Potentials of Telehealth Devices for Speech Therapy in Parkinson's Disease

Lilian Beijer[1] and Toni Rietveld[1,2]
[1]Sint Maartenskliniek Nijmegen, Section Research, Development & Education
[2]Radboud University Nijmegen, Faculty of Arts, Department of Linguistics
The Netherlands

1. Introduction

Rapidly evolving technological developments have been influencing our daily lives for a few decades now. In particular information and communication technologies enable more sophisticated and faster ways of communication than ever before. These developments have far reaching consequences for professional, domestic and leisure activities. Use of computers, whether or not with an internet connection, has become very common also in the field of education and health care. The primary benefits in these areas are believed to lie in cost reduction and enhanced efficiency.

In this chapter we will focus on the possibilities of technological developments in health care, particularly for patients with Parkinson's disease. This patient group is believed to benefit considerably from innovative applications of information and communication technologies since the number of parkinsonian patients is dramatically growing due to demographic developments. Moreover, Parkinson's disease pre-eminently concerns a chronic and progressive illness, increasingly disabling these patients in almost all domains of their lives. In this chapter we will explore how telehealth technology and speech technology relates to the maintenance of their communicative competence.

2. General aspects of speech disorders associated with Parkinson's disease

2.1 General motor functions and oral motor control

Parkinson's disease is caused by the progressive impairment of neurons in an area of the brain known as substantia nigra. This is due to an imbalance in two brain chemicals (i.e. dopamine and acetylcholine) which are responsible for the transmission of nerve messages from the brain to the motor nerves in the spinal cord which control muscle movement. As a result, the communication between the substantia nigra and the corpus striatum, required for coordinating smooth and balanced muscle movement, is distorted.

The diminished functioning and coordination of respiratory, laryngeal and supralaryngeal muscles obviously affects speech, swallowing and saliva control in parkinsonian patients (Ziegler, 2003). As a consequence, the quality of speech in patients with Parkinson's disease tends to be deteriorated to some degree. As it is, dysarthria is a common manifestation of Parkinson's disease (PD) which increases in frequency and intensity with the progress of the

disease (Streifler, 1984). Hypokinetic dysarthria is mainly associated with PD; mixed dysarthria tends to occur in atypical parkinsonism.

General motor symptoms such as rigidity, bradykinesia (reduced speed of muscles), tremors or trembling are reflected in typical speech symptoms of hypokinetic dysarthria. Bradykinesia associated with Parkinson's disease causes difficulty in the initiation of voluntary speech. This can result in delay in starting to talk as well as very slow speech. According to Duffy (1995), there may be freezing of movement during speech. Rigidity can also occur. Additionally, parkinsonian patients have reduced loudness, imprecise consonant production, reduced pitch variability and festinating speech. The latter can result in extremely fast speech together with short rushes of speech (Ferrand and Bloom, 1997). Perceptual features of parkinsonian speech associated with hypokinetic dysarthria, are a weak, breathy (hoarse) voice, monotone and monoloud speech, low volume, articulatory imprecision and rate disturbances (Darly et al., 1969). The syndrome of parkinsonian dysarthria is by no means homogeneous with respect to speech rate. This might be due to different consecutive stages in the development of parkinsonian dysarthria or to different degrees of impairment (Ackermann & Ziegler, 1991). In general, more pronounced phonatory than articulatory disturbances tend to occur as far as clinical-perceptual ratings are observed.

2.2 Speech intelligibility

As a consequence of distorted oral motor functions and coordination, speech intelligibility in patients with PD obviously tends to be reduced. Prevalence studies point out that about 70% of patients living at home with Parkinson's disease have speech complaints (Kalf et al., 2008a), which are mainly associated with hypokinetic dysarthria. Diminished communication skills, in addition to the fact that these patients are increasingly disabled in their physical condition and motor abilities in the course of their disease, frequently lead to patients experiencing a deteriorating quality of life (Slawek et al., 2005). Improving or at least maintaining speech quality as long as possible is essential to enable optimal social participation and to maintain relationships. As a consequence, patients with PD are often eager to practice and improve their speech quality.

Although the majority of patients with PD are dysarthric speakers, only a minority of this group with diminished speech quality receives speech therapy (20-30%). The small percentage of PD patients under 'speech therapeutical control' might be partially due to the fact that a majority of speech language therapists consider themselves not capable to adequately treat dysarthric speakers with PD. Another reason might be the fact that speech therapy often is provided to patients that have only recently been diagnosed with PD. Therapists lose sight of their patients once face-to-face therapy sessions have been completed. Thus, dysarthric patients tend to be deprived of speech therapy in the chronic and deteriorating course of their disease. As it is, this tendency of 'undertreatment' of dysarthric patients with PD is likely to continue in the course of the coming years. That is, the incidence and prevalence of PD will increase due to our aging population. It is estimated that in 2030 about 30% of our population will be 65 years of age or older.

2.3 Guidelines for diagnostic and treatment procedures

Apart from the observed 'undertreatment' of patients with PD, large variability exists in therapeutical approaches of this patients group. In the Netherlands, evidence based guidelines for diagnostic and treatment procedures for patients with PD were developed in order to

provide speech-language therapists with recommendations for their clinical practice (Kalf et al., 2008b). It should be noted that methodological quality of comparative studies is often insufficient to meet the conditions of highest level of evidence (Deane et al., 2001). Therefore, evidence is mainly based on comparative studies of less methodological quality, noncomparative studies or experts' opinions. Five key points for treating dysarthric speech in patients with PD were formulated in the evidence based guidelines (Kalf et al., 2008b):

1. Patients with PD have basically normal motor skills, requiring to be elicited in an adequate way.
2. Hypokinesia increases when duration and complexity of motor acts increase. Therefore, complex acts should be divided into more simple acts.
3.   Separate acts should therefore compensate for failing automatic motor acts
4. External cues could support initiation and continuation of motor acts.
5. Simultaneous execution of motor and cognitive tasks should be avoided, since execution of motor tasks already puts considerable demands on cognitive functions.

For diagnosis and treatment of dysarthria in patients with PD, two procedures are strongly recommended. 1) As far as diagnostic procedures are concerned, the initial situation should be assessed by documenting spontaneous speech and establishing to what extent speech can be stimulated by means of maximum performance tests. 2) For treatment, the Lee Silverman Voice Treatment (LSVT) (Ramig et al., 2001) and the Pitch Limiting Voice Treatment (PLVT) (de Swart et al., 2003) are strongly recommended. The LSVT focuses on tasks to maximize respiratory and phonatory functions in order to improve respiratory drive, vocal fold adduction, laryngeal muscle activity and synergy, laryngeal and supralaryngeal articulatory movements, and vocal tract configuration. The PLVT also aims at increasing loudness but at the same time sets vocal pitch at an adequate level. The LSVT and the PLVT produce the same increase in loudness but PLVT limits an increase in vocal pitch and claims to prevent a strained or stressed voicing (de Swart et al., 2003). Both therapy programs concern intensive training periods of four sessions weekly during a training period of four weeks. Intensive speech therapy is preferred if diagnostic results allow highly intensive and frequent training. That is, voice quality, intrinsic motivation, physical condition and cognitive abilities are vital conditions for intensive training of newly acquired speech techniques. In case a patient's condition does not allow intensive training, augmentative and alternative procedures or devices could provide a solution to the communication problems experienced by dysarthric speakers with PD.

In the next sections, we will go into more detail with respect to current trends in speech therapy for patients with PD. These result from rapidly evolving developments in information-, communication- and speech technology. Not only will these developments provide patients with new therapy facilities; they are also expected to bring about some crucial changes for health care providers (i.e. speech therapists) and influence health care processes.

3. Current trends in the therapy of speech disorders related to Parkinson's disease

3.1 Increased need for speech training

A considerable percentage of PD patients experience oral motor disorders, causing problems with swallowing, speech and saliva control. With 70% of PD patients being dysarthric, it is obvious that therapeutical interventions are required. That is, the speech of PD patients with

predominantly hypokinetic dysarthria, needs treatment in order to improve speech intelligibility. Since communication skills are vital for adequate social participation, improvement of these abilities can significantly contribute to quality of life.

