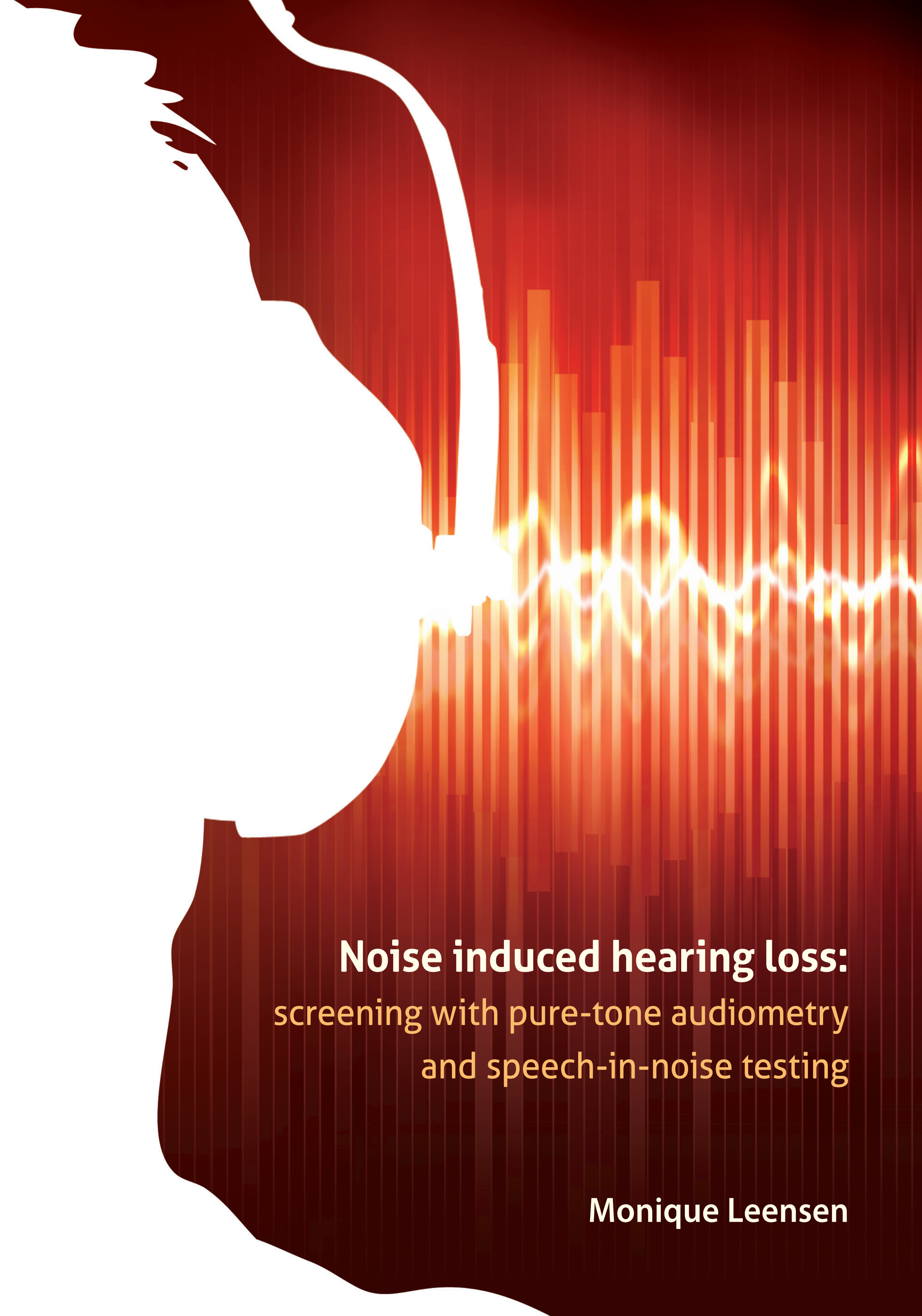
Noise induced hearing loss: Screening with pure-tone audiometry and speech-in-noise testing

Leensen, M.C.J.

Publication date
2013

Document Version
Final published version

Citation for published version (APA):
Leensen, M. C. J. (2013). *Noise induced hearing loss: Screening with pure-tone audiometry and speech-in-noise testing*. [Thesis, fully internal, Universiteit van Amsterdam].



Noise induced hearing loss:

screening with pure-tone audiometry and speech-in-noise testing

Monique Leensen

ISBN: 978-94-6191-702-7

Design cover: Promotie In Zicht, Arnhem
Lay-out: Promotie In Zicht, Arnhem
Print: Ipskamp Drukkers, Enschede

Financial support for printing this thesis was kindly provided by Stichting Atze Spoor Fonds and Nationale Hoorstichting.

Noise induced hearing loss:
screening with pure-tone audiometry and
speech-in-noise testing

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 23 mei 2013, te 14.00 uur

door

Monique Christina Johanna Leensen
geboren te Nijmegen

Promotiecommissie

Promotor: Prof. dr. ir. W. A. Dreschler

Copromotor: Dr. K. S. Rhebergen

Overige leden: Prof. dr. S. van der Baan
Prof. dr. F. J. H. van Dijk
Prof. dr. ir. J.M. Festen
Dr. ir. J.A.P.M. de Laat
Dr. B. Sorgdrager

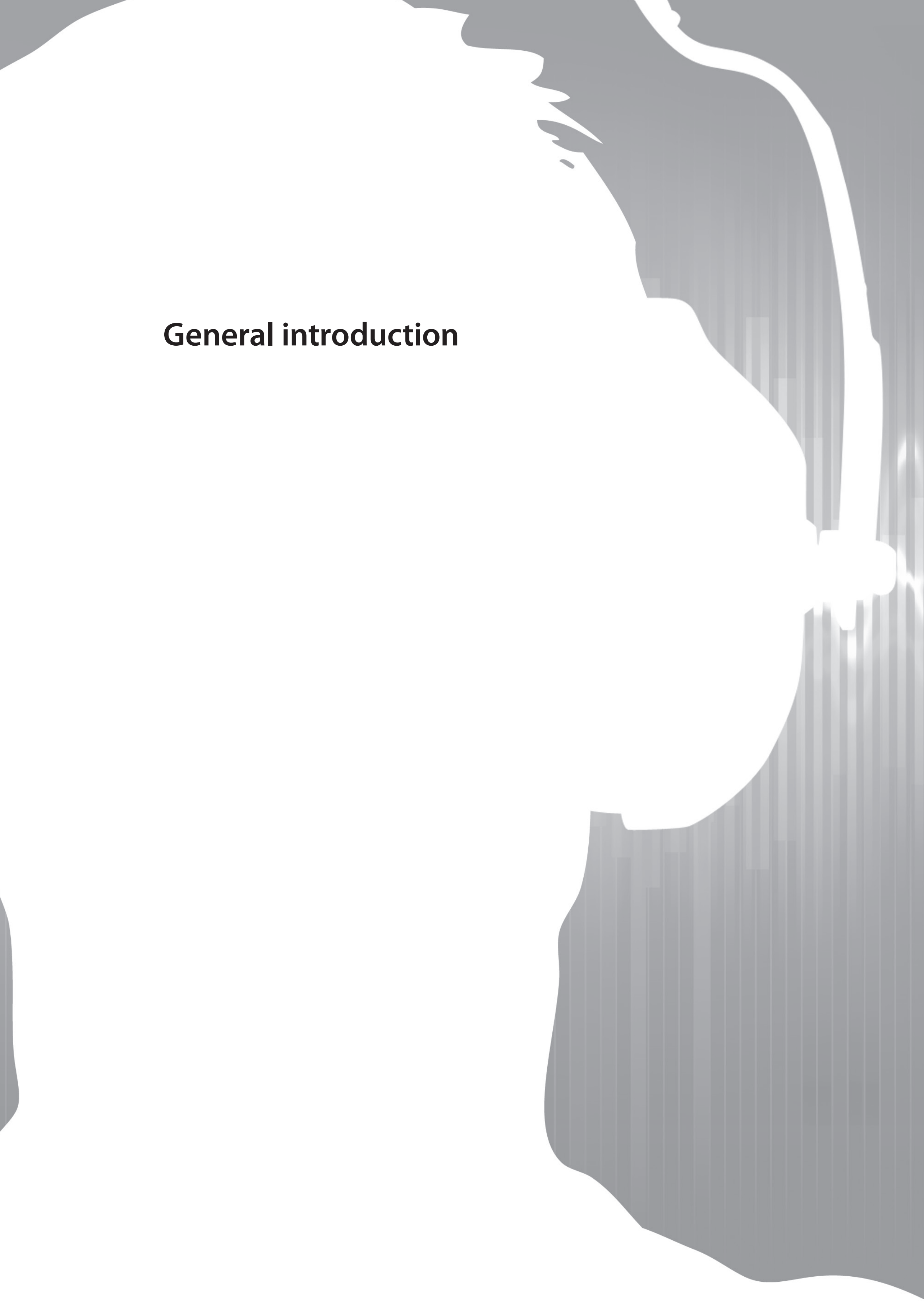
Faculteit der Geneeskunde

Voor mijn ouders

Contents

1

General introduction



Hearing impairment is a highly prevalent sensory deficit in the human population. The World Health Organization estimated that globally 278 million people have a permanent hearing loss of more than 40 dB HL (WHO, 2012), and when including milder losses (> 25 dB HL), this number increases to an estimated 642 million, which is almost 10% of the world population (WHO, 2006).

Since good hearing is essential for daily communication and social interaction, hearing damage can be seriously disabling. Worldwide, hearing loss is the second leading cause of disability (Mathers et al, 2003). Hearing impairment negatively affects physical, cognitive, and psychosocial function, by generating burdening effects such as distress, loneliness, depression, and social isolation (Mulrow et al, 1990; Carabellese et al, 1993; Cacciatore et al, 1999; Kramer et al, 2002; Arlinger, 2003; Nachtegaal et al, 2009). As a result, hearing impairment can have important implications for the quality of life (Arlinger, 2003; Chia et al, 2007)

Hearing loss is commonly classified as conductive, sensorineural, or mixed. Conductive hearing loss is caused by a mechanical defect interfering with sound transmission through the external and middle ear to the cochlea, affecting the mobility of the drum and/or the ossicles, thereby reducing hearing sensitivity (Sataloff & Sataloff, 1993). See Figure 1.1 for an overview of the anatomy of the ear. When hearing impairment is due to pathology in the cochlea or in the auditory nerve, the loss is referred to as a sensorineural hearing loss. In addition to reduced hearing ability for soft sounds, persons with sensorineural hearing loss can have suprathreshold deficits leading to the distortion of sounds (Plomp, 1986), causing difficulty in understanding speech, especially in adverse conditions such as in noise and reverberation. This type of hearing loss is largely irreversible and cannot be medically or surgically corrected.

Sensorineural hearing loss can be caused by a wide range of etiologies and its characteristics vary accordingly. The leading causes of acquired sensorineural hearing loss are age-related hearing loss, also referred to as presbyacusis, followed by noise-induced hearing loss (NIHL) (Rabinowitz, 2000; Mathers et al, 2003). Presbyacusis is a multifactorial hearing loss initially affecting the high frequencies and becoming progressively worse with advanced ageing (Albera et al, 2010).

Epidemiology of NIHL

Exposure to excessive noise causes a sensorineural hearing impairment referred to as noise-induced hearing loss. About 16% of the acquired hearing loss in adult workers worldwide is attributable to occupational noise exposure (Nelson, 2005). In the Netherlands, this is estimated to be 13 to 22% (Hoeymans et al, 2005). In addition, estimations demonstrate that 10 to 15% of the Dutch labor force is exposed to damaging noise levels during their work (Hoeymans et al, 2005). As a result, NIHL is the most frequently reported occupational disease in the Netherlands (Van der Molen

& Lenderink, 2012). Averaged over the past five years, 39% of occupational disease reports concerned NIHL, the majority of which came from the construction industry (Van der Molen & Lenderink, 2012).

