intensity. Meanwhile, the monoloudness dimension is cited to be an important speech characteristic for listeners’ perception and recognition (chapter 3) and, at the same time, is one of the main targets in speech therapy of HD (Ramig et al., 2004). Therefore, future studies focused on listeners’ impressions of dysarthric speech or on monitoring speech of PwPD longitudinally should consider exploiting recording settings that would allow for reliable analysis of the speech intensity (e.g., headsets used by Jacobi et al. (2019)). Moreover, including a broader set of conventional acoustic features would further enhance the existing picture of a correlation between acoustic features and listeners’ responses in recognition tasks. While studies in this dissertation did not take features related to consonant articulation into account, it
is a highly ranked deviant dimension in the description of HD according to Darley et al. (1969b). Future studies should include acoustic features reflecting consonants articulation to gain insights into how this characteristic so highly affected by HD influences cross-linguistic recognition of dysarthric speech. Taking the research into acoustic features of speech of PwPD even further, a more detailed analysis of listeners’ recognition failures and successes with subsequent analysis of stimuli would provide insights into the nature of prosody-related compensatory strategies employed by Dutch PwPD.

The scarcity of comprehensive longitudinal studies with dense observations over time and/or targeting bi- and multilingual speakers with PD calls for additional research. Even though part three of this dissertation presents studies with dense observations over time, the resolution of such measurements is insufficient to account for the noise in the data. Therefore, it is worthwhile to explore the possible effects of as many external variables as possible in this and similar longitudinal studies with dense measurements, because it would paint a picture with enhanced resolution for HD-related changes. Future studies should also investigate the ways for non-linear analysis of speech changes in the longitudinal studies with time-series designs.

Yet another possible direction for future research is for the ideas and arguments put forward in this dissertation to be re-interpreted within a different theoretical paradigm. For example, one may further explore the underlying reasons for the trained listeners’ differing intra-rater reliability or less accurate recognition of healthiness within the scope of the prototype theory of categorical perception. A study by Latinus and Belin (2011) provided evidence for the similarity of coding stimulus identity between visual and audio modalities with contrastive mechanisms relative to the prototype concept, while a study by Bruckert et al. (2010) showed that perceived attractiveness of brief voice samples is a function of their acoustical distance to the average voice. Therefore it is possible to hypothesize that any listener might have certain cognitive prototypes for the “healthy” and “unhealthy” voices which shape recognition. Since prototypes originate from perceptual experiences, listeners with significantly differing experiences would presumably differ in their recognition strategies (Kreiman et al., 1990).

9.7. Conclusions

This dissertation adopted a multifaceted approach towards speech of PwPD which provided us with insights about dysarthric speech that are grounded in communication issues experienced by PwPD. The evidence reported in the studies of this dissertation demonstrates the presence of intricate relations between various factors influencing recognition of speech of PwPD. The insights gained from the studies on acoustic speech production by PwPD and the recognition of such speech by different listener groups highlights the variability of such factors from cross-sectional and longitudinal perspectives.

Findings in this dissertation are also insightful with regards to a broader picture of dysarthric speech recognition, demonstrating that whether or not dysarthric speech is recognized as “healthy” is highly dependent on listener’s expertise and language background. The majority of the studies that explore the effect of listeners’ experiences in relation to dysarthric speech recognition, focus on speech intelligibility issues, while we hardly know anything about more global and more abstract impressions of dysarthric
speech (Kreiman et al., 1993; Weismer et al., 2001; Sussman and Tjaden, 2012). Such global and abstract assessments, while being closer to the speech recognition decisions that listeners might make every day, border with attitudes and opinions, as it was pointed out by Maryn and Debo (2014). Building upon the evidence about healthiness recognition presented in this dissertation, additional research may provide us with new perspectives on the causes of negative attitudes triggered by such factors as “unhealthiness” of voices of PwPD (Maryn and Debo, 2014; Miller, 2017).

In essence, from the combined findings of this dissertation, it becomes clear that the speech changes that reflect the manifestations of HD are expressed in the acoustics of speech irrespective of its message content. Evidence from the studies demonstrates that HD-related changes in speech are apparent to a wide population of listeners that differ in their training, native languages or familiarity with speech disorders. This dissertation also reports on the findings that support the existing evidence about the presence of language-universal speech characteristics of HD (Pinto et al., 2017). According to the results from the predictive modelling experiment, while some acoustic features are important for all listener groups, a number of demographic factors such as speakers’ accent, age or gender have very low power for predicting recognition of a certain voice as "healthy" or "unhealthy". This provides optimistic evidence for drawing conclusions about HD-specific acoustic markers rather than questions of the influence of age-, gender- or accent-related features. Thus, the universal HD speech characteristics reported in this dissertation include changes in voice quality, prosody and in vowel articulation.

Regarding the listener-related factors, expertise and language background of listeners – with respect to the task – do matter significantly. The findings on sentence type recognition accuracy are in agreement with previously reported benefits of experience for understanding speech of PwPD (Carvalho et al., 2020). Nonetheless, when it comes to more global and abstract impressions about speech of PwPD and broader concepts such as speech healthiness, the benefits of expertise and language familiarity may disappear. Comparison of the task types on the one hand, and comparison of listeners’ expertise and language background on the other hand, provide valuable insights into the complex nature of expertise phenomenon: in one task it is an asset, in the other it may lead to confusion. This highlights the methodological implications accentuating the need for adapting a rigorous approach and attention to the experimental design whenever it comes to the concept of expertise and training.

Regarding the longitudinal perspective, the densely measured acoustic change investigated simultaneously in two working languages of one bilingual speaker with PD provides unique insights into the speech change development in bilingual PwPD. The reported findings indicate a clear benefit to speech tracking in speakers with PD, which may help in the therapy management or even in the early detection of dysarthria in PD. It is precisely the case studies with many measurements that allow us to explore the process of development over time, rather than “just the results” as in pre-test/post-test studies.

This dissertation’s multifaceted ways of looking at the speech change in PwPD complement each other and forward our understanding of dysarthric speech recognition and the underlying communication issues experienced by PwPD. The findings have implications for improving speech therapy in dysarthria in PD and for raising awareness of HD in the multilingual world. Viewing dysarthric speech recognition as a multidimensional system
of variables opens the field of communication issues of PwPD towards more holistic and fine-grained speech analyses within the context of multilingualism and healthy ageing.