A number of current trends seem to influence the developments in speech therapy for parkinsonian patients. Firstly, there is an increased attention for dysarthria and its treatment. This is partially due to the results of scientific research in the field of PD, enhancing care givers' awareness of the relevance for long lasting communication skills in parkinsonian patients. Secondly, recent social and demographic developments have caused patients to be more aware of possibilities for treatment and to be more assertive in their call for adequate information. Patient centred health care has even gained considerable importance for reimbursement companies that find themselves increasingly confronted with clients searching for the best quality of care. Apart from this, the economic instability in this decade urges the health care community to treat the growing number of elderly patients with neurological diseases with less financial means. It is obviously a challenge to maintain a sound balance between the need (and call) for speech training on the one hand, and the availability of professionals and financial means for speech training on the other hand. Particularly with current speech training programs for PD such as the LSVT (Ramig et al., 2001) and the Dutch PLVT (de Swart et al., 2003), involving intensive speech training for several weeks to enhance speech intelligibility, it becomes clear that traditional speech therapy does no longer meet the actual needs of our current society.

3.2 Telehealth in the field of speech-language pathology

Telehealth applications, resulting from recently developed information and communication technologies in health care, could provide solutions to overcome barriers of access to therapy services caused by factors such as decreasing financial resources, shortage of professionals and increasing number of clients. The terms 'telemedicine' and 'telehealth' are sometimes used interchangeably. Telemedicine is considered a subset of telehealth. Telemedicine uses communication networks for delivery of healthcare services and medical education from one geographical location to another, primarily to address challenges like uneven distribution and shortage of infrastructural and human resources (Sood et al., 2007). 'Telehealth' is a broader term and does not necessarily involve clinical services. It can be defined as the use of telecommunication technologies both to provide health care services and to enable access to medical information for training and educating health care professionals and consumers. As such, telehealth concerns all applications of information and communication technologies, enabling the retrieval, recording, management and transmission of information to support health care. In this chapter, we refer to this latter definition when discussing telehealth.

Mashima and Doarn's (2008) overview of telehealth activities in the field of speech-language pathology provide a strong foundation for broader applications of telehealth technologies in this area. Also telehealth applications for treatment of patients with neurogenic communication disorders have been reported. Theodoros et al. (2006) report an online speech training for PD patients which turned out to be effective. Ten patients with PD followed the LSVT online using video conferencing, during a four-week program of intensive training, involving 16 therapy sessions. Comparison of sound pressure level, pitch measurements and perceptual ratings from audio recordings pre- and posttreatment, containing participants' reading and conversational monologue, showed significant

improvements, comparable to previously reported outcomes for the LSVT when delivered face-to-face. This example shows that remote diagnosis and treatment of speech in parkinsonian patients has vital benefits, in particular for patients who are less mobile and easily fatigued due to their deteriorated physical condition. Ziegler and Zierdt (2008) report an online version of a computer-based intelligibility assessment tool: the Munich Intelligibility Profile. The web based MVP-version is reported to have potentials for dysarthric speech of patients with PD and other underlying neurological diseases such as stroke.

3.3 E-learning based Speech Therapy (EST)

Quite recently, in the Netherlands a web based speech training device 'E-learning based Speech Therapy' (EST) has been developed (Beijer et al., 2010a). EST primarily aims at patients with dysarthric speech resulting from acquired neurological impairment such as stroke and Parkinson's disease. According to our clinical experience, these patients suffer from their deteriorating quality of speech. Particularly in the chronic phase of their disease, once therapy sessions have been completed, the lack of practice results in diminished speech intelligibility. With verbal communication being a vital condition for adequate social participation, diminished abilities in this field can be considerably invalidating. A vital benefit of EST is the possibility to follow a tailor-made speech training program in the patients' home environment. That is, time, energy and costs normally involved with speech training can be reduced for these patients who tend to be less mobile and easily fatigued due to their physical condition. In addition, the possibility to practice speech in the home environment at any moment, allows intensive speech training, which is known to be effective in patients with acquired neurological diseases (Kwakkel et al., 1999). Repetitive training in chronic phases also has been proven to have positive effects on speech intelligibility (Rijntjes et al., 2009).

Since telehealth applications tend to differ in many respects, Tulu et al., (2007) made an effort to provide insight into the large number of innovative web based devices that are available. They introduced a taxonomy of telehealth applications along five dimensions: communication infrastructure, delivery options, application purpose, application area and environmental setting. According to this classification, EST concerns a store-and-forward web application for treatment (i.e. training) purposes in the area of speech pathology that is commonly used in the home environment.

The keystone of the EST infrastructure is formed by a central server. The server hosts two types of audio files: target speech files in MP3 format and recorded speech files uploaded by patients in wav format (Figure 1). A desktop computer or a laptop with internet connection provides users with access to the server. Using their EST therapist account, therapists are able to, at a distance, provide their patients with a tailor-made speech training program, which is compiled from audio examples of target speech, stored at a central server.

Patients have access to this program using their client account. In the EST training procedure patients listen to audio examples of target speech which is downloaded from the server. Subsequently they imitate the audio example, in order to approach the target speech. The target and the own speech are then aurally compared. Finally the patients' speech is uploaded and stored at the server (Figure 2).
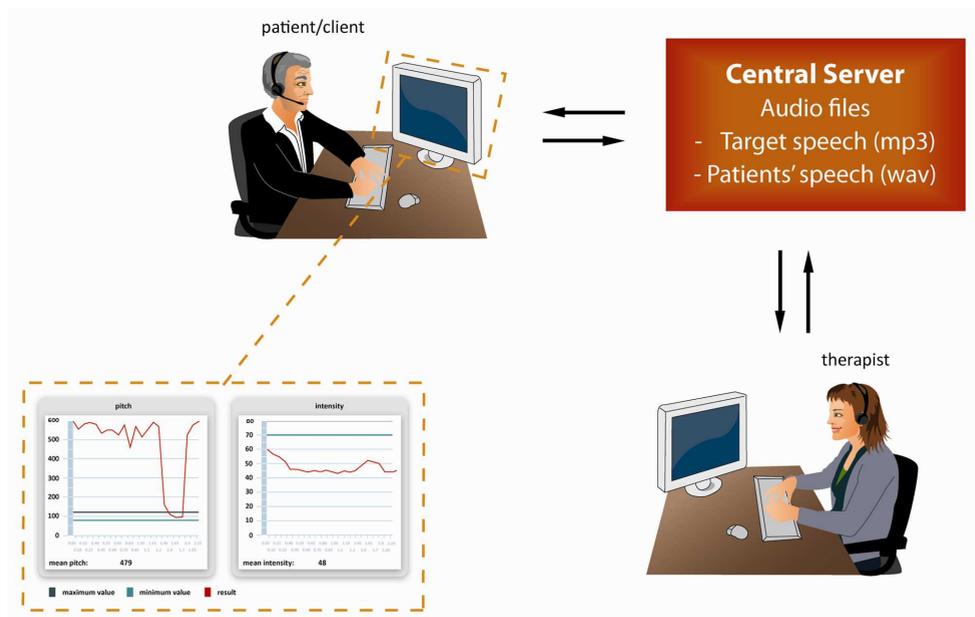
Fig. 1. Infrastructure of E-learning based Speech Therapy (EST). (Reproduced with credit of the Telemedicine and eHealth journal)

Obviously, this training procedure puts considerable demands on patients' auditory speech discrimination skills. However, indications have been reported that patients with PD experience problems with estimating the own speech volume (Ho et al., 2000) and with auditory speech discrimination (Beijer, Rietveld & van Stiphout, in press). Although this diminished auditory discrimination might be caused by cognitive problems and hearing loss, these patients would benefit from additional visual feedback on their own speech realization. That is, visualization of speech might support them in the auditory discrimination task of the EST training procedure. Although this visualization is already implemented in EST, the abstract graphs (Figure 1) and the delayed, post hoc display of visual feedback did not appear suitable for all patients (Beijer et al., 2010b). Therefore, the development of an intuitive visualisation of loudness and pitch is currently underway in order to apply to patients with various backgrounds (i.e. educational levels, age, gender). Not only should the graphic form of the visual feedback apply to the patients, indicating into what direction a new speech attempt should be adjusted to approach the target. It should also be assessed to what extent different visualisations contribute to the improvement of speech intelligibility. In section 4.4. we will go into some detail with respect to visual feedback on pitch and intensity (loudness).