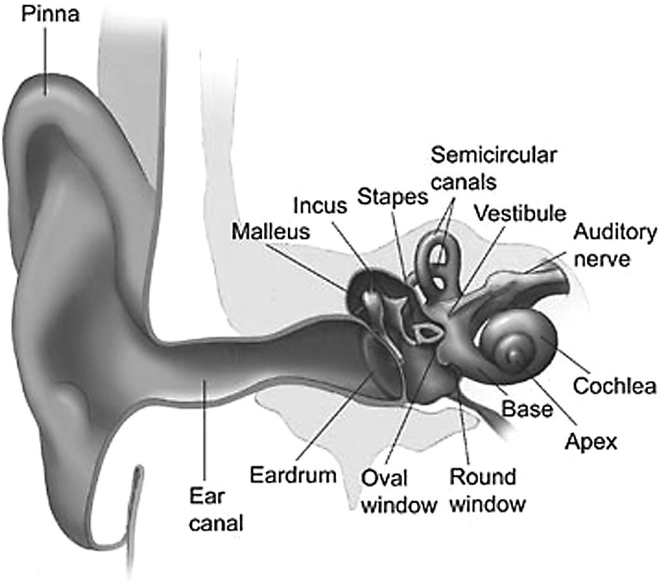


Figure 1.1. Anatomy of the ear.

Outside of work, loud sounds during recreational activities, such as visiting music concerts or dance events and listening to personal music players, may reach excessive noise levels as well. The addition of these effects is of growing concern, because an increasing percentage of noise-exposed employees also experiences exposure to noise during leisure time (Sorgdrager & Dreschler, 2010). Although evidence supporting the relationship between exposure to leisure noise and hearing loss remains ambiguous (Meyer-Bisch, 1996; Mostafapour et al, 1998; Niskar et al, 2001; Biassioni et al, 2005; Shah et al, 2009; Zhao et al, 2009), any exposure to noise of significant intensity and duration increases the risk of hearing damage. Average leisure noise levels are high enough to theoretically cause NIHL when exposed to for longer periods of time (SCENIHR, 2008). This is particularly important among those with higher susceptibility to noise (Biassoni et al, 2005) or among those who also work in a job with significant noise exposure.

Moreover, hearing losses from many causes are additive (ISO-1999, 1990; Albera et al, 2010). As a result, NIHL has become a major cause of hearing loss in the ageing population, producing hearing impairment sooner than would occur from ageing alone.

NIHL pathology and symptoms

NIHL is usually a bilateral symmetrical sensorineural hearing disorder, arising from damaged structures in the inner ear due to prolonged and repeated exposure to loud noise. The mechanism of noise-induced hearing loss involves the destruction of hair cells in the organ of Corti within the cochlea. See Figure 1.2. for a schematic representation of the organ of Corti.

The organ of Corti contains approximately 15,000 hair cells arranged in rows; one row of inner hear cells (IHCs) and three to five parallel rows of outer hair cells (OHCs). Each hair cell has tiny hair-like structures called stereocilia. When these stereocilia are deflected, ion channels are opened, causing the release of neurotransmitters by depolarization of the hair cells. By this mechanism, the IHCs are responsible for converting the mechanical vibrations caused by the movement of the basilar membrane into electrochemical impulses in the auditory nerve (Sataloff & Sataloff, 1993; Plack, 2005). The outer hair cells, on the other hand, contribute to the cochlear amplifier; they amplify the movement of the basilar membrane by contracting when stimulated by sound (electromotility), increasing the input for the IHCs in case of low-level sounds (Brownell, 1990; Plack, 2005; Gorga et al, 2007). Thus, the outer hair cells are extremely important to hearing. However, they are also very fragile, and OHCs are the structures most susceptible for damage due to noise (Henderson et al, 2006) (see Figure 1.3).

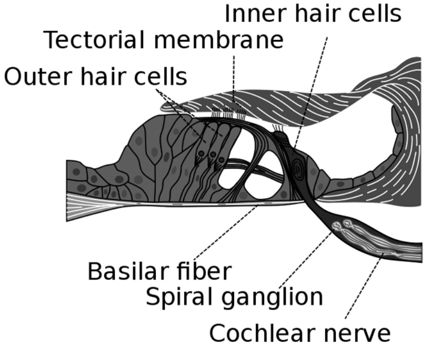


Figure 1.2. Schematic representation of the organ of Corti.

By Madhero88 2009 (Own work) [CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons

Noise can injure the ear in two different ways, depending on the type of exposure (Clark & Bohne, 1999; Dobie, 2001; Sataloff & Sataloff, 1993). Exposure to impulse noise such as explosive events, with peak levels exceeding 140 dB SPL, can directly cause mechanical damage (Clark & Bohne, 1999; Henderson et al, 2006). More common however, is the damage that develops over a longer period of chronic noise exposure that leads to several physical changes in the structures of the organ of Corti (Henderson et al, 2006; LePrell et al, 2007). Excessive noise increases the shearing movement between the basilar membrane and the tectorial membrane. As a direct result, mechanical changes in stereocilia arise (Sliwinska-Kowalska & Jedlinska, 1998); they are bended or floppy, or their tips are detached from the tectorial membrane (Gao et al, 1992; Nordmann et al, 2000). These stereociliary abnormalities are reversible over time, hence they are associated with a temporary threshold shift (TTS) (Gao et al, 1992; Henderson et al, 2006).

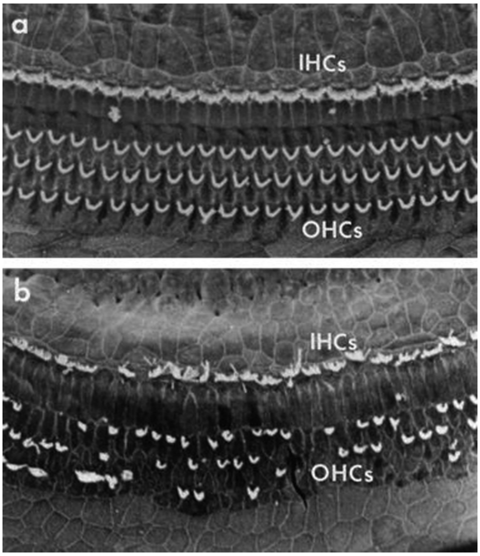


Figure 1.3. Scanning electron micrographs of the normal (a) and damaged (b) cochlear sensory epithelium. In the normal cochlea, the stereocilia of a single row of inner hair cells (IHCs) and three rows of outer hair cells (OHCs) are present in an orderly array. In the damaged cochlea, hair cells are missing, and stereocilia are abnormal, leading to hearing loss.

From Allen F. Ryan. 2000. Protection of auditory receptors and neurons: Evidence for interactive damage. PNAS, 97 (13), 6939-6940. Copyright © 2013 National Academy of Sciences, USA. http://www.pnas.org/content/97/13/6939/F1.expansion.html.

If the ear is not given a chance to rest and recover, cells experience metabolic overload and go through a cascade of chemical events that leads to cell death (Talaska & Schacht, 2007). Intense metabolic activity of the hair cells generates an overproduction of reactive oxidative species (ROS) (Henderson et al, 2006; LePrell et al, 2007). Although these are natural byproducts of normal cellular life processes, they damage cells when present in excess (Bielefeld et al, 2005). Damage from ROS triggers hair cell death due to either necrosis or apoptosis (Hu et al, 2002). Although ROS formation is not limited to hair cells, the primary damage is concentrated on the OHCs (Sliwinska-Kowalska & Jedlinska, 1998; Talaska & Schacht, 2007).

The loss of outer hair cells leads to elevated hearing threshold levels, indicating a permanent threshold shift (PTS). However, only few OHC are required for normal hearing and according to several studies, up to 30-50% of OHCs can be absent before any measurable level of hearing loss is detected by audiometry, a phenomenon called OHC redundancy (LePage & Murray, 1993; Hamernik et al, 1996; Daniel, 2007). After continued exposure to noise, the audiogram displays a classic pattern of early NIHL, showing a notch in the area of 3-6 kHz, centred at 4 kHz (Sataloff & Sataloff, 1993; Dobie, 2001; Plack, 2005). The human ear is more susceptible to cochlear damage from sound in this specific frequency region, due to primary resonances of the external ear. Hearing damage progresses steadily over the initial decade of exposure, followed by a slowing increase in hearing loss (Rösler, 1994). With more severe noise exposures, the pathology spreads to include IHC death and degeneration of auditory nerve fibers and spiral ganglions (Talaska & Schacht, 2007).

Total OHC loss causes a reduction of 50-70 dB in hearing sensitivity (Kemp, 1986; Hamernik et al, 1989; Norton, 1992; Gao et al, 1992; Henderson et al, 2006). However, a beginning hearing loss in this frequency region usually does not significantly affect speech understanding in quiet, hence it is rarely perceived. With prolonged noise exposure, damage spreads to adjacent frequencies, affecting the lower frequencies that are important for speech (Taylor et al, 1965). At this point the person becomes aware of the irreversible hearing damage that has been progressing for years (Clark & Bohne, 1999; Daniel, 2007).

From a functional perspective, noise-induced hearing loss not only leads to reduced hearing sensitivity but also to loss of cochlear frequency tuning and hence impaired frequency selectivity, reduced temporal resolution, and an abnormal increase in loudness sensitivity known as recruitment (Sataloff & Sataloff, 1993; Dobie, 2001). This usually implies poor speech intelligibility in noise (Chung & Mack, 1979; Smoorenburg, 1992; Sliwinska-Kowalska & Davis, 2012). In addition, noise exposure frequently leads to tinnitus (May, 2000; Daniel, 2007), an ongoing ringing or buzzing in the ear.

Both NIHL and tinnitus constitute major limitations in relation to hearing-critical jobs. Hearing-impaired workers have a reduced ability to detect warning signals, to communicate with coworkers, and to localize sound sources (May, 2000; Suter, 2002). Sound attenuation from the use of personal hearing protective devices in this setting is essential to prevent further damage, but may augment these implications even more (Hetu & Fortin, 1995).

ISO standards

The intensity and the duration of noise exposure both determine the degree of NIHL. Higher exposure levels and longer exposure durations cause more severe hearing losses (Taylor et al, 1965; Rösler, 1994; Dobie, 2007), although a very large inter-individual variability in susceptibility to NIHL is observed (Henderson et al, 1993). For a population exposed to noise, this relationship is mathematically described in the widely used international standard ISO-1999 (1990). With this model, the expected noise-induced permanent threshold shift (NIPTS) after a certain exposure to noise can be predicted for each frequency. These effects of noise are considered additive to age-related hearing loss (Dobie, 2001). ISO-1999 also incorporates a database for hearing thresholds as a function of age, in order to predict the total amount of hearing loss for individuals exposed to noise. This mathematical model, indicated as database A is derived from data of an otologically screened non-noise-exposed population, and allows the prediction of hearing threshold levels in relation to age, for males and females seperately (ISO-1999, 1990). Because hearing levels span a range of values, the ISO tables report median audiometric values and percentiles for a given frequency.

Occupational standards

Sound intensity is measured as sound pressure level in a logarithmic decibel (dB) scale. Noise exposure measurements are often expressed as dBA, where the 'A' represents a filter mimicking the frequency response characteristics of the human auditory system (Dobie, 2001; ANSI S1.42-2001, 2011). The logarithmic scale means that a 3-dB increase in sound level represents a doubling of the sound energy. The 3-dB doubling factor is known as the exchange rate (Dobie, 2001). The equal energy principle of noise exposure states that the amount of hearing loss caused by a sound is directly proportional to the average amount of sound energy received over time. Therefore, a doubling in noise level (i.e. +3 dB) can be offset by halving the permissible exposure duration. For example, an exposure of 88 dBA for 4 hours is considered equivalent to an 8-hour exposure to the same sound at 85 dBA (ISO-1999, 1990; Rabinowitz, 2000).

Occupational safety standards do not allow unprotected exposure to noise levels exceeding a certain limit for 40 hours a week. By exceeding these levels, a person runs a risk of hearing damage. Nelson et al (2005) report that, based on data of the National Institute of Occupational Safety and Health (NIOSH), the theoretical minimum exposure was defined as 80 dBA; a level found not to have an increased risk of causing hearing loss exceeding 25 dB HL in PTA$_{1,2,3,4}$, after 40 years of exposure (Nelson et al, 2005). A limit of 85 dBA was associated with a risk for hearing impairment of 8% and this risk was estimated to be 25% for a 90 dBA limit (NIOSH, 1998).

Specific measures for the prevention and control of exposure to noise in the Netherlands are based on the European Directive 2003/10/EC (EPC, 2003), which was adapted by Dutch national law in 2006 (Staatsblad 56, 2006). This directive states that control measures should be taken to protect workers' safety and health from the risks arising from noise exposure. These measures should be implemented in a hierarchical order, which in occupational hygiene is called the 'hierarchy of controls' (EPC, 2003; Staatsblad 56, 2006). Priority should be given to the reduction of noise exposure at its source, by implementing quieter machinery and equipment, and maintaining them properly. If this is not reasonably possible, technical (e.g. isolation of machines) or organizational measures (e.g. adaptions in the layout of workplace or work schedule) should be taken to reduce the noise exposure level or the duration of the exposure. If the risks arising from noise exposure cannot be prevented by other means, appropriate and properly fitted individual hearing protective devices (HPDs) shall be made available to workers.

The directive defines three exposure values with requirements for action, depending on the equivalent noise level for 8-hour working day:

1) A lower action level of 80 dBA, measured at ear level: employees exposed to noise at or above this level should receive information and training on the risks of exposure to noise, preventive audiometric testing should be provided and individual hearing protectors should be made available to these workers;
2) An upper action level of 85 dBA, measured at ear level: employers are required to reduce noise to intensities below this level, by elimination at its source whenever reasonably practicable or implementing technical and/or organizational measures. Workplaces should be marked with appropriate signs, and employees have the right to have their hearing checked. Individual hearing protectors should be made available to workers and should be used by them;
3) An exposure limit of 87 dBA, measured in the ear canal before the tympanic membrane: when applying this exposure limit, the attenuation provided by individual hearing protection is taken into account. Hence, this exposure level is to be measured in the ear canal before the tympanic membrane, when wearing hearing protection devices. A worker's noise exposure shall under no circumstances exceed this exposure limit. If so, immediate action should be undertaken.

Hearing ability is tested by audiometric screening. When demonstrable hearing impairment is observed, its most likely cause is determined, and the worker receives adequate audiological referral if needed (Arbouw, 2006). When the hearing loss is most probably caused by exposure to occupational noise, measures should be taken to prevent further development of NIHL in the individual worker as well as in the specific department of the company. Noise levels should be reassessed and preventative measures should be revised, and employees working in similar circumstances must be given the opportunity to check their hearing (again) (Staatsblad 56, 2006). Moreover, measures to compensate for worker's functional loss, such as technical or organizational adaptations, should be taken as well.

Prevention of NIHL

The vast majority of noise-induced hearing losses is preventable. Primary prevention can be accomplished by eliminating or reducing exposure to excessive noise. Although the hierarchy of controls should be the leading principle for reducing environment levels below the lower action level (EPC, 2003), this is often impractical and costly. Therefore prevention often relies on employee's use of individual hearing protectors rather than controlling noise exposure at its source (Neitzel & Seixas, 2005). The effective attenuation of HPDs depends on the condition of the material, the fit, and consistency of usage. Discomfort, interference with any other equipment or hinder to communication cause irregular use of HPDs (Suter, 2002; Neitzel & Seixas, 2005). Workers who selectively wear their HPDs experience greatly reduced effective protection as a result of noise exposure received during time of non-use (Gerges et al, 2001). For example, if a hearing protector has an effective attenuation of 20 dBA, and it is worn in an daily ambient noise of 100 dBA for 8 hours, then the worker will be exposed to 80 dBA if the protection is worn 100% of the time. If the same hearing protector is not used during 10% of the working day, the worker will be exposed to a time-weighted average noise level of 90 dBA.

Occupational hearing conservation also incorporates ways for secondary prevention, by means of preventative hearing testing that provides early diagnosis of NIHL. Because of its gradual development NIHL is often unnoticed; listeners are unaware that a hearing disorder is developing until hearing thresholds have dropped markedly in the range of speech frequencies. Early detection of hearing loss is therefore a crucial aspect of hearing conservation; this can increase awareness about the risk for hearing damage caused by noise and can help to prevent further hearing loss development.

Awareness and an objective assessment of hearing ability might induce behavioral changes in order to prevent NIHL. Workers who are demonstrated to have hearing loss after audiometric testing, may be much better motivated to use HPDs properly (Royster, 2003; Hong & Csaszar, 2005). However, construction workers' use of HPDs is influenced by various factors, such as workers' perceived benefits and barriers of using HPDs, perceived risk of hearing damage associated with noise exposure, and safety climate (Melamed et al, 1996; Lusk et al, 1997). Most of these are described in Penders revised health promotion model, a model shown to be useful for explaining the workers' use of HPDs (Lusk et al, 1997). Some studies have established a direct positive effect of information about the status of an individual's hearing ability on HPD use (Zohar et al, 1980; Widén et al, 2009), while other studies showed no or only limited effects (Lusk et al, 1998; Lusk et al, 1999; Williams et al, 2004; Edelson et al, 2009). Although the direct association between hearing status and HPD use is not equivocally proven, knowledge about a worker's hearing ability can affect different factors in Penders revised health promotion model, such as perceived risk of noise exposure, and benefits of reducing workplace noise, thereby indirectly affecting HPD use (Melamed et al, 1996; Purdy & Williams, 2002; Williams et al, 2004; Azeres & Miguel, 2005).

A literature review by El Dib et al. (2012), showed that interventions to influence the wearing of hearing protection improve the mean use of hearing protective devices compared with non-intervention, especially when they are individually tailored and contain mixed aspects. Hearing testing is a very important aspect of these interventions in hearing conservation; it provides an opportunity to educate workers about NIHL and motivate them to change behaviors regarding hearing protection, it is a starting point for taking (individual) precautionary measures, and it monitors hearing health of the workforce (Royster, 2003).

Although pure-tone air conduction audiometry is the general hearing screening method incorporated in occupational standards, several other possible methods for NIHL screening in occupational heath can be considered as well.

Methods for hearing screening

Pure-tone audiometry

The pure-tone audiogram is considered the gold standard for describing hearing sensitivity (Sataloff & Sataloff, 1993; May, 2000). The audiogram determines the lowest signal level a person can hear over a range of frequencies. The pure-tone hearing thresholds are used to identify and qualify hearing loss, and determine its cause. Screening audiometry involves an assessment of the hearing thresholds using air conduction under headphones only, carried out under specified conditions given in ISO-6189 (ISO, 1983). This audiometric assessment is usually part of a hearing conservation program.

However, pure-tone audiometry does not have perfect precision. Behavioral thresholds vary somewhat from one test to the next, because of tester and patient experience and motivation (Schlauch & Carney, 2012). Clinical test-retest variability, expressed as standard deviation of the difference, varies from 3 to 6.8 dB depending on frequency (Hétu, 1979; Dobie, 1983; Hall & Lutman, 1999). These values increase somewhat with larger interval lengths (Dobie, 1983).

When audiometric testing is applied in industrial screening programs, the variability may increase even more due to a number of sources of systematic and random errors. These sources may be calibration errors of audiometric equipment, excessive background noise in the testing room, residual TTS at the time of testing, partial or complete obstruction of the external auditory canal (e.g. by cerumen), interfering signals from the test equipment, differences in earphone placement, bias introduced by the tester or the examination procedure, familiarization with the examination procedure and the presence of tinnitus (Hétu, 1979). Many of these error sources can be minimized by careful control of the testing environment, cautiously following the protocol and giving good instructions (Hétu, 1979; Franks, 2001).

Adequate audiometric testing requires a quiet environment with acceptable ambient noise levels during testing, since audiometry involves determination of the lowest signal level that a person can hear (Franks, 2001). The maximum permissible ambient noise levels are specified in ISO-6189 (1983). These are rarely achieved without an audiometric soundproof booth, which is not always available in occupational assessment (HSA, 2007). The audiometers must meet ISO standard 8253-1 (1989) and need to be tested for proper function prior to each day's use, and calibrated according to ISO-389-1 (1998) annually (May, 2000). Employees need to be advised to have a quiet period of ideally 16 hours preceding audiometry, without exposure to either occupational or non-occupational noise, in orderto reduce the likelihood of TTS (Franks, 2001). Finally, otoscopic examination should be performed before testing, and findings should be noted. If significant amounts of earwax are present it may be better to advise removal of wax before performing the test (HSA, 2007), as partial obstruction of the ear leads to higher air conductive thresholds (Schlauch & Carney, 2012).

In occupational screening settings these requirements are not easily met, therefore test-retest reliability becomes reduced. Indeed, occupational audiometry is found to be less reliable than clinical audiometry; industrial test-retest variability ranged from 6.7-10.1 dB depending on frequency (Dobie, 1983; Helleman & Dreschler, 2012). As a result, small early threshold shifts for an individual employee cannot easily be distinguished from normal measurement variability (Royster & Royster, 1986), so alternative (or additional) methods were sought to improve early detection of NIHL in occupational health surveillance.

Otoacoustic emissions

One of these proposed alternatives is the measurement of otoacoustic emissions. Healthy ears generate low-level sounds that are by-products of the active, non-linear properties of the cochlea arising from the OHCs (Kemp, 1978). These sounds are known as otoacoustic emissions (OAEs) and can be recorded by a sensitive microphone inserted in the ear canal (Kemp, 2007). The presence of these emissions provide information on the function of OHCs (Lonsbury-Martin et al, 1995), the structures most vulnerable to high level noise.

The most common application of OAEs is in newborn hearing screening (Lonsbury-Martin et al, 1995; Kemp, 2007), but OAE recording is also suggested to be a sensitive method to screen for NIHL (Lapsley Miller et al, 2006; Lapsley Miller & Marshall, 2007, Marshall et al, 2009). The added value of evoked OAE (EOAE) recording in an occupational audiology environment is that it is a non-invasive objective technique that is not influenced by the patients state of consciousness, it is simple, quick and cost-effective (Chan et al, 2004; Lapsley Miller & Marshall, 2007) and it does not require a sound-proof booth but only a relatively quiet test room.

As OAEs are able to indicate small changes in cochlear function, OAE amplitude reduction can reflect OHC damage due to noise exposure (Sliwinska-Kowalska & Kotylo 2007). EOAEs may provide a more direct measurement of early changes to the inner ear than audiometry (Lapsley Miller & Marshall, 2007), and findings of audiometrically normal-hearing noise-exposed individuals having lower OAEs than non-noise controls suggests that OAEs may show noise-induced changes before they are detectable in the regular pure-tone audiogram (Lapsley Miller & Marshall, 2007). However, evidence for OAE sensitivity to detect so-called preclinical damage is equivocal (Lapsley Miller & Marshall, 2007). Most of the findings are reported by cross-sectional studies (LePage & Murray, 1993; Attias et al, 1998; Desai et al, 1999; Attias et al, 2001), and findings of longitudinal studies could not sufficiently establish this enhanced sensitivity (Engdahl et al, 1996; Seixas et al, 2005; Lapsley-Miller et al, 2004; Konopka et al, 2005; Lapsley Miller et al, 2006; Helleman & Dreschler, 2012). Nevertheless, many studies found reduced OAE amplitudes or absent OAEs as a result of exposure to noise (LePage & Murray 1993; Hotz et al, 1993; Engdahl et al, 1996; Attias et al, 1998; Desai et al, 1999; Attias et al, 2001; Lapsley Miller et al, 2004; Konopka et al, 2005; Lapsley Miller et al, 2006). In addition, high test-retest variability is observed for OAEs, which was lower than for audiometry (Hall & Lutman, 1999; Keppler et al 2010; Helleman & Dreschler, 2012). However, several aspects limit the application of detecting OAE changes in NIHL screening purposes.