Therapists are allowed to download audio files of their patient's speech from the server. Thus, they are able to listen to their patient's speech at different points across time. In addition they may analyze the acoustic speech signal for objective measures of speech dimensions that are relevant for an individual patient.
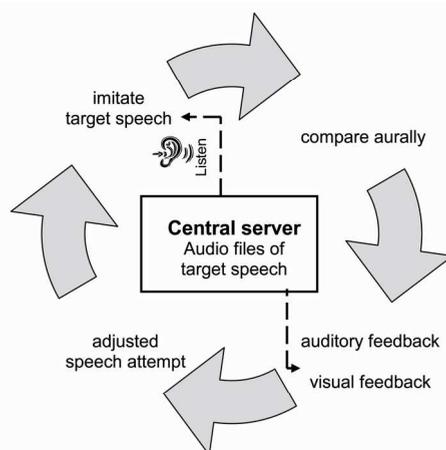
Fig. 2. EST training procedure. (Reproduced with credit of Telemedicine and eHealth)

Despite the improvements to be made, a case study conducted with a male patient with PD suggested that EST is a suitable web based speech training device with potential efficacy for patients with PD (Beijer et al., 2010b). The patient had completed face-to-face sessions of PLVT practice, and was able to conduct the training program that he was already familiar with, independently at home. He followed an intensive, protocolized four-week program, involving the PLVT (de Swart et al., 2003) by means of EST. His speech intelligibility had significantly improved immediately after the EST training period. Speech intelligibility was measured by the percentage correctly (orthographically) transcribed words in semantically unpredictable sentences (SUS). After several weeks without practice, the patient's speech intelligibility declined. Apparently, practice of speech maintained speech quality. As mentioned in section 1.3, the PLVT primarily aims at improving speech intensity (the acoustical correlate of perceived loudness) with, for vocal hygiene reasons, limited vocal pitch raise and laryngeal tension. It appeared that the participant appreciated weekly contact via telephone with his speech therapist. Apparently there was a need for additional therapeutical suggestions and a therapeutical relationship. Nevertheless, the results of this case study are hopeful. Currently, the efficacy and the user satisfaction of the web application EST are subject of investigation.

3.4 Research issues for EST

It will be clear that innovative web based applications for diagnostic and treatment purposes should be evaluated from several perspectives. First of all, the technological feasibility should be proven. Secondly, patients as well as therapists should be able to operate the web based devices. Hence, user satisfaction should be evaluated since this is obviously vital for successful implementation. The term 'user satisfaction' needs to be accurately defined to ensure comparison of user satisfaction across time and across different web based devices. This brings us to the need to establish minimum user requirements regarding physical condition, motor coordination skills and auditory or cognitive abilities. Assessing these conditions for successful use of web based devices is vital for parkinsonian patients who tend to experience constraints in more domains than communication or speech alone. Thirdly the efficacy and the

effectiveness of EST should be evaluated. This brings us to the vital issue of reliable outcome measures for treatment outcomes. In the case of parkinsonian speech, these treatment outcomes primarily concern speech intelligibility. Most of these outcome measures concern subjective, perceptual measures of speech quality (FDA, rate scaling, etc.). Along with health care reimbursers' call for objective outcome measures however, current trends point into the direction of objective acoustical measures of speech quality as a vital outcome measure for speech intelligibility in addition to traditional perceptual measures.

Employment of web based applications for diagnosis and treatment tend to go perfectly along with the need for speech technological developments. That is, speech data can be easily collected, thus generating an automatic data base of pathological speech. We will elaborate on this in section 4.5.

4. The role of speech technology in speech therapy

4.1 Introduction

For more than 25 years phoneticians, speech technologists and speech therapists have systematically investigated the phonetic correlates of speech disorders. These investigations were carried out with a number of explicit or implicit objectives: a) to corroborate subjective judgements of speech therapists, b) to find objective evidence for progress as a result of therapy, c) to facilitate the distinction of subgroups of pathologies, and d) to find evidence for theories on the nature of pathologies, which could not be obtained on the basis of subjective measurements. As has been the case in phonetics sciences, the progress of computer and information technology made available a number of additional applications, which form the core of the current chapter:

a.   Gathering objective evidence based on acoustic and/or physiological data,
b.   The development of systems which can be used by patients to obtain direct or indirect feedback on their realizations in a training program,
c.   The implementation of feedback systems in telehealth applications in order to facilitate intensive training at home.

In the following we will focus on a number of applications of speech technology to be used in the assessment and treatment of dysarthria in general and that of dysarthria associated with Parkinson's disease in particular. We should be aware of the fact that phonetics and the associated speech technology are language bound, that is to say that phenomena which are relevant in one language, may be irrelevant in another.

Kent et al. (1999) published a seminal overview of acoustic correlates of quite a number of phenomena associated with dysarthria. The overview distinguishes the conventional phonetic components of the speech production process: Initiation, Phonation, Articulation, Velopharyngeal functioning and Prosody. As is the case with most acoustic correlates of speech disorders, stochastic relations between perceptually distinctive disorders on one side and acoustic correlates on the other are more evident than inferential procedures which boil down to statements like: the F1 and F2 (first and second formants) of segment X are higher/lower than 'normal', so we can be sure that this segment was not realized in a canonical way. The fact that trade-off relations exist between speech production and speech perception is the nuisance factor. Trade-off relations in phonetics occur when an effect in one domain – say segment duration – can compensate for the absence of a feature in another domain, say voicing. In English, for instance, a relatively long pre-consonantal vowel can perceptually compensate for an obstruent which is incorrectly realized as voiceless.

The presence of this kind of trade-off relations is an obstacle in finding clear and unambiguous acoustic correlates of the perception of speech and speech disorders. This fact is not a direct problem in group studies, which aim at finding tendencies in signal characteristics between pathological groups and a control group. In set-ups in which the aim is to provide stable and robust feedback to a patient, the presence of trade-offs can be disturbing.

Providing instrumental feedback to speakers has quite a long history. As a matter of fact, there are two parallel developments. One development focuses on learning a foreign language ("L2"), and the other on correcting speech disorders. It is quite obvious why these developments are parallel: the dimensions on which deviations of target speech can be projected are – most of the time – equal or similar: prosodic dimensions, dimensions of segmental quality, phonatory dimensions and dimensions of velopharyngeal functioning. Like in speech pathology, it is hardly ever the case that all dimensions are equally relevant. For French as L2, nasality is more important than for English or Dutch. Intonation and tone is extremely important for languages like Chinese, and much less important for English, French or Dutch. These facts have directed research both in systems which provide feedback in L2 learning and in speech therapy. Until now, it has proven not to be worth the effort in this context to assess all characteristics of speech which are imitations of target speech. It is better to direct efforts to specific segments which are known to be vulnerable and/or relevant in L2 learning and speech pathology. This brings us to two different   lines in feedback:

a.   Direct, quasi real time feedback on the realization of global parameters like intensity, tempo and intonation and parameters associated with segmental quality, and

b.   Indirect, post-hoc feedback on the realization of speech parameters.

Direct feedback is meant to help the patient in non-face-to-face training sessions; indirect, post-hoc feedback is often only needed when the therapist has to have access to assessment scores; it is only available after quite a number of speech materials have been collected.

There might be a misunderstanding when it is decided to provide web based speech therapy, in the sense that it is often assumed that a computer system which provides feedback is immediately applicable in an e-health application. That is not true. Supervised training/learning often cannot be directly applied in an environment in which direct assistance is absent. Supervised learning is much more robust than its non-supervised counterpart. An example is provided by Carmichael (2007). In his study, which aimed at the development of objective acoustic measures for the Frenchay Dysarthria Assessment Procedure (Enderby, 1980), the calibration of the 'loudness' measurements might be somewhat complicated to be performed at home. It involves the use of a Sound meter at a standard distance. In his set-up the test administrator performs the calibration procedure. In order to avoid this kind of calibration at home, we opted in our telehealth application for the production of a long nasal consonant [*mmmm*]. As the production of this consonant does not involve any mouth opening and variable jaw movements, the radiated sound can be assumed to be quite constant, and to function as a reliable calibration.

Speech technology in the context of speech pathology can be divided in a number of approaches, which also depend on the objectives to be achieved. In this chapter we restrict our review to applications in assessment, therapy and training. We distinguish five dimensions on which the approaches can vary:

Dimension I: Either the parameters focus on global parameters like intensity, tempo and pitch, or on characteristics which reflect segmental quality.

Dimension II: There are two types of results to be obtained, viz. global assessment, or direct feedback.

Dimension III: Types of speech: the assessment of free speech, or the assessment of read, known speech.

Dimension IV: The inclusion of physiological parameters, like reflexes and respiration.

Dimension V: The inclusion of facial expressions as parameter(s).

A more general dimension is the user-interface. Of course, the interface for the therapist requires less attention, but the one for the patient asks for robustness and psychological validity.

Dimension I: In specific therapies, like the PLVT (de Swart et al., 2003), global parameters are of great importance. As explained in section 1.3 of this chapter, it involves two therapy goals: "speaking loud" while not increasing pitch at the same time. The rationale is that speaking loud generally leads to an increase of articulatory precision, while increasing pitch above habitual level may harm the vocal cords.