First of all, the high test reliability of OAE measurements can be affected by equipment limitations and methodological issues, such as adequate calibration, the stimulus parameters used and environmental noise (Kemp, 2007; Keppler et al, 2010). Adequate probe placement is highly important for adequate OAE recording, and larger test-retest variability is found after probe refitting (Keppler et al, 2010). In addition, EOAEs are highly dependent on the forward and reverse transmission through the middle and external ear (Keefe, 2007), and tympanometric pressure has an impact on EOAE amplitudes (Kemp et al, 1990; Marshall et al, 1997). So a reduction in OAE amplitude might also reflect measurement error or a (temporary) conductive hearing loss (Kemp, 2007).

Second, OAEs only reflect OHC function and their presence does neither exclude hearing impairment caused by IHC dysfunction, nor by a retrocochlear dysfunction (Robinette et al, 2007).

Third, and most important, this method is applicable only where reliable OAEs can be recorded. There is a good correlation between OAE sensitivity and hearing threshold up to 30-40 dB HL. Above this level, there is often no recordable OAE (Kemp, 2007). This excludes the investigation of most cases of moderate to severe hearing loss, which is an important limitation for the use of OAE recordings for monitoring purposes. People in hearing conservation programs often have very low emission levels, due to presbyacusis , NIHL or both. It is important to ensure that these low-level emissions are still well above levels of ambient noise (Lapsley Miller et al, 2004). The signal-to-noise ratio (SNR), which refers to the difference between response level and the level of the background noise, can be used as a reliability estimate. Recent investigations in two Dutch hearing conservation programs showed that according to a criterion of SNR ≥ 0 dB, OAEs could not be reliably recorded in 10-45% of the noise-exposed employees investigated, depending on the frequency measured (Helleman et al, 2010; Leensen et al, 2011). For monitoring purposes, an even higher SNR criterion would be more appropriate, as that leaves enough room for deterioration over time (Helleman et al, 2010). However, using a higher SNR as reliability criterion reduces the number of valid OAE data points even more (Helleman et al, 2010; Leensen et al, 2011).

These findings indicate that OAEs can only be used as a reliable monitoring tool for a subset of an industrial population with good baseline hearing. This means that pure-tone audiometry remains necessary when a pre-existing hearing loss is present.

Speech-in-noise testing

A speech-in-noise test is a functional hearing test that also may provide a valuable method for NIHL screening. Measuring the ability to understand speech in a background noise has become a commonly used method to quantify everyday communication performance. Difficulty in understanding speech, especially in the presence of background noise, gives rise to the largest number of complaints of sensorineural hearing loss in general (Arlinger, 2003). Since speech reception in noise is highly correlated with the pure-tone average of 2 and 4 kHz (Smoorenburg, 1990; Smoorenburg, 1992), it is often the first problem experienced by subjects with NIHL. Some individuals experience these complaints even in the absence of clinically significant hearing loss in the pure-tone audiogram (Badri et al, 2011; Kumar et al, 2012).

Multiple forms of speech-in-noise testing exist, with different parameters that may influence tests results (Theunissen et al, 2009). The most important properties of a speech-in-noise test are the speech material (e.g. sentences, monosyllables, spondees), the type of masking noise (stationary noise, fluctuating noise, multitalker babble etc), and the presentation mode (fixed or adaptive presentation levels). Adaptively presenting a closed set of words in noise, makes speech-in-noise testing very suitable for automated administration, thereby offering opportunities for self-testing. Based on the adaptive up-down procedure introduced by Plomp & Mimpen (1979a), one of the first automated speech-in-noise tests was the National Hearing Test, developed by Smits et al (2004; 2006a), presenting digit-triplets in stationary noise. This fully automatic self-test for screening purposes can be administered by telephone or internet and has been very successful in the Dutch population in general (Smits et al, 2006b). However, the bandwidth of this test was limited to 0.3-3.4 kHz to mimic the telephone network frequencies. Because NIHL predominantly affects the high frequency region, another Dutch broadband online speech-in-noise test was generated; ‘Earcheck’ (Albrecht et al, 2005).

These online tests all measure the speech reception threshold (SRT); the SNR that corresponds to the ratio at which 50% of the speech is correctly understood. Because the test measures a SNR rather than absolute thresholds, this kind of testing is fairly insensitive to poor acoustics due to transduction or background noise (Smits et al, 2004; Culling et al, 2005), placing less demands on the testing environment (Jansen et al, 2010). Moreover, the SRT is not influenced by the absolute presentation level in stationary noise (Plomp & Mimpen, 1979b), requires little calibration, and is very quick (Smits et al, 2004; Culling et al, 2005; Jansen et al, 2010). Finally, speech-in-noise testing is, when presented at a sufficiently high presentation level, insensitive to conductive hearing losses (Plomp, 1986). Due to these factors, speech-in-noise tests can be implemented as an easily accessible and reliable self-screening test that can even be completed in a home setting.

The rapid growth of online screening tests on health status illustrates that the internet is a suitable medium to contact the general public (Koopman et al, 2008). Hence, the greatest advantage of an internet-based self-test for hearing screening is that it offers widespread access to testing (Swanepoel & Hall, 2010), providing a fast way to reach many employees at risk (Stenfelt et al, 2011). As a result, online hearing screening might lead to higher participation rates in hearing conservation. It also offers the opportunity to check hearing ability more frequently e.g. when complaints arise (Koopman et al, 2008), and it can be performed more easily after a period free of occupational noise, reducing possible TTS effects.

Nevertheless, although reduced speech intelligibility in noise was shown for listeners with NIHL (Chung & Mack, 1979; Smoorenburg et al, 1982; Smoorenburg, 1992; Bosman & Smoorenburg, 1995), the sensitivity of speech-in-noise testing for early NIHL has to be established. Since listeners with NIHL often exhibit (near) normal hearing thresholds in the low to mid frequencies, they can benefit from their preserved hearing when recognizing words in noise (Quist-Hansen et al, 1979). Patients with high-frequency hearing loss above 2 kHz, showed word recognition in stationary noise similar to normal performance (Pekkarinen et al, 1990; Philips et al, 1994). The sensitivity of the online speech-in-noise test Earcheck for NIHL, and its applicability in occupational health will be studied in this thesis.

Outline of this thesis

This thesis studies methods for NIHL screening and monitoring in occupational hearing conservation, as is practiced in the Dutch construction industry.

In the first part of the thesis, the value of the traditional method of pure-tone audiometry in detection of NIHL is investigated, by analysing audiometric data obtained in regular occupational health examinations of a large cohort noise-exposed workers.

Chapter 2 describes a cross-sectional analysis of audiometric thresholds of approximately 30,000 noise-exposed construction workers. Their hearing threshold levels are compared to the ISO-1999 standard, in order to assess excessive hearing loss relative to normal presbyacusis and the correspondence between observed and predicted NIPTS.

Chapter 3 presents the results of a longitudinal analysis of audiometric thresholds of about half of the baseline cohort that obtained a follow-up assessment of hearing. The development of hearing loss over a period of 4 years is investigated, and compared to ISO-1999 predictions for noise and aging. By examining consecutive datasets, the quality of audiometric data collected during screening assessments can be judged as well.

The second part of this thesis aims to develop an alternative approach for NIHL screening. Because an internet application can provide easily accessible hearing screening to a broad public, the application of online speech-in-noise testing for NIHL screening purposes is evaluated and improved.

Chapter 4 evaluates the sensitivity of three versions of a Dutch online speech-in-noise test for detecting NIHL. Since this sensitivity turned out to be rather low, ways to improve its sensitivity, and consequently the applicability for NIHL screening, were investigated.

Chapter 5 describes the results of this investigation of Earcheck adaptations, by concerning homogenization of the speech stimuli and various spectral and temporal modulations to the masking noise. Sensitivity for NIHL increased extensively when using low-pass filtered masking noise.

The third part of this thesis aims to investigate the value of this newly developed internet-based speech-in-noise test for NIHL screening purposes in hearing conservation. Because of the filtering of the masking noise, the SNR in the high-frequency region is changed, and test results might be affected by uncontrollable parameters of domestic testing. In Chapter 6, these effects are described and investigated.

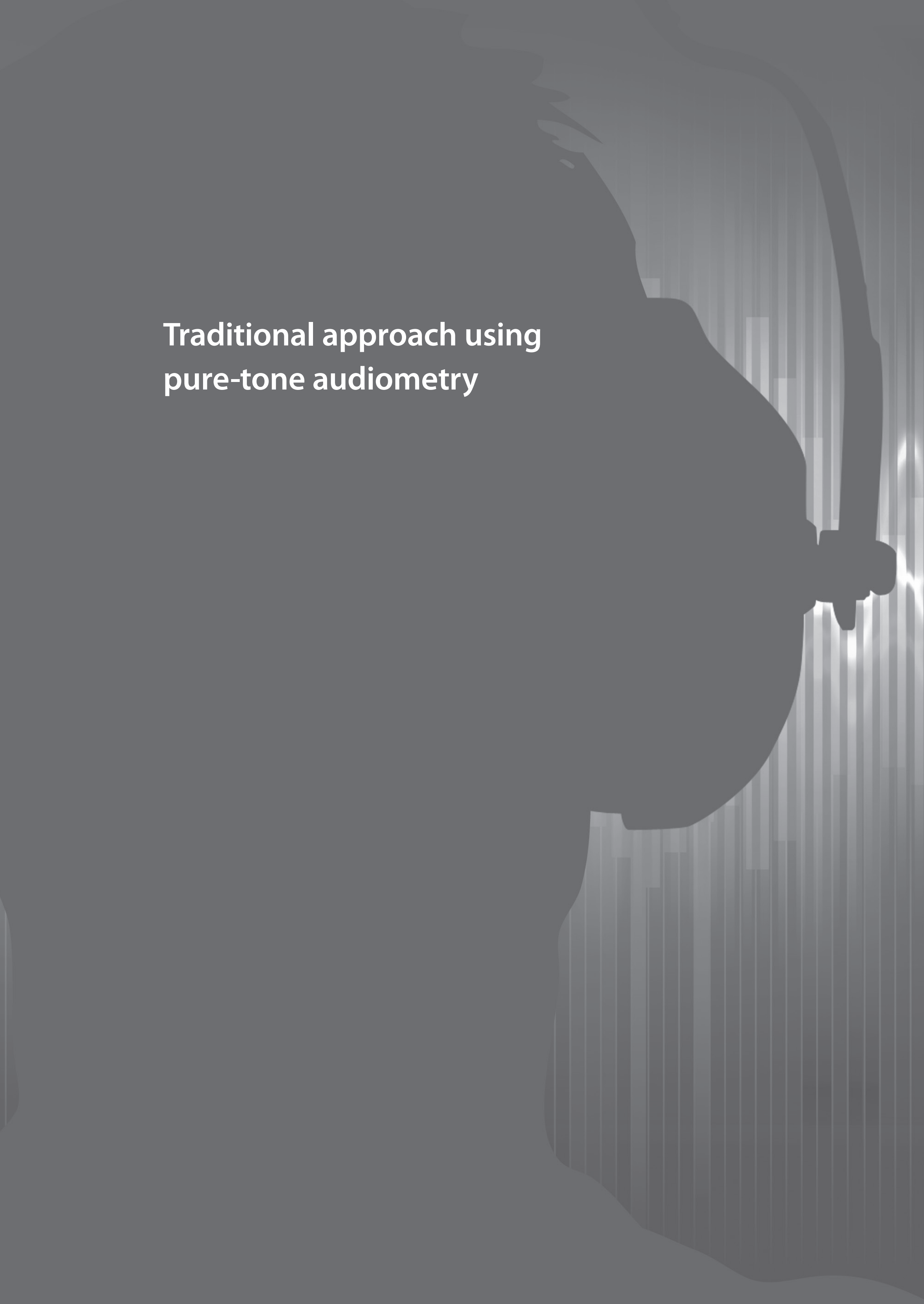
Finally, Chapter 7 describes the validity of the online speech-in-noise test compared to screening audiometry, and determines the applicability of such a screening test in occupational health.

Chapter 8 presents the general conclusions from this thesis and discusses the relevance of the current findings for occupational hearing conservation. Moreover, some suggestions are given for future research to increase the reliability and applicability of this new testing technique even more.

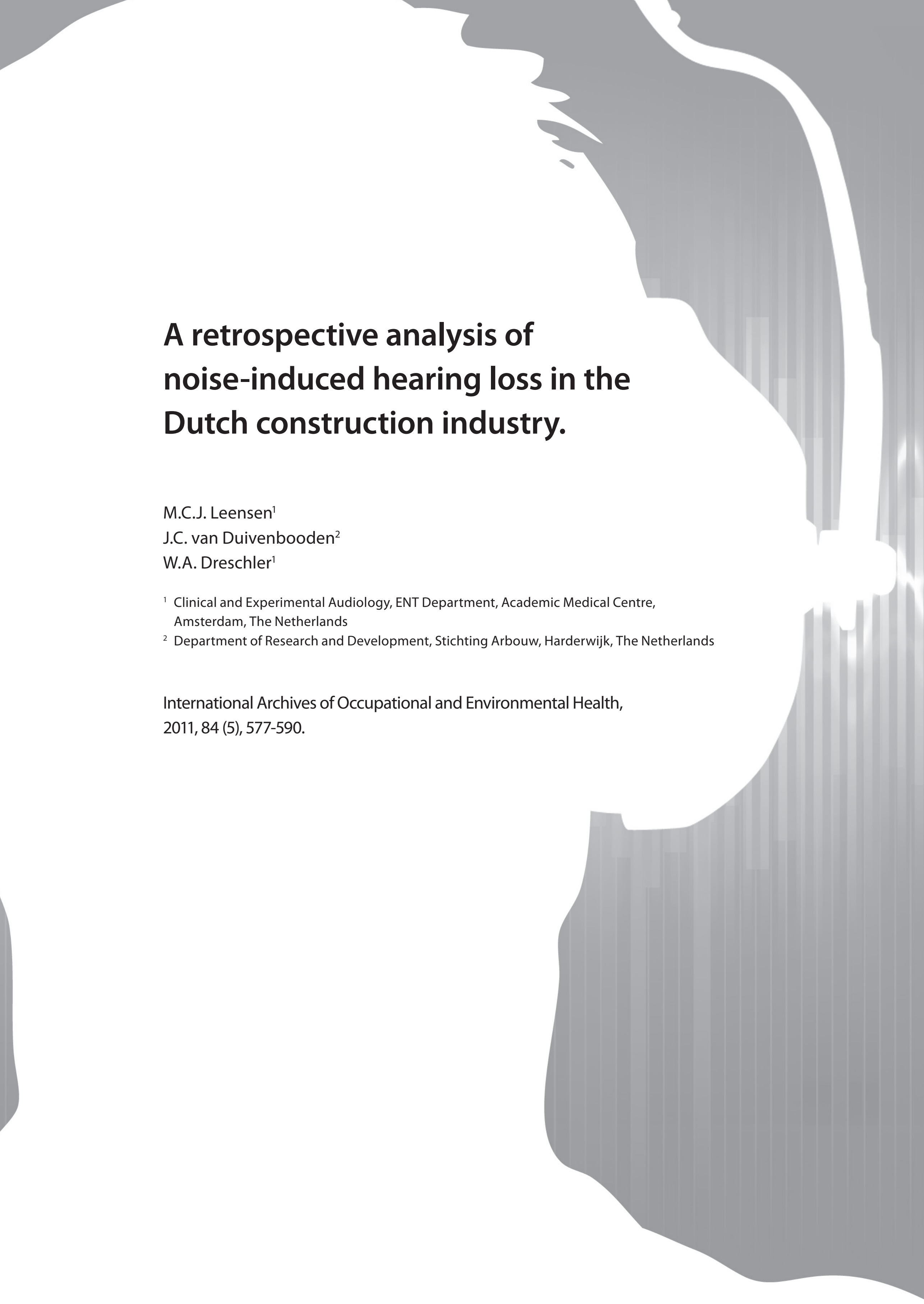
It should be noted that this thesis is composed of five papers (Chapters 2, and 4 to 7), published or submitted for publication as research paper. This means that these chapters can be read separately, but as a consequence there may be some overlap in the some sections of these chapters.

1

Traditional approach using pure-tone audiometry



2



A retrospective analysis of noise-induced hearing loss in the Dutch construction industry.

M.C.J. Leensen[1]
J.C. van Duivenbooden[2]
W.A. Dreschler[1]

[1] Clinical and Experimental Audiology, ENT Department, Academic Medical Centre, Amsterdam, The Netherlands
[2] Department of Research and Development, Stichting Arbouw, Harderwijk, The Netherlands

International Archives of Occupational and Environmental Health,
2011, 84 (5), 577-590.

Abstract

Purpose: Noise exposure is an important and highly prevalent occupational hazard in the construction industry. This study examines hearing threshold levels of a large population of Dutch construction workers and compares their hearing thresholds to those predicted by ISO-1999 (1990).
Methods: In this retrospective study, medical records of periodic occupational health examinations of 29,644 construction workers are analysed. Pure-tone audiometric thresholds of noise-exposed workers are compared to a non-exposed control group and to ISO-1999 predictions. Regression analyses are conducted to explore the relationship between hearing loss and noise intensity, noise exposure time and the use of hearing protection.
Results: Noise-exposed workers had greater hearing losses compared to their non-noise-exposed colleagues and to the reference population reported in ISO-1999. Noise exposure explained only a small proportion of hearing loss. When the daily noise exposure level rose from 80 dBA towards 96 dBA, only a minor increase in hearing loss is shown. The relation between exposure time and hearing loss found was similar to ISO-1999 predictions when looking at durations of 10 years or more. For the first decade, the population medians show poorer hearing than predicted by ISO-1999.
Discussion: Duration of noise exposure was a better predictor than noise exposure levels, probably because of the limitations in noise exposure estimations. In this population, noise-induced hearing loss was already present at the beginning of employment and increased at the same rate as is predicted for longer exposure durations.

Introduction

Noise is an important occupational health hazard, with a high prevalence in the construction industry. The noise exposure of construction workers varies greatly with the activities performed and the equipment used on the worksite (Hong, 2005), frequently exceeding daily noise exposure levels of 80 dBA, which the European Directive 2003/10/EC defines as lower action level (EPC, 2003). This directive also considers an upper action level of 85 dBA, at which the use of hearing protection is mandatory, and an exposure limit of 87 dBA that takes the attenuation of individual hearing protectors into account. Long-term exposure to daily noise levels above the lower action level of 80 dBA may eventually cause noise-induced hearing loss (NIHL), a bilateral sensorineural hearing impairment. Typically, the first sign of NIHL is a notching of the audiogram at 3, 4 or 6 kHz, with a recovery at 8 kHz (May, 2000). This audiometric notch deepens and gradually develops towards the lower frequencies when noise exposure continues (Rösler, 1994).

As a result of the high noise exposures in construction, NIHL is one of the major occupational health problems in this industry. It may have a great impact on a workers' quality of life (May, 2000), and it also influences workers' communication and safety (Suter, 2002). NIHL is the most reported occupational disease in the Dutch construction sector, with a prevalence of 15.1% in 2008 (Van der Molen et al, 2009). In other countries NIHL is one of most prevalent occupational diseases among construction workers as well (Arndt et al, 1996; Hessel, 2000; Hong, 2005) and prevalence estimations range from 10% in the USA (Dobie, 2008) to 37% in Australia (Kurmis & Apps, 2007). A large US analysis of self-reported hearing impairment in industrial sectors showed that the largest number of employees with hearing difficulty attributable to employment was found in the construction industry (Tak & Calvert, 2008).

Previous studies showed a dose-response relationship of exposure to noise and hearing loss. Higher exposure levels and longer exposure durations cause greater hearing impairment (Rösler, 1994; Prince, 2002; Rabinowitz et al, 2007; Dobie, 2007). This relationship is mathematically described in the international standard ISO-1999 (1990), predicting both the distribution of the expected noise-induced threshold worsening in populations exposed to continuous noise, and the total hearing levels resulting from NIHL in combination with age-related hearing loss. Hence, the standard also incorporates a database for hearing thresholds as a function of age, for male and female populations separately. This algorithm, indicated as database A, is an internationally well-accepted reference, derived from data of an otologically screened non-noise-exposed population.

The expected noise-induced threshold shift is a function of noise exposure level and exposure time. Characteristically, NIHL develops progressively in the first 10-15 years of noise exposure, followed by a slowing rate of growth with additional exposure to noise (Taylor et al, 1965; ISO-1999, 1990; Rösler, 1994). This pattern is represented in the ISO-1999 model. However, these predictions are based on data from subjects exposed for 10 years or more. The algorithm to predict hearing damage in the first ten years is interpolated from the predicted median NIHL after 10 years of exposure and the assumed hearing threshold of 0 dB HL at the beginning of exposure (ISO-1999, 1990), resulting in a steep linear increase in hearing loss during the first years of exposure. A study of NIHL in railway workers showed that 20% of final hearing loss at 2 and 4 kHz was already established after the first year of noise exposure. This highly exceeded the predictions of the ISO-model, yet after 3-4 years of exposure data and model are in close agreement (Henderson & Saunders, 1998). On the contrary, another study found only a slight increase in hearing threshold levels (HTLs) of construction apprentices after the first three years of employment in construction industry (Seixas et al, 2005), which was much smaller than predicted by ISO-1999.

Because NIHL is preventable, hearing conservation programmes are established, often relying on employee's use of hearing protection devices (HPDs) rather than on controlling the noise exposure at its source (Neitzel & Seixas, 2005). Protection from HPDs depends largely on the consistency of usage, because noise exposure during non-use greatly reduces their effectiveness (Neitzel & Seixas, 2005). Discomfort, hinder to communication and highly variable noise levels, which are common in construction, can cause irregular use of HPDs (Suter, 2002; Neitzel & Seixas, 2005). Several studies focusing on the use of hearing protectors in construction demonstrated low level of HPD usage; Lusk et al. (1998) found that workers in different construction trades reported to wear protection during only 18-49% of time exposed to self-reported high noise. In a more recent study this percentage was 41% (Edelson et al, 2009). Neitzel & Seixas (2005) reported an even lower percentage of usage of less than 25% of the time, which combined with the amount of attenuation resulted in negligible effective protection. Nevertheless, a study examining hearing loss in Canadian construction workers showed that HPD usage was common (>90%) and resulted in a protective effect on hearing (Hessel, 2000). These different findings underline the complicating effects of the consistency of HPD usage in assessing the relationship between occupational noise exposure and NIHL.

In addition, there is also a great variability in individual susceptibility to hearing loss (Henderson et al, 1993; Sliwinska-Kowalska et al, 2006), partly explained by other possible causes of hearing loss. These are both intrinsic and external factors (Prince et al, 2003; Sliwinska-Kowalska et al, 2006). Intrinsic factors are for example gender, race, genetics, medical history, and hypertension (De Moraes Marchiori et al, 2006). External factors concern ototoxicity, leisure noise exposure, HPD usage and smoking (Mizoue et al, 2003; Wild et al, 2005).