Dimension II: Global assessment – not to be confused with global parameters – involves the assessment of speech on a long-term scale. That is, direct feedback is not provided, only feedback after some amount of speech materials has been realized, recorded and analyzed. In an application for direct feedback, the user is provided with quasi real time feedback on the quality of the speech parameters at issue: global parameters like intensity and F0, or feedback on specific segments, like vowels or consonants.

Dimension III: Very often, known and consequently read speech will be used in assessment and therapy sessions. The automatic recognition of speech and the detection of deviations from it, are enormously facilitated by the use of this kind of texts (it implies what is called forced recognition). The drawback is that the use of read speech may decrease the ecological validity of the measures and indices thus obtained.

Dimension IV: As is well-known, the initiation phase in speech, which refers mainly to respiration, is crucial for the generation of speech. There are, to our knowledge, no applications available yet which provide assessment and feedback on initiation (respiration) parameters.

Dimension V: In a number of speech pathologies the assessment of facial expressions is a relevant issue. This is also the case with dysarthria. The recognition and assessment of facial expression demand dedicated software, which is quite difficult to tune to the demands of the patient and/or therapist.

4.2 Realization of assessment and feedback systems

For the realization of feedback and assessment on global parameters (F0, Intensity), relatively simple algorithms are needed, often implemented in current software packages for signal analysis, like PRAAT (Boersma & Weenink, 2011). The problem there is not the analysis of the parameters itself, but the display of the results and the feedback on deviations from the goal values. No significant changes in the detection of the global speech parameters are to be expected, but work has to be done in order to provide displays which facilitate insight in possible errors and stimulate improvements.

For the realization of feedback and assessment of segmental quality, speech technology comes to play. There are a number of approaches, depending again, on the objectives of the application: direct/indirect feedback on the realization of each target speech sound

(phoneme), direct/indirect feedback on the realization of single words, direct/indirect feedback on the realization of a short text, direct/indirect feedback on fixed or free texts, and feedback on the overall intelligibility of words and texts. The two main technical approaches are: the analysis of speech based on Automatic Speech Recognition (ASR) and ASR-free analysis of speech (Middag et al., 2010). If ASR is used, the Hidden Markov Model (HMM) is the main tool. HMMs constitute the default tool for automatic speech recognition, although other approaches are also possible (Middag et al., 2010). Hidden Markov Models are based on probabilities of states of speech segments and transitions from one state to another, or to the same state. In light of the popularity of HMMs we present a short account of this approach below. To illustrate HMMs we give a fictitious example of the use of an HMM for the recognition of the vowel /i/. The acoustic parameter used is the second formant (F2) only. In that respect the example is already fictitious: HMMs hardly ever use formants as acoustic parameters, let alone just one formant. The default parameters used to reflect the spectra of the sound segments are Mel-Frequency-Cepstral Coefficients (MFCCs), which also take into account human perception (Davis & Mermelstein, 1980).
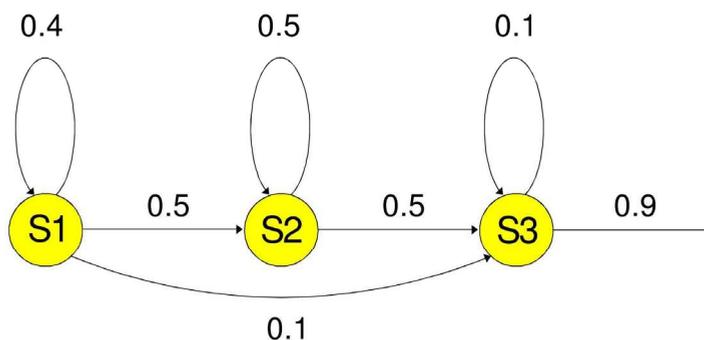

Fig. 3. A hypothetical Hidden Markov Model of the vowel /i/. (Reproduced from Rietveld & van Heuven, 2009, with permission).

In the above figure we display a model of the second formant (F2) of the vowel /i/. We see an inner loop from state 2 to state 2, which occurs with a probability of 0.5; this means that the model has a relatively high probability of staying in state 2, which boils down to the realization of a long vowel. The probability of going from state 1 (initial state) to the final state (S3) is small, only 0.1. The probabilities of obtaining discrete values of F2 for state 2 (low, rather low, rather high and high: E1, E2, E3, E4) are 0.05, 0.10, 0.15 and 0.70 respectively. Thus the probability of finding a high value of F2 in the middle of the vowel /i/is rather high.

In reality we observe small speech frames, with – in our restricted and hypothetical example – only values of F2. The sequence of values of the F2, for instance E2, E4, E4, E4, E3 (each frame covering 20 ms), has to be compared with the probabilities implied by the model. The model which has the largest probability of having generated the observed sequence of

states, will be labelled as the 'realized segment'. Of course, there will be differences in confidence that a sequence X should be labelled as segment /x/.

A confidence index is a possible measure of the quality of the realized segment. This procedure is used in an ASR application for the detection of errors in L2 (Cucchiarini et al., 2009). There is a complication. In most speech recognition algorithms, a so-called 'language model' plays a role. That language model contains transitional probabilities of going from one word to another. In Dutch, for instance, the probability of finding a word with the neutral gender – like '*house*' – is extremely low after the non-neutral determiner 'de' ('the'). In some ASR approaches language models even pertain to phoneme sequences. In Dutch, for instance, the sequence /l r/ is very low, whereas the probability of observing a sequence /s t/ is relatively high. Of course, this language knowledge should be used in speech recognition, but not in a system that aims at assessing the quality of speech segments. Knowledge of the language model is a well-known obstacle to the subjective assessment of speech. We all know that in '*the cat p...*' '*p*' will be followed with a high probability by '*urrs*'. That is why subjective measurement has an intrinsic problem: the expectation of the listener.

HMMs are used in a number of formats, depending on:

- the number of states used in modelling speech segments,
- the amount of training material needed,
- the dimensionality of the statistical distributions of the parameters used.

A word-based account of errors is often not informative, even if the language model in the HMM is "switched off". The reason is that segment mispronunciations have to be weighted in order to obtain a valid error score for an utterance (Preston, Ramsdell, Oller, Edwards & Tobin, 2011). That is why most systems developed for providing feedback on the adequacy of pronunciation are segment based. Before a robust ASR system can be set up to provide feedback on speech performance, it has to be established to which extent target speech segments meet the following criteria, which, as a matter of fact, are quite similar to the criteria used in systems for error detection in second language acquisition (see Cucchiarini et al. 2009).

a. The influence of an incorrect realization has an impact on intelligibility and communication; this implies that for every language different segments and features are important. Tonal movements are less important for languages like English or German, but crucial in Chinese.
b. The errors are perceptually salient;
c. The errors are frequent;
d. The errors occur in the speech of relatively many speakers;
e. The errors are persistent;
f. Robust automatic error detection is possible;
g. Unambiguous feedback is available.

4.3 Speech materials to be used

An often neglected subject is the nature of the speech materials to be used. Of course, the materials should contain language samples which are prone to be incorrectly realized – see above -, but there are also other aspects which should be considered. There are a multitude of factors which affect the realization of speech segments and global parameters

of speech. We mention: the prosodic position: in the English word *rhododendron*, for example, '*den*' carries word stress. Word stress affects duration, intensity and spectral characteristics, the length of the utterance: the longer an utterance is, the shorter the speech segments which make it up are (the 'i' in '*stride*' is shorter than in '*side*', the distinction between function words and content words (for instance '*in*' vs. '*bin*') is influential in speech tempo, intensity and spectral characteristics. If speech materials are to be used in subsequent assessment procedures, it is worthwhile to have the speech segments realized in balanced conditions.

4.4 Technical requirements

The American Telemedicine Association published a valuable list of "Core Standards for Telemedicine Operations" (2007). For speech applications an additional number of technical requirements have to be met, as a function of the goals of the application. Most of them are self-evident. We mention a number of requirements, and will give some more information on requirements which are less self-evident:

1. *The presence of a reliable and robust server*, with personnel that can answer technical questions at well-defined time intervals and can update the system as required by IT-developments (firewalls, browsers etc.);
2. *A cross-platform browser-based application* which delivers uncompromised viewing of applications;
3. *A clear distribution of roles* with an associated system for authorizations: user, therapist and administrator;
4. *Quick uploading* of target utterances;
5. *A psychologically valid and quick presentation of feedback with sufficient screen resolution.*
   The configuration of visual feedback for speech is not self-evident, and needs some scrutiny in order to adapt it to the user population. In this domain, speech technologists should be supplemented with experts in the integration of auditory and visual perception (Sadakata et al., 2008). For patients with neurogenic communication disorders such as PD, who are likely to suffer from other disorders than distorted speech alone, visual or cognitive distortions might be a serious constraint in the perception of visual displays that aim at providing feedback on different speech dimensions, such as pitch and intensity (loudness). This is particularly the case when two or more speech dimensions of a dysarthric speaker are displayed, such as pitch and intensity in the case of the PLVT for parkinsonian speakers. Rather than abstract graphs, which are difficult to interpret for a large number of predominantly elderly speakers, visualization should be simple and intuitive. That is, the form of visual feedback should apply to patients and give cues for approaching and adequate realization of speech.
   An example: how should one display the time course of pitch and intensity, as a simple graph (see Figure 4a), or as a picture which might intuitively be more appealing (Figure 4b)? The solution might obviously lie in an integrated, multidimensional picture of more speech dimensions.
   Currently, in the Netherlands, web based experiments are set up in order to evaluate what graphic form appeals to healthy controls. In addition, it will be evaluated whether or not preferences for visual forms in healthy controls also goes for neurological patients such as speakers with PD or after stroke.

Fig. 4a. Separate displays of the time course of pitch and intensity (female speaker)
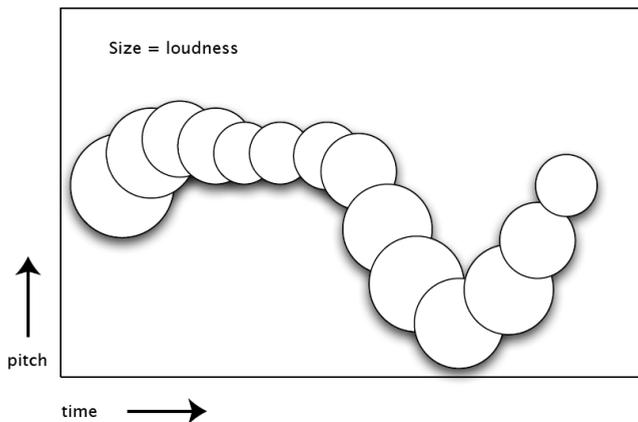

Fig. 4b. Integrated display of the time course of pitch and intensity

6. *Easy use of the PC/laptop;*
7. *Adjustable text fonts;*
8. Possibility to *personalize* protocols for exercises;
9. *Privacy guarantees;*
10. *Adequate format of speech files.*

    If (subsets of) realized utterances are stored for subsequent assessment by a speech therapist or an automated computer procedure, the format of the speech files should be suited for those procedures. The main element of the format is the sampling frequency (22.05 kHz or 44.1 kHz). Preferably no signal coding should be used to reduce the amount of data. MP3 coding and the associated data reduction (with bit rates of 128 or 192) does not have any effects on perception, but may lead to some effects on the

spectral representation. WAV-files have an advantage: they are files without any data-coding. Important factors in the decision on the sampling frequency and the possible data reduction are the characteristics of the parameters to be extracted from the signal. The upper bound of the frequency range relevant for the acoustic description of vowel-like sounds is around 3 kHz. For the analysis of fricatives – for instance speech sounds like /s/ and /ʃ/ - we need a wider frequency range, with an upper bound of at least 6 kHz (Olive et al., 1993).

11. *Robust and well-defined recording conditions.*

The basic principle underlying eHealth applications is that patients can use the application at home. Conditions at home vary to a great extent. Some people will use the application in a quiet office, others in a kitchen with neon tubes, or in a garden with traffic noise in the background. In the following figure we show the waveform of an utterance with an added 50 Hz-signal; the latter is not a pure sinus wave. The sinusoidal signal might hinder subjective judgement, and create biases in the spectral representation of the realized utterances, see panel (b) in Figure 5, where a strong 50 Hz component and the associated harmonics are displayed. The presence of the harmonics is due to the fact that the sinus wave was not 'pure'.
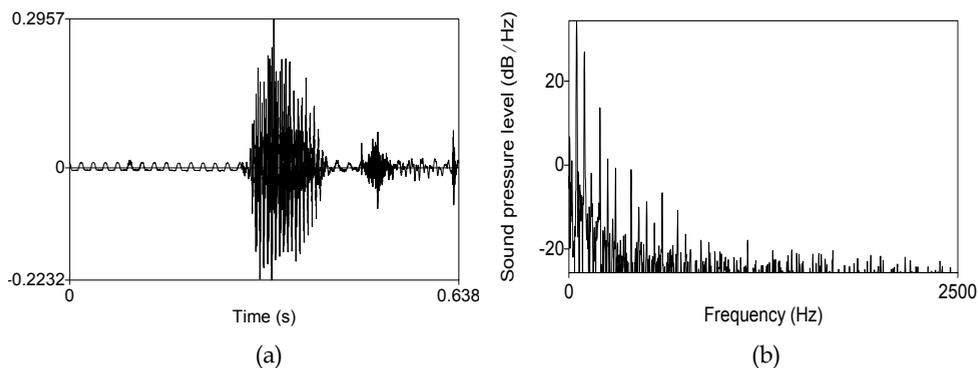

(a)   (b)

Fig. 5. (a) Waveform of a fragment of the Dutch word 'buitenboord' ([bœytəb...]:'English: 'outboard') with an added 50 Hz signal generated by an external source and
(b) spectrum of the (intended silent) initial part of the waveform.

In many approaches to the (semi-automatic) assessment of the segmental quality of speech segments, the silent interval associated with the closing phase of stop consonants like /p, t, k/ is relevant. If this interval is filled with some 'humming' it may be an indication that the speaker was not able to firmly close his/her lips for the stop consonant (Kent et al., 1999).

4.5 Secondary outcomes of telehealth applications for speech technology

Apart from therapeutical aims, which mainly focus on the benefits for clients and therapists, telehealth applications such as EST provide a vital source of data for researchers in the field of speech pathology and speech technology. That is, uploading patients' speech by means of web based systems, automatically generates a data base of pathological speech (Figure 6). This data base is vital for clinical outcome research in the field of speech pathology. That is, a data base allows perceptual and acoustical measurements of speech across time in order to evaluate therapy outcomes. This will increasingly gain importance in the context of

decreasing financial resources for health care, where evidence based treatments finally will prevail. Although guidelines for diagnosis and treatment have been formulated (section 2.3), these are not based on the highest level of evidence. Objective outcome measures on the basis of a central data base of pathological speech are likely to enhance evidence based guidelines. Government policies and hence requirements of health care reimbursers will be based on objective therapy results, to be derived from data sources as generated by web based applications such as EST. For therapists the objective speech data over time of an individual patient is expected to provide useful information to evaluate therapy results and to adjust therapy focus if necessary.

In addition to its relevance for clinical outcome research, a data base of pathological speech contains vital information for speech technological and language technological research. In general, speech and language technology has considerably gained importance in health care during the last few decades. Particularly for patients with communicative problems this research area is of vital importance. These problems can be due to cognitive disorders (e.g. aphasia or dyslexia), sensory disorders (blindness or hearing loss) or voice and speech disorders (e.g. dysarthria, stuttering, dysphonia). Speech and language technology also applies to the needs of patients with communication problems in a broader sense. That is, constraints in the interaction with their environment as a result of motor system disorders such as Repetitive Strain Injury (RSI) or movement disorders such as paralysis after stroke or distorted arm movement coordination due to PD. Only recently, a report on needs and future possibilities for speech and language technologies for patients with communicative problems appeared (Ruiter et al., 2010).

Applications of speech and language technology are expected to contribute to more efficient and more effective health care. Patients can stay longer in their home environment without putting demands on health care givers and, hence, on financial resources. For example, patients with PD could benefit from speech synthesis applications for text-to-speech conversions, facilitating patients with severely diminished speech intelligibility in their verbal communication. Automatic error detection could provide parkinsonian patients who are eager to practice their speech in their home environment using web applications such as EST, with automatic feedback on segmental speech quality (i.e. articulation of speech sounds). An ASR application in dysarthric speech for example would lie in the field of domotica. PD patients with severe motor constraints could gain considerable independence from remotely (i.e. speech) controlled domestic equipment.