In this study, a large audiometric dataset of 29,216 construction workers is used to describe their hearing status. The effect of noise exposure on hearing is observed by comparing hearing threshold levels of noise-exposed workers to thresholds of references. The relationship between hearing and noise intensity and noise exposure time is examined, with particular interest in the hearing loss established during the first 10 years of noise exposure. The observed relationships are compared to ISO-1999 predictions. In addition, the influence of wearing hearing protection and other factors collected in periodic occupational health surveys on NIHL is considered.

Methods

This cross-sectional study is based on data collected by Arbouw, the Dutch national institute on occupational health and safety in the construction industry. These data are derived from medical records of periodic occupational health examinations (POHE), performed between 1 November 2005 and 20 July 2006 throughout The Netherlands.

A POHE consists of an extensive self-administered questionnaire and a physical examination, including standardized audiometric testing. POHEs are provided for all employees in the construction industry, irrespective of occupational noise exposure. The right to participate is laid down in the collective labour agreement and participation is completely voluntary.

Demographic, occupational and health-related data are extracted anonymously from the medical records. This includes information regarding job title, use of HPDs (yes/no), self-reported hearing complaints, noise disturbance at work, and the number of years employed in both the construction industry and the current occupation. Cigarette smoking status (non-/ex-/current smoker), alcohol intake (gl/wk) and blood pressure are also recorded. Hypertension is defined as systolic blood pressure ≥ 140 mmHg combined with diastolic blood pressure ≥ 90 mmHg (De Moraes Marchiori et al, 2006). Independent ethical approval is not needed for this type of retrospective analyses in the Netherlands.

Participants

The eligible study population contains all 29,216 construction workers who had undergone a POHE in the given period. Hearing threshold levels of the noise-exposed construction workers are compared to different reference groups, in order to separate the effects of occupational noise from those due to ageing and other non-occupational causes of hearing loss. The ISO-1999 standard (1990) provides two reference databases: database A, based on a highly screened non-noise-exposed population free from otologic disease, which is used in this study to correct for median age-related hearing loss; and annex B, an alternative database representing a typical otologically unscreened population of an industrialized country, not occupationally exposed to noise. This database derived from representative population-based samples can serve as an appropriate comparison group (Dobie, 2006).

The participants of the study population currently exposed to daily noise exposure levels below 80 dBA, such as office workers, can be considered as a comparison group as well. These non-noise-exposed employees are recruited from the same companies, and are examined in the same period and according to the same protocol as the noise-exposed subjects. However, almost two-third of these currently unexposed workers (65.8%) reported prior employment in the construction industry. Their past job titles, and corresponding exposure history, are unknown, but past occupational noise exposure cannot be excluded for each of these workers. Since an unscreened industrialized population should not be occupationally exposed, only the 1.016 non-exposed employees without prior employment are considered as an appropriate control group.

These controls show hearing threshold levels (HTLs) very similar to ISO-1999 database B, especially in the high frequency region (3-6 kHz). Since these non-exposed employees match the workers under consideration, they form an ideal comparison group (Prince, 2002; Prince et al, 2003). Thus, this internal comparison group is preferred over the unscreened ISO-1999 annex B to be used as control group in this study.

Audiometric measurement

Hearing ability is assessed by a qualified medical assistant using standardized audiometric examination procedures according to ISO-6189 (1983). Pure-tone audiometry is conducted at the workplaces, if possible in a mobile unit equipped with a soundproof booth, using a manual audiometer (Madsen Electronics, Taastrup, Denmark) coupled with TDH-39 headphones. Audiometers are annually calibrated according to the ISO-389.1 standard (1998). Testing is done during the work shift, but subjects had at least a noise-free period of approximately 2-3 hours prior to testing. Pure-tone air-conduction thresholds are determined at frequencies 0.5, 1, 2, 3, 4, 6, and 8 kHz in both ears, in 5-dB increments. A hearing threshold level of 90 dB HL is the upper limit of the equipment and hearing threshold is marked as 95 dB if the participant does not respond to this maximum sound signal. Because of this ceiling effect, only HTLs up to 90 dB HL or better are preserved in this analysis.

Noise exposure estimation

Duration of exposure is defined as the years employed in construction industry, as is reported in the questionnaire. If the number of years employed in construction sector exceeds the number of years in the current job, it is assumed that the former job had equivalent exposure levels.

Sound levels are expected to vary more from day to day for the individual workers than between different workers in the same trade. Therefore, workers are classified by the time weighted average (TWA) noise exposure levels estimated for standardized job titles. These daily noise exposure levels were extracted from a database of Arbouw. Most of the estimates reported in this database are retrieved from findings of Passchier-Vermeer et al. (1991). Their findings were based upon a collection of audiometric hearing thresholds of a large population of construction workers. For each profession, the noise levels were derived from the observed HTLs, using a maximum-likelihood fitting procedure in conjunction with the algorithm given in ISO-1999. A comparable approach is used more recently in a military population (Tufts et al, 2009). This way, hearing thresholds can be predicted for populations, even when noise exposure levels are not precisely known. The calculated noise level estimates are a result of all unknown aspects that may have influenced the workers' noise exposure, such as HPD use, non-occupational noise exposure, individual susceptibility and other factors. Therefore, these predictions were verified by noise measurements in 1983, 1991, 2002 and 2007. These measurements are generated by Arbouw and include full-shift personal dosimetry and sound level measurements during specified job-related tasks. Sound level measurements are combined logarithmically in order to calculate an 8-hour equivalent noise level, using the duration and frequency of each task. The daily noise exposure levels obtained by dosimetry are arithmetically averaged to obtain job-specific exposure estimations. Table 2.1 provides an overview of the available data on noise exposure estimates for the twenty most prevalent jobs in the current dataset.

The results of the noise measurements showed good agreement with the noise level calculations for the majority of job titles (Table 2.1). In case of a deviation, the result of the noise measurements was considered the appropriate noise exposure level to be used in this study. Also, the different measurements performed in different periods showed great similarity.

Exclusion criteria

Of the 29,216 participants included in this study, all 951 female workers are discarded because of their concentration in non-noise-exposed jobs. Furthermore, one subject lacks all audiometric data and 173 participants show HTLs of 95 dB HL at one or more frequencies in both ears. In addition, 357 subjects show HTLs of 95 dB in one ear and hearing threshold levels of 90 dB HL or better at all frequencies in the other ear. For these subjects, only the latter ear is preserved in the dataset.

Table 2.1. Noise exposure level estimates for the 20 most prevalent job titles, deriving from calculations and different measurements. Noise exposure levels are expressed as equivalent 8-hour, A-weighted sound-pressure levels ($L_{A,eq(8h)}$), calculated using an exchange rate of 3 dB.

Top	Job title	n	Calcu-lations	Measured sound level	Dosimetry	Noise level used
1	carpenter	10225	91		84 – 95	91
2	bricklayer	2394	91	87 – 92		91
3	painter	2082	88	80 – 90		88
4	contractor	1748	88	84 – 89		88
5	hodman	635	90	80 – 90		87
6	engineer (civil)	582	92		81 – 99	88
7	navvy	518	91	81 – 95		91
8	paver	508	91	86 – 93		92
9	plasterer	412	90	85 – 108		93
10	tiler	344	91	87 – 91		91
11	crane operator	323	92	79 – 98		92
12	driver/chauffeur	283	91			91
13	mechanical woodworker	282	93	83 – 96	87 – 95	91
14	concrete bender	237	89	82 – 89		89
15	concrete scraper	224	91	87 – 92		91
16	mecanic (machines)	214	92	90 – 95		92
17	pipelayer	200	91	85 – 95		91
18	mecanic	192	92	82 – 96		92
19	pile driver	145	96		80 – 103	86
20	destructor	140	89		81 – 109	96

Data are excluded for 447 workers with insufficient noise exposure data; they miss either information on job title (n = 19) or on duration of employment (n = 428). Finally, the 1,958 currently non-exposed workers that reported prior employment in construction are excluded from the internal control group.

The excluded participants do not differ significantly from the included subjects, except for younger age (-3.3 ± 0.5 years) and shorter employment duration (-6.0 ± 2.9

years). However, age-corrected hearing loss is similar in both groups ($p = 0.908$). The study population thus comprises 27,644 men and 54,931 ears.

Data analysis

All statistical analyses are performed using SPSS for windows software, version 15.0. Binaural average thresholds are computed for each test frequency and for each subject. If threshold levels of only one ear are available, these are regarded as the binaural thresholds and are used for analyses. Audiogram data usually have a positively skewed distribution. However, the tested sample is assumed to be large enough to approach a normal distribution and parametric tests are used (Dawson-Saunders & Trapp, 1994). The mean binaural hearing threshold levels of exposed workers are compared to age-matched ISO-standard values using a paired Student's t test, and to HTLs of the non-exposed control group using an independent Student's t test. In order to compare hearing thresholds of the exposed workers to those of controls and to NIHL predictions by ISO, HTLs of each participant are corrected for age effects by subtraction of the age-matched median HTL predicted by annex A of ISO-1999. This ISO-model assumes that noise-induced permanent threshold shift (NIPTS) and age-related hearing loss (ARHL) are additive, according to the following empirical formula:

$$\mathrm{HTL} = \mathrm{ARHL} + \mathrm{NIPTS} - (\mathrm{ARHL} * \mathrm{NIPTS})/120 \qquad \textit{(Equation 2.1)}$$

The correction term (ARHL * NIPTS)/120 starts to modify the result significantly when NIPTS+ARHL is more than approximately 40 dB HL (ISO-1999, 1990). To avoid underestimation of NIPTS in this study, this correction term was taken into account in calculating the age-corrected thresholds for measured HTLs exceeding 40 dB HL.

To simplify the results, hearing loss is also evaluated using pure-tone averages calculated for 1, 2 and 4 kHz ($\mathrm{PTA}_{1,2,4}$) and for the noise-sensitive frequencies 3, 4 and 6 kHz ($\mathrm{PTA}_{3,4,6}$). These parameters are used in multiple linear regression analyses, to investigate the dependence of PTA-values on noise intensity and exposure time. Since there is an important dependence between age and hearing loss, age is also considered as an explanatory variable. The possible statistically significant interaction of noise intensity and noise exposure time is tested by adding a product term in regression analyses.

In addition, multiple linear regression analysis is used for the analysis of combined action of different parameters on $\mathrm{PTA}_{3,4,6}$ values. Modelling proceeded in several steps. First, bivariate relationships of the covariates with $\mathrm{PTA}_{3,4,6}$ are checked by simple linear regression. All analyses are adjusted for age by including age as a covariate. Most of the categorical variables are dichotomous, and others are converted into dummy variables before inclusion into the analysis. Variables are retained for further modelling if the age-adjusted p-value of the individual testing was < 0.10. Second, a multiple linear regression model is created using the selected set of potential predictive variables. Relevant variables are selected using a backward stepwise elimination procedure, with p < 0.05 for inclusion and p < 0.10 for exclusion.

The use of hearing protection devices reduces noise exposure, which may lead to overestimation of exposure levels and attenuation of the exposure-response relationship (Sbihi et al, 2010). To reduce the effects of hearing protection, some analyses are adjusted for reported HPD use by performing stratified analyses for the subgroups of HPD users and non-users.

The level for statistical significance is taken as p < 0.01 for all analyses.

Results

General population characteristics

The total population of 27,644 men, is divided into a large group of noise-exposed employees (n = 24,670) and an internal non-exposed control group (n = 1,016). The exposed group is slightly older than the control group (average age 44.3 years and 40.9 years, respectively, see Table 2.2). Noise-exposed workers are significantly longer employed in both the construction industry and their current occupation than controls. Mean employment differences are 12.4 years and 6.7 years, respectively. More than half of the exposed workers has always been employed in the current job (55.5%). Of the exposed employees, 75.5% claim to use hearing protection, 22.1% has complaints of worsened hearing and 39.1% is bothered by noise during work. Smoking status, alcohol intake, and blood pressure do not differ between the groups.