In general, speech synthesis, as applied in text-to speech conversions, is usually relatively simple and is not dependent on features of pathological speech. Automatic speech recognition (ASR) and automatic error detection of pathological speech however, are complex issues in the field of speech technology. This is primarily due to the large variability within and between pathological speakers, in particular in the case of neurogenic speech disorders such as dysarthria. Hence, large amounts of data are required for the development of ASR and automatic error detection in pathological (i.e. dysarthric speech). Applications of automatic error detections concern for instance feedback on segmental speech quality in EST, in addition to feedback on loudness and overall pitch. This would enhance patients' independent web based speech training.

Obviously, apart from research in the field of speech technology, an automatically generated data base of pathological speech, is an essential source for additional fundamental research into acoustical features of parkinsonian dysarthria for instance. Outcomes of acoustical studies might even lead to adjustments of speech training programs for patients with PD.
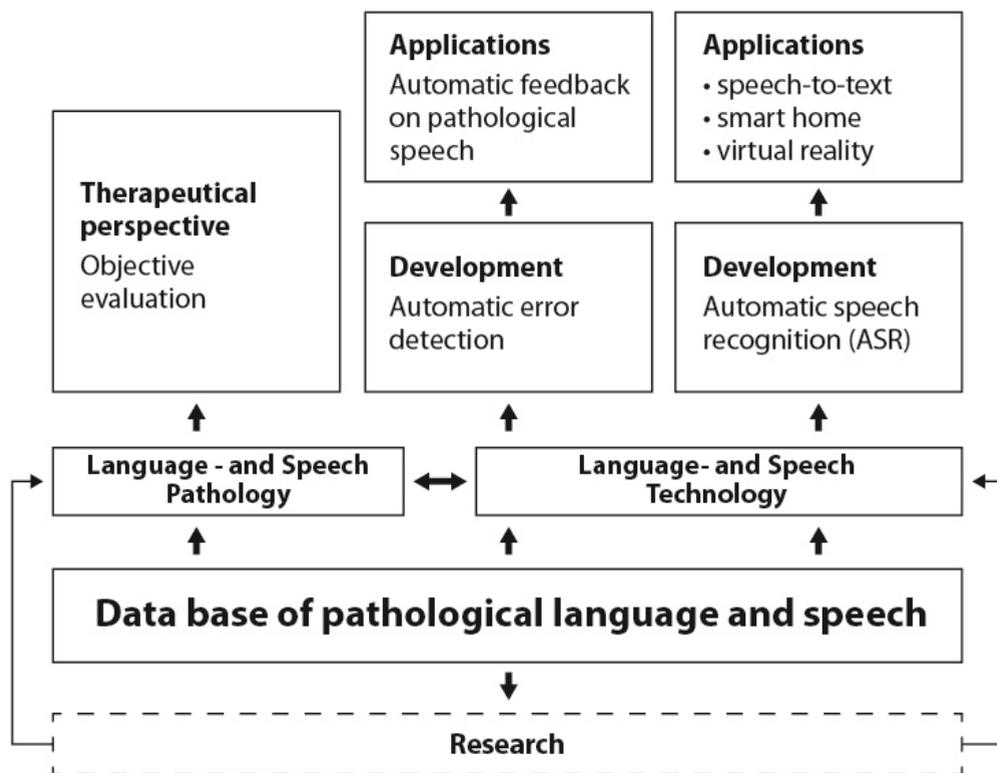
Fig. 6. A data base of pathological speech as a vital source for clinical outcome research and speech and language technological research.

A data base of pathological speech should contain speech at various linguistic levels. Audio recordings stored at the data base should be adequately annotated. That is, identification of standardized speech tasks, orthographic and phonetic annotation and linguistic level should be well documented. In addition, anonymous speaker identification should be ensured. An adequately structured data base should facilitate researchers' search for audio files of pathological speech. In Belgium, the Corpus of Pathological and Normal Speech (COPAS) (Middag et al., 2010) has been collected. Researchers employ the COPAS data base for the development of an automated intelligibility assessment, based on phonological features. These phonological features refer to articulatory dimensions. This information should reveal underlying articulation problems in dysarthric speakers. The Nemours Data Base of Dysarthric Speech is another example of a corpus of pathological speech (Menendez-Pidal et al., 1996). It should be noticed however that the COPAS and the Nemours data bases were not generated by a web based system, whereas a data base generated by means of EST involves upload and storage of audio files by means of a telehealth application. Vital conditions must be met to ensure audio recordings with adequate quality for perceptual and acoustical assessment. Obviously, a data base of pathological speech is language specific. Cross-language comparisons however should be enabled by similar structures of speech data bases for different languages.

5. Research areas with respect to the development, implementation and evaluation of telehealth applications for speech training of patients with Parkinson's disease

Telerehabilitation has a potential in a large number of fields. We are still in the first phase of a development which may revolutionize medical care and cure. The heterogeneous applications – be they in psychiatry, asthma or diabetes care, speech and language therapy – share a number of factors which have to be fulfilled in order to warrant success, but which are not always met yet. A dangerous aspect of the phase we are in now is that we focus on technology and just admire its realized or promised possibilities. Thus we might overlook the key human factors for telerehabilitation applications in general as reviewed by Brennan and Barker (2008). In this section we give an overview of the four research areas formulated by the American Telemedicine Association (Krupinsky et al., 2007), and will zoom in on those aspects which are relevant for telehealth applications for people with Parkinson's disease.

a) Attention should be paid to definition of infrastructure and integration of various infrastructural components of web based devices

Of course, the definition of infrastructure and the integration of various infrastructural components of web based devices is a prerequisite for the application and evaluation of results obtained with web based devices (see section 3.3), but the central component of the applications in our field remains the availability of robust speech recognition and/or error detection systems, if at least providing automatic feedback on realized utterances in speech training is (one of) the goal(s) of the web-application. Both speech and language pathology and speech and language technology are language bound. That is, they share underlying principles (HMMs, for instance, are used for very diverse languages like Chinese, English or Russian), and pathological reduction of vowels, consonants and pitch excursions occur in all languages, but the details of the technologies have to be developed for every specific language and the phenomena associated with speech pathology of specific languages have to be studied. In this context much attention has to be paid to the form of feedback, as was already pointed out in section 4.4.

Speech disorders and the associated symptoms may show considerable variability between and within speakers. While intensive training may help some patients to partially recover and improve their speech skills, other patients will show no improvement or perhaps even deterioration. Therefore, novel speech recognition and natural language processing techniques will have be developed that can cope with the dynamics of the speech disorder.

b) Clinical utility of telehealth should be established

The establishment of the clinical utility of a telehealth application for speech therapy of patients with PD is not a simple one. The first prerequisite is the presence of a robust infrastructure and a robust feedback system. While students using a feedback system on the quality of their pronunciation of a second language may be "robust" themselves, we cannot assume the same extent of robustness in patients with PD. That is why a very small number of studies have been conducted on the clinical utility of eHealth applications in our field. In order to test the clinical utility in phase III or phase IV studies, quite a number of conditions have to be fulfilled:

- A clear definition of the effect one wants to attain with the application. There are different possible and/or positive effects: (1) after a pre-specified time interval the

speech quality of the telehealth group has improved more than that of the control group, (2) after the time interval the speech quality of the telehealth group was more stable than that of the control group, (3) in the long run, maintenance of achieved results is made possible by the telehealth application. Each of these possible outcomes has to be crossed with outcomes regarding the cost-effectiveness of the treatment (see under d) and user satisfaction.

- Relatively large numbers of patients are needed in order to ensure sufficient power to detect possible differences between a treatment and a control group. In light of the heterogeneity of the patients with respect to a large number of relevant background variables (age, SES, cognitive and motor skills, hearing and vision, computer skills, mobility, home situation), matching on these variables is a crucial issue in the effect studies to be carried out. As it is not very simple to include large numbers of patients, it is difficult to obtain a complete overview of the importance of these variables (see under c).

- An important aspect of effect studies is a clear definition of the outcome variables to be used. For telehealth applications for patients with Parkinson's disease we mention five variables which are directly related to speech dimensions:
  - Articulatory precision
  - Intelligibility
  - Naturalness of speech
  - Speech effort
  - Listening comfort

  Much research is needed to find generally accepted operationalizations of the above-mentioned variables. Reviewing the relevant international journals in this domain make**s** clear that research is still going on, and that final results are not in view. This is in contrast to related questions on the intelligibility and naturalness of synthesized speech, where researchers agreed on a number of well-documented protocols to assess these aspects of computer speech (see the Blizzard Challenge, a yearly competition among speech synthesis systems based on corpora, see http://www.cs.cmu.edu/~awb/). For a review of problems encountered in subjective methods to assess intelligibility we refer to Beijer, Clapham & Rietveld (submitted) and Hustad (2006).