Hearing threshold levels

To examine the hearing ability of the construction employees, median hearing threshold levels of the noise-exposed workers are compared to median HTLs of the non-exposed controls and age-matched thresholds reported in annex A of the ISO-1999 standard (Figure 2.1). All curves show the well-known deterioration of hearing with age, which is most prominent in the high frequency region. Both the exposed workers and the internal controls show significantly poorer hearing threshold levels relative to the ISO-1999 predicted values, across the complete range of test frequencies. In addition, both groups show a slight worsening in the high frequencies in the two youngest groups. In the older age groups, the differences between median HTLs of the exposed workers and the internal controls increase. These differences are greatest for hearing thresholds at 4 and 6 kHz. With increasing age, the exposed group develops a typical NIHL notching pattern in the high frequency range, which broadens from 4 to 6 kHz to the lower frequencies.

Table 2.2. Demographics and hearing loss risk factors, by subject group. Italic values represent percentages.

Variables		Exposed	Controls
n		24,670	1,016
Age, yrs (mean ± SD) *		44.3 ± 11.4	40.9 ± 11.5
Years in construction (mean ± SD)*		24.3 ± 12.6	11.9 ± 10.2
Years in current job (mean ± SD)*		18.6 ± 12.8	11.9 ± 10.2
Always employed in current job (%)*		*55.5*	-
Usage of HPD (%) *		*75.3*	*9.9*
Complaints of worsened hearing (%)*		*22.1*	*11.7*
Bothered by noise during work (%)*		*39.1*	*4.5*
Smoking	Never (%)	*35.0*	*36.4*
	Current (%)	*32.8*	*33.5*
	Ex (%)	*32.2*	*30.1*
Cigarettes/day (mean ± SD)		14.7 ± 9.9	14.2 ± 9.2
Years of smoking (mean ± SD)		18.9 ± 11.8	18.9 ± 11.7
Alcohol intake, glasses/week (mean ± SD)		9.8 ± 10.3	9.8 ± 10.3
Hypertension (%)		*21.6*	*19.7*
$L_{Aeq, 8h}$ (dBA)	80-84 (%)	*0.6*	-
	85-89 (%)	*29.0*	-
	90-94 (%)	*68.7*	-
	> 95 (%)	*1.7*	-

** difference between groups is significant at 0.01 level.*

Figure 2.1. shows that hearing thresholds strongly depend on age. Therefore, measured HTLs are corrected for age effects. After these corrections, the differences between the noise-exposed workers and controls remain statistically significant for all frequencies ($p < 0.001$). These differences are relatively small at 0.5 and 1 kHz (<1 dB) but become more pronounced at higher frequencies, with a maximum mean difference of 7.0 dB at 4 kHz.

Relationship of noise and hearing loss

In order to assess the relationship between hearing loss and noise exposure, multivariate regression analyses are performed, with age as covariate. Both noise parameters and the interaction term show a significant bivariate association with the

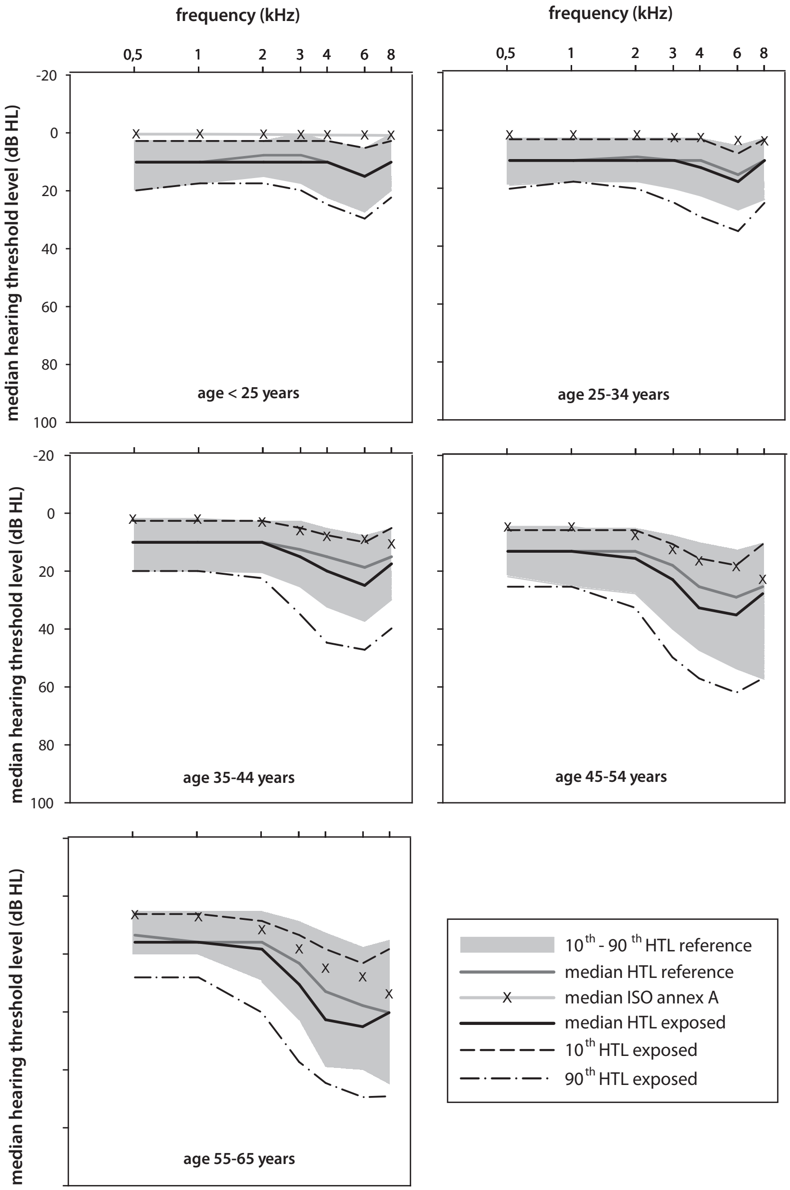


Figure 2.1. Measured hearing thresholds levels of the exposed workers (thick black lines), compared to the non-exposed internal controls (grey area and line) and age-matched ISO-1999 predictions of annex A (crosses), for five age groups.

PTA-values. However, the interaction term does not contribute significantly to both multivariate regression models and is excluded from further analyses. For $PTA_{1,2,4}$, the model accounts for 24.3% of the variance. The age-adjusted regression coefficient for noise level is 0.14 (99% CI 0.11-0.19), for years of exposure this is 0.07 (99% CI 0.05-0.09). The regression model for $PTA_{3,4,6}$ accounts for 32.4% of the variance. Also the age-adjusted regression coefficients for noise level and exposure time are higher for $PTA_{3,4,6}$, 0.27 (99% CI 0.22-0.32) and 0.12 (99% CI 0.09-0.15) respectively.

To gain more insight into the relationship between hearing loss and noise exposure, the impact of both parameters on hearing loss is further explored in separate analyses. The age-corrected hearing thresholds enable comparison to the noise-induced permanent threshold shift (NIPTS) described in ISO-1999. These NIPTS values are functions of audiometric frequency, exposure level and exposure time. For each individual construction worker, his expected median NIPTS is computed. $PTA_{3,4,6}$ is most affected by noise, and this age-corrected pure-tone average is examined as function of exposure duration. For exposure times between 10 and 40 years, the median value of expected NIPTS and its distribution can be calculated. For exposure times shorter than ten years, median expected NIPTS values are interpolated from the value of NIPTS for ten years, according to ISO-1999 (Figure 2.2).

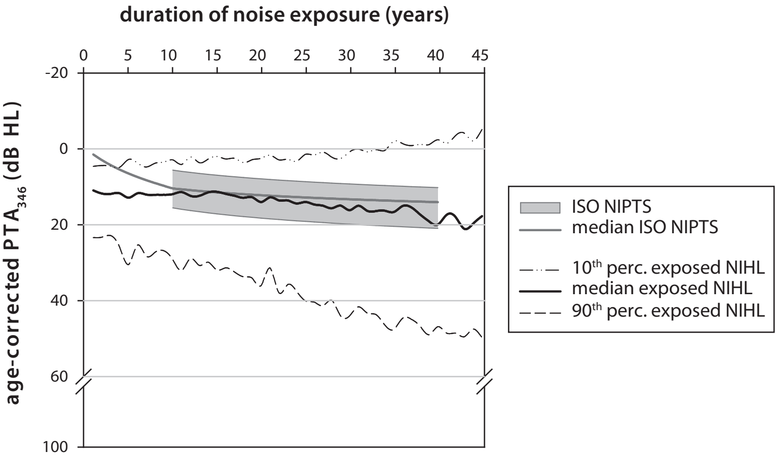


Figure 2.2. Median, 10th and 90th percentile age-corrected $PTA_{3,4,6}$ values of exposed population (black lines) and NIPTS distribution calculated using ISO-1999 (gray area) as a function of exposure time.

Although the inter-individual variation in the age-corrected hearing thresholds is larger in the exposed construction workers than predicted by ISO-1999, the median values of both groups follow a similar pattern for exposure times ranging from 10 to 40 years. However, this is not the case in the first 10 years of exposure. Where median values of ISO-1999 are interpolated to a NIPTS of 0 dB HL at the start of noise exposure, the population of noise-exposed construction workers shows age-corrected $PTA_{3,4,6}$ values that are approximately 10 dB HL higher at the beginning of occupational noise exposure without the steep increase as is predicted by ISO-1999.

Similarly, age-corrected $PTA_{3,4,6}$ values as function of daily noise exposure level are examined (Figure 2.3). The non-exposed control group accounted for the starting point at 80 dBA. There are large differences between the distributions of age-corrected hearing thresholds of the exposed study group and the ISO-1999 reference population. Hearing loss variation is, again, much greater in exposed employees, and their $PTA_{3,4,6}$ values are almost evenly distributed over the range of noise intensities. Hearing loss increases only slightly with increasing noise exposure level in this population, resulting in an almost flat curve that deviates strongly from the NIPTS predicted by ISO-1999. Up to exposure levels of 91 dBA, construction workers exhibit a greater hearing loss than predicted, while at higher noise levels less hearing loss is observed.

Other variables of influence

Data collection during periodic occupational health examinations also provides information about various factors possibly associated with NIHL, such as, the use of hearing protection, smoking and hypertension. To investigate the association between these risk factors and hearing loss, bivariate and multivariate regression analyses are performed. These analyses focus on $PTA_{3,4,6}$ only and are adjusted for the confounding effect of age. Results are displayed for the overall population and for both HPD subgroups separately in Table 2.3.

Age, noise intensity and exposure time have shown to be significant contributors to the regression model. The addition of other potential risk factors improves the model fit statistic from 32.6 to 42.0%. For the overall population, the additional variables that remain significant in the multivariate model include the use of hearing protection, a change in job history, noise nuisance at work and the presence of hearing complaints. The use of hearing protection shows a positive association with $PTA_{3,4,6}$ values, meaning that employees using hearing protection exhibit slightly more hearing loss than participants never using HPDs. Always being employed in the current job is associated with significantly greater hearing loss, and there is a strong association between the subjective complaints about poor hearing and the degree of hearing loss.