  Even if the correct operationalizations are available, a number of other questions have to be answered before ecologically valid effectiveness studies can be carried out. Here are two examples of the questions still to be answered: (1) Articulatory precision can be achieved at the cost of naturalness ("speak loud", the message of the Lee-Silverman therapy): what is more important: articulatory precision or naturalness? (2) Intelligibility is obviously related to articulatory precision, but to what extent can outcomes of current intelligibility tests and tests of articulatory precision be generalized to daily life?

c) Human and ergonomic factors should be taken into account in research activities.

At first sight the condition under c) is stating the obvious; however, enthusiasm for technology might obscure the importance of these factors. It is well known that diseases like PD may come with other problems: comorbidity is not uncommon (Lowit et al., 2005). The problems may be such that a telehealth application is not suitable for a patient, neither in daily practice, nor in a research setting. For a research setting which aims at finding evidence for the effectiveness of an application itself – out of the social/psychological

context - the problem may be that some of the participants are not suited to fully appreciate the application. There are a multitude of possible reasons for this; we mention impaired auditory processing and impaired vision as possible and obvious obstacles to the use of a telehealth application. In section 3.3 we describe an Auditory Discrimination Test on a number of speech dimensions (Beijer, Rietveld, & van Stiphout, in press) to assess participants' suitability to use auditory feedback.

There are also less obvious factors which should be taken into account, and which are often also region-determined. In densely populated areas like the Netherlands, with short distances, a patient has the choice to opt for face-to-face sessions or to stay at home with an eHealth application. A number of aspects may influence the choice: (1) finances – in the Dutch context hardly ever a factor for a patient, as even taxi expenses will often be reimbursed, (2) the wish to see other people at a regular basis; this opportunity is provided by face-to-face sessions and not by a telehealth application, (3) mobility; some people are home-bound, while others may be mobile, (4) the need to be intelligible for people others than direct partners.

d) Economic analysis should point out whether the balance of costs and benefits is beneficial to the actual economic and social situation.

This research area becomes an important one against the background of an ageing population and limited financial resources. A prerequisite for cost-effectiveness studies is the availability of effectiveness measures accepted by the community of speech and language pathologists and therapists. The question how to decide whether a number of beneficial units of web based speech therapy (less effort for a listener, less repetitions needed to achieve complete understanding, less absence from home, less transport, better maintenance of communication etc.) are in a positive trade-off relation with additional costs (implementation and maintenance of infrastructure, availability of a help-desk etc.) is a matter of politics and society.

6. Conclusion

Employment of telehealth devices for dysarthric patients with PD seems promising with respect to their possibilities to practice their speech independently. As such, they might provide a solution for the foreseen imbalance between the need (and call) for speech training on the one hand, and the reduced availability of therapeutical resources on the other hand. Although technological feasibility of various web based training devices has been established, user requirements for parkinsonian patients with frequently observed deficits in cognitive and motor functioning demand further adjustments. Apart from therapeutical goals, web based training devices such as EST provide the possibility of generating a data base of pathological speech. This data base not only provides required information for clinical outcome research. It is also of vital importance for the development of automatic speech recognition and automatic error detection of pathological (i.e. parkinsonian) speech.

7. Acknowledgement

8. References

Ackermann, H. & Ziegler, W. (1991). Articulatory deficits in parkinsonian dysarthria: an acoustic analysis. *Journal of Neurology, Neurosurgery & Neuropsychiatry*, 54 (12), pp. 1093-1098, ISSN 00223050.

American Telemedicine Association. http://atmeda.org/news/definition.html. Last accessed 20.2.2011.

Beijer, L.J., Clapham, R.P., & Rietveld, A.C.M. 'Evaluating the suitability of orthographic transcription and intelligibility scale rating of semantically unpredictable sentences (SUS) for speech training efficacy research in dysarthric speakers with Parkinson's disease.' *Journal of Medical Speech-Language Pathology* (Accepted for publication, November 2011).

Beijer, L., Rietveld, T., van Beers, M, Slangen, R., van den Heuvel, H., de Swart, B. & Geurts, S. (2010a). E-learning based Speech Therapy (EST) as a web application for speech training. *Telemedicine and e-Health*, 16(2), pp. 177-180, ISSN 1530-5627.

Beijer, L., Rietveld, T., Hoskam, V., Geurts, A. & de Swart, B. (2010b). Evaluating the feasibility and the potential efficacy of e-Learning-Based Speech Therapy for speech training in dysarthric patients with Parkinson's disease: a case study. *Telemedicine and e-Health*, 16(6), pp. 732-738, ISSN 1530-5627.

Beijer. L., Rietveld, A. & van Stiphout, A. (in press). Auditory discrimination for E-learning based Speech Therapy: a proposal for an Auditory Discrimination Test (ADT) for adult dysarthric speakers. *Journal of Communication Disorders*, ISSN 0021-9924.

Blizzard Challenge: http://www.cs.cmu.edu/~awb/).

Boersma, P. & Weenink, D. (2011). *Praat: Doing Phonetics by Computer*. www.praat.org. Downloaded 8.3.2011.

Brennan, D.M. & Barker, L.M. (2008). Human factors in de the development and implementation of telerehabilitation systems. *Journal of Telemedicine and Telecare*,14, pp. 55-58, ISSN 1758-1109.

Carmichael, J.N. (2007) *Introducing Objective Acoustic Metrics for the Frenchay Dysarthria Assessment Procedure*, Unpublished PhD-thesis, University of Sheffield.

Cucchiarini, C., Neri, A., Strik, H. (2009). Oral Proficiency Training in Dutch L2: the Contribution of ASR-based Corrective Feedback. *Speech Communication*, 51(10), pp. 853-863, ISSN 0167-6393.

Davis, S. & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28, pp. 357-366, ISSN 0096-3518.

Darly, F.L., Aronson, A.E., & Brown, J.R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12*, pp. 246-269, ISSN 1092-4388.

Deane, K.H.O., Whurr, R., Playford, E.D., Ben-Shlomo, Y., Clarke, C.E. (2001). Speech and language therapy versus placebo or no intervention for dysarthria in Parkinson's disease. Cochrane *Database of Systematic Review*s, Issue 2. Art. No.: CD002814, ISSN 1469-493x.

Duffy, J.R. (1995). *Motor Speech Disorders: Substrates, Differential Diagnosis and Management.* Mosby-Year Book, ISBN 0801669448, St. Louis.

Enderby, P. (1980). Frenchay Dysarthria Assessment, *British Journal of Disorders of Communication*, 21, pp.165-173, ISSN 0007-098x.

Ferrand, C. & Bloom, R. (1997). *Introduction to Organic and Neurogenic Disorders of Communication.* Allyn and Bacon, ISBN10: 0205168671, Needham Heights, MA.

Ho, A.K., Bradshaw, J.L. & Iansek, R. (2000).Volume perception in parkinsonian speech. *Movement Disorders* 15 (6), pp. 1125-1131, ISSN 1531-8257.

Hustad, K.C. (2006). Estimating the intelligibility of speakers with dysarthria. *Folia Phoniatrica et Loopaedica,* 58, pp. 217-228, ISSN 1021-7762.

Kalf, J.G., de Swart, B.J.M., Zwarts, M.J., Munneke, M., Bloem, B.R. (2008a). Frequency of oral motor impairments in Parkinson's disease and implications for referral to speech therapists. *Movement Disorders*; 23:S328, ISSN 1531-8257.

Kalf, J.G., de Swart, B.J.M., Bloem, B.R., Munneke, M. Guidelines for speech-language therapy in Parkinson's disease (2008b). *Movement Disorders*; 23:S328, ISSN 1531-8257.

Kent, R.D., Weismer, G., Kent, J.F., Vorperian, H.K. & Duffy, J.R. (1999). Acoustic studies of dysarthric speech: methods, progress, and potential. *Journal of Communication Disorders*, 32, pp. 141-186, ISSN 0021-9924.

Krupinsky, E, Dimmick, S., Grigsby, J., Mogel, G., Puskin, D., Speedie, S., Stamm, B., Wakefield, B., Whited, J., Whitten, P., Yellowlees, P. Research recommendations for the American Telemedicine Association. *Telemedicine and e-Health*, 12(5), pp. 579-589, ISSN 1530-5627.

Kwakkel, G., Wagenaar, R.C., Twisk, J.W.R., Lankhorst G.J., Koetsier, J.C. (1999). Intensity of leg- and arm-training after primary middle-cerebral- artery stroke. Lancet, 354 (9174), pp.191-196, ISSN 0140-6736.

Lowit, A., Howell, P. & Brendel, B. (2005). Cognitive impairment in Parkinson's disease. Is it a unfied phenomenon? *Journal of Brain Impairment*, 6 (3), pp. 191-204, ISSN 1443-9646.

Mashima, P. & Doorn, C. (2008). Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health,* 14, pp. 1101-1109, ISSN 1530-5627.

Menéndez-Pidal, X., Polikoff, J.B., Peters, S.M., Leonzio, J.E. & Bunnell, H.T. (1996). The Nemours Database of Dysarthric Speech. *Proceedings of the Fourth International Conference on Spoken Language Processing.* Philadelphia, PA, USA.

Middag, C., Saeys, Y. & Maryens, J-P. (2010). Towards an ASR-free objective analysis of pathological speech. *Proceedings Interspeech 2010, Makuhari, Japan*, pp. 294-297 ISSN 1990-9772.

Neri, A., Cucchiarini, C. & Strik, H. (2006). Selecting segmental errors in L2 Dutch for optimal pronunciation training. *International Review of Applied Linguistics in Language Teaching*, 44, pp.357-404, ISSN 1613-4141.

Olive, J.P., Greenwood, A. & Coleman, J. (1993). *Acoustics of American English Speech: A Dynamic Approach.* Springer Verlag, ISBN 0-387-97984, New York.

Preston, J.L., Ramsdell, H.L., Oller, D.K, Edwards, M.L. & Tobin, S.J. (2011). Developing a Weighted Measure of Speech Sound Accuracy. *Journal of Speech, Language, and Hearing Research*, 54, pp.1-18, ISSN 1558-9102.

Ramig, L.O., Sapir, S. Countryman, S., Pawlas, A.A., O'Brien, C., Hoehn, M. & Thompson, L.L. (2001). Intensive voice treatment (LSVT) for patients with Parkinson's disease:

a 2 year follow up. *Journal of Neurology Neurosurgery and Psychiatry*, 71, pp. 493-498, ISSN 00223050.

Rietveld, T. & Stoltenberg, I. (2005). *Taal- en spraaktechnologie en communicatieve beperkingen*. ('Language and speech technology for people with communication disorders', in Dutch). Dutch Language Union, ISBN-13: 9789070593063, The Hague.

Rietveld, T. & van Heuven, V.J. (2009). *Algemene Fonetiek ('General Phonetics', in Dutch)*, 3rd edition. Coutinho, ISBN 978 90 469 0163 2, Bussum.

Rijntjes M, Haevernick K, Barzel A, van den Bussche H, Ketels G, Weiller C.(2009). Repeat therapy for chronic motor stroke: a pilot study for feasibility and efficacy. *Neurorehabilitation & Neural Repair*. 23(3), pp. 275-80, ISSN 0888-4390.

Ruiter, M.B., Rietveld, Toni C.M., Cucchiarini, C., Krahmer, E.J. & Strik, H. (2010). Human Language Technology and Communicative Disabilities: Requirements and Possibilities for the Future in Dutch. *Proceedings Seventh International Conference of Language Resources and Evaluation* (LREC), pp. 2839-2846, ELRA: ISBN 2-9517408-6-7, Malta.

Sadakata, M., Hoppe, D., Brandmeyer, A., Timmers, R. & Desain, P. (2008). Real-Time Visual Feedback for Learning to Perform Short Rhytms with Variations in Timing and Loudness. *Journal of New Music Research*, 37, pp. 207-220, ISSN 1744-5027.

Sanders, E., Ruiter, M., Beijer, L. & Strik, H. (2002). Automatic recognition of Dutch dysarthric speech: A pilot study. *Proceedings of ICSLP-2002*, Denver, USA, pp. 661-664.

Slawek, J., Derejko, M., Lass, P. (2005). Factors affecting the quality of life of patients with idiopathic Parkinson's disease – a cross-sectional study in an outpatient clinic attendees. *Parkinsonism & Related Disorders*, 11 (7), pp. 465-468, ISSN 1353-8020.

Sood, S., Mbarika V., Jugoo, S., Dookhy, R., Doarn, C.R., Prakash, N., Merrell, R.C. (2007). What is telemedicine? A collection 104 peer-reviewed perspectives and theoretical underpinnings. *Telemedicine and e-Health*,13 (5), pp. 573-590, ISSN 1530-5627.

Strik, H., Truong, K., de Wet, F., Cucchiarini, C. (2009). Comparing different approaches for automatic pronunciation error detection. *Speech Communication*, Volume 51, Issue 10, October 2009, pp. 845-852, ISSN 0167-6393.

Swart, B.J.M. de, Willemse, S.C., Maassen, B.A.M. & Horstink, M.W.I.M. (2003). Improvement of voicing in patients with Parkinson's disease by speech therapy. *Neurology*, 60, pp. 498-500, ISSN 0028-3878.

Streifler, M. & Hofman, S. (1984). Disorders of verbal expression in parkinsonism. *Advances in Neurology*, 40, pp. 385, ISSN 0091-3952.

Teasell, R.W. & Kalra, L. (2004). Advances in stroke 2003. What's new in stroke rehabilitation? *Stroke*, February, pp. 383-385, ISSN 1524-4628.

Von Campenhausen, S., Bornschein, B., Wick, R., Botzel, K., Sampaio, C., Poewe, C., et al. (2005). Prevalence and incidence of Parkinson's disease in Europe. *European Neuropsychopharmacology*, 15(4), pp. 473-490, ISSN 0924-977x.

Theodoros, D.G., Constantinescu, G., Russell, T.G., Ward, E.C., Wilson, S.J.& Wootton, R.(2006). Treatin the speech disorder in Parkinson's disease online. *Journal of Telemedicine and Telecare*, 12: S3, pp. 99-91, ISSN 1758-1109.

Tulu, B. Chatterjee, S. and Maheshwari, M. (2007). Telemedicine Taxonomy: a classification tool. *Telemedicine and e-Health*, 13 (3), pp.349-358, ISSN 1530-5627.

Ziegler, W. (2003). Speech motor control is task-specific: Evidence from dysarthria and apraxia of speech. *Aphasiology*, 17 (1), pp. 3-36, ISSN 1464-5041.

Ziegler, W. & Zierdt , A. (2008). Telediagnostic assessment of intelligibility in dysarthria: A Pilot investigation of MVP-online. *Journal of Communication Disorders,* 41(6), pp. 553-577, ISSN 0021-9924.
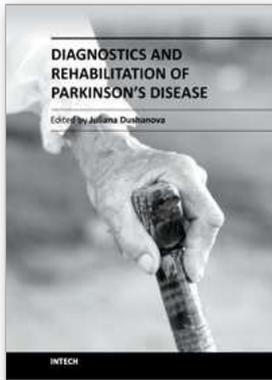
Diagnostics and Rehabilitation of Parkinson's Disease

Edited by Dr. Juliana Dushanova

Diagnostics and Rehabilitation of Parkinson's Disease presents the most current information pertaining to news-making topics relating to this disease, including etiology, early biomarkers for the diagnostics, novel methods to evaluate symptoms, research, multidisciplinary rehabilitation, new applications of brain imaging and invasive methods to the study of Parkinson's disease. Researchers have only recently begun to focus on the non-motor symptoms of Parkinson's disease, which are poorly recognized and inadequately treated by clinicians. The non-motor symptoms of Parkinson's disease have a significant impact on patient quality of life and mortality and include cognitive impairments, autonomic, gastrointestinal, and sensory symptoms. In-depth discussion of the use of imaging tools to study disease mechanisms is also provided, with emphasis on the abnormal network organization in parkinsonism. Deep brain stimulation management is a paradigm-shifting therapy for Parkinson's disease, essential tremor, and dystonia. In the recent years, new approaches of early diagnostics, training programmes and treatments have vastly improved the lives of people with Parkinson's disease, substantially reducing symptoms and significantly delaying disability. Written by leading scientists on movement and neurological disorders, this comprehensive book should appeal to a multidisciplinary audience and help people cope with medical, emotional, and practical challenges.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Lilian Beijer and Toni Rietveld (2011). Potentials of Telehealth Devices for Speech Therapy in Parkinson's Disease, Diagnostics and Rehabilitation of Parkinson's Disease, Dr. Juliana Dushanova (Ed.), ISBN: 978-953-307-791-8, InTech, Available from: http://www.intechopen.com/books/diagnostics-and-rehabilitation-of-parkinson-s-disease/potentials-of-telehealth-devices-for-speech-therapy-in-parkinson-s-disease

INTECH
open science | open minds