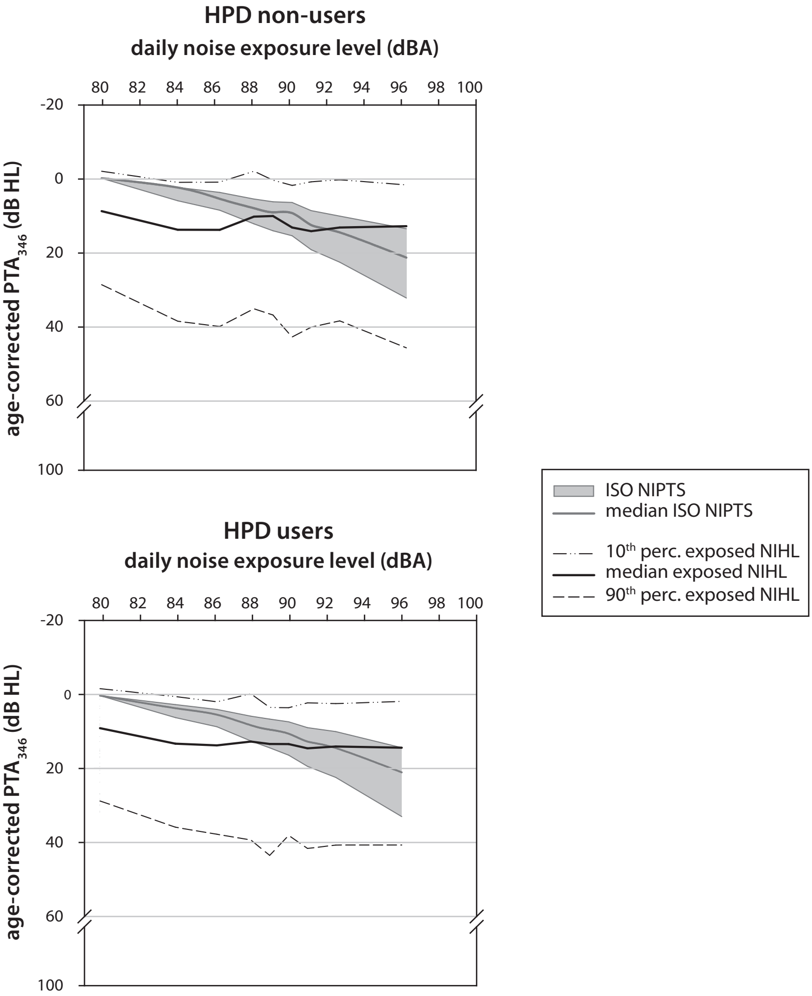


Figure 2.3. Median, 10th and 90th percentile age-corrected $PTA_{3,4,6}$-values of exposed population (black lines) and NIPTS distribution calculated using ISO-1999 (gray area), as a function of daily noise exposure level. Above: NIHL in HPD non-users. Below: NIHL in HPD users.

Hearing protection

Only 77% of the employees exposed to daily noise levels exceeding 80 dBA report to wear hearing protection devices at work, meaning that 23% of the exposed workers state to never use protection. Regression analyses show that employees using HPDs have an average increase in $PTA_{3,4,6}$ of 1.4 dB with regard to employees never using protection, after adjusting for relevant covariates. To gain more insight into the differences between participants using hearing protectors and participants not using protection, both groups are analysed separately. These analyses show that HPD users are employed in construction for a slightly shorter period (24.0 vs. 25.4 years) and are significantly younger than non-users (43.7 and 46.1 years, respectively). The percentage of HPD users declines with increasing age from 83.2% in employees younger than 25 years to 68.5% of the workers 55 years or older. Of the HPD users 44.8% indicated to be bothered by noise in their jobs, which is twice as much as the 21.6% in the non-user group. More importantly, the intensity of noise exposure differs significantly between HPD users and HPD non-users (90.6 and 89.5 dBA, respectively).

Stratified regression analyses for both subgroups of HPD users and HPD non-users did not show any differences between the results of the subgroups and of the overall population, except for the insignificant contribution of job history to the model for the non-users (Table 2.3). However, the regression coefficient found for noise intensity in the non-user group was slightly higher than in the user group. Nevertheless, Figure 2.3 does not show a stronger relationship of noise exposure level with age-corrected $PTA_{3,4,6}$ values in the non-user group compared to HPD users.

When dividing the noise exposure levels into high noise intensities (> 90 dBA) and moderate noise levels (between 80 and 90 dBA), it is shown that 84.4% of the highly exposed workers report to use HPDs versus 53.6% of the employees exposed to moderate noise levels. A stratified regression analysis for these two groups showed that HPD use only showed significant association with $PTA_{3,4,6}$ in workers exposed to noise levels between 80 and 90 dBA (data not shown).

Discussion

The results of this study confirm the adverse effect of noise exposure on hearing threshold levels; the construction workers exposed to noise have poorer hearing thresholds compared to their non-exposed colleagues and to an international reference population, especially in the 3-6 kHz region.

Audiometric results

This study shows a maximum mean deviation of 16.5 dB at 6 kHz from the ISO-1999 reference population. Compared to the internal control group, the greatest average difference is 7.0 dB, at 4 kHz. Although these differences are not as large as expected, the findings are in agreement with a study of Suter (2002). That study reports hearing threshold levels of carpenters and equipment operators that were approximately 5 dB worse than the HTLs reported in annex B of ISO-1999 in the high frequency region.

Table 2.3. Bivariate and multivariate predictors of hearing loss.

	Total				HPD non-users				HPD users			
	Bivariate		Multivariate ($R^2 = 0.42$)		Bivariate		Multivariate ($R^2 = 0.41$)		Bivariate		Multivariate ($R^2 = 0.43$)	
	B	99% CI	B	99% CI	B	99% CI	B	99% CI	B	99% CI	B	99% CI
age	0.80	0.79 - 0.81	0.61	0.58 - 0.64	0.76	0.72 - 0.79	0.64	0.61 - 0.67	0.82	0.80 - 0.84	0.59	0.55 - 0.63
noise intensity	0.31	0.26 - 0.36	0.18	0.13 - 0.23	0.24	0.18 - 0.29	0.19	0.13 - 0.24	0.30	0.25 - 0.35	0.20	0.15 - 0.25
years of exposure	0.16	0.13 - 0.19	0.09	0.06 - 0.12	0.12	0.07 - 0.17	0.05	-0.01 - 0.12	0.20	0.16 - 0.23	0.12	0.09 - 0.16
use of HPD	2.92	2.43 - 3.41	1.44	0.95 - 1.95	-				-			
no job change	0.30	-0.14 - 0.74	0.72	0.30 - 1.14	-0.89	-1.70 - -0.03	0.37	-0.45 - 1.18	0.18	-0.33 - 0.69	0.79	0.31 - 1.27
hearing complaints	12.80	12.33 - 13.27	12.38	11.98 - 12.91	13.16	12.19 - 14.13	12.76	11.79 - 13.73	12.54	11.96 - 13.12	12.20	11.61 - 12.79
bothered by noise	2.97	2.52 - 3.42	0.60	0.16 - 1.04	3.91	2.89 - 4.94	1.26	0.283 - 2.23	2.55	2.05 - 3.06	0.51	0.03 - 0.99
smoking - never	**reference**				**reference**				**reference**			
smoking - current	0.04	-0.49 - 0.57			-0.44	-1.42 - 0.55			0,18	-0.43 - 0.78		
smoking - ex	0.05	-0.48 - 0.58			-0.37	-1.36 - 0.63			0,17	-0.44 - 0.78		
cigarettes/day	-0.005	-0.04 - 0.03			0.000	-0.05 - 0.05			-0.01	-0.04 - 0.02		
years smoked	0.000	-0.03 - 0.03			0.03	-0.02 - 0.07			-0.01	-0.04 - 0.01		
alcohol intake	-0.001	-0.02 - 0.01			-0.01	-0.05 - 0.03			0.002	-0.25 - 0.26		
hypertension	0.11	-0.43 - 0.65			0.13	-0.85 - 1.12			0.21	-0.40 - 0.81		

Bivariate predictors are age-adjusted. Dependent variable is binaural $PTA_{3,4,6}$*. Variables are included in the multivariate regression analysis if bivariate regression shows a significant predictor at the 0.10 level.*

The unscreened reference population of annex B reports HTLs that are comparable to the high frequency thresholds measured in our internal control group. Nevertheless, the small group effects do not rule out significant threshold shifts in the ears of individuals that are more susceptible to noise-induced hearing loss than on average.

Study limitations

Although the main strength of this study was the size of the study population showing only a small percentage of missing values, some limitations in test administration and data collection cannot be avoided.

When comparing hearing threshold levels of construction workers to ISO-1999 standard values, both noise-exposed workers and controls show a deviation of about 10 dB HL at the lower frequencies. This deviation is reported in other studies as well, either in control groups used to analyse hearing ability of construction employees (Hessel, 2000; Hong, 2005) or in a general occupational population (Dobie, 2007). In this study, some aspects of test administration may have been responsible for this difference.

The available audiometric data are retrieved from screening assessments, omitting measurements of bone conduction. Therefore, we cannot correct for the possible presence of conductive hearing losses (e.g. due to permanent middle ear problems or temporarily conductive losses caused by a cold) that may be responsible for the elevated thresholds at the lower frequencies. Moreover, audiometric measurements are carried out on location, if possible in a mobile unit equipped with a soundproof booth. Nevertheless, possible exposure to background noise during the hearing test, which could produce elevated thresholds at 0.5 kHz, and to a lesser extent at 1 kHz (Suter, 2002), cannot be ruled out completely.

Furthermore, in this study no fixed noise-free period prior to audiometric measurements is defined. However, minimal time between possible occupational noise exposure and hearing tests was 2-3 hours. Guidelines in literature recommend a longer noise-free period, varying from 6 to 14 hours (NCvB, 1999; May, 2000). Consequently, the noise-free period of 2-3 hours may not be sufficient to fully recover from a possible temporary threshold shift (TTS) (Melnick, 1991; Strasser et al, 2003), and a complete absence of TTS cannot be guaranteed.

Moreover, collecting the appropriate data for noise exposure in this large population appears to be another limitation in this study. This study lacks individually measured noise exposure levels. Because construction workers are highly mobile and perform several different tasks, it is extremely difficult to obtain accurate estimates of the individual noise exposure levels.

Noise exposure estimations

Although regression analyses confirm a significant relationship between noise intensity and PTA-values, the hearing thresholds increase only marginal with increasing noise exposure level. This relationship follows a much flatter curve than predicted by ISO-1999.

A previous examination of Dutch industry workers compared single frequency threshold levels to ISO-1999 predictions (Passchier-Vermeer, 1986) and obtained a similar pattern, suggesting that ISO-1999 underestimates hearing loss at lower exposure levels and overestimates hearing loss at higher noise levels. In a more recent study, the shift between baseline and follow-up audiograms showed good agreement with model predictions (ANSI 3.44, 1996) at lower noise exposure levels, while at higher noise intensities less hearing loss than predicted was observed (Rabinowitz et al, 2007).

In the current study, individual noise exposure intensities are assigned based on job titles. This may have been too simplistic. It does not take into account that exposure may vary extensively between workers and over time. The diversity in specific tasks and the variety of equipment used at different workplaces introduces uncertainty in the calculations of noise exposure (Passchier-Vermeer, 1986; Rabinowitz et al, 2007). As a consequence, the resulting estimates are inaccurate in obtaining a reliable dose-effect relationship. Although the majority of the noise level estimates used in this study are mainly based upon carefully conducted sound level measurements and/or on personal dosimetry, noise levels are determined during a limited period of time. Therefore, the noise estimations are only samples and this limited sampling in complex and variable job situations, may have resulted in less accurate estimations.

Finally, the present noise exposure levels are also used as estimations of past exposure. Noise exposure levels of the construction workers may have varied considerably over their career. Regression analyses show only a small effect of prior employment on hearing, but the changes within jobs overtime may have limited the validity of the noise intensity estimations.

All these uncertainties in noise level estimations may have obscured a clear dose-effect relationship for the individual construction worker. However, for groups of workers with a sufficient number of employees, we may assume that most of the uncertainties mentioned above, e.g. the day-to-day variability and variations between individual workers, will be averaged out. Although the relations found in such an approach may be prone to some bias, we did not expect to find such a weak dose-effect relationship.

Attenuation of noise exposure from the use of hearing protection might partly explain the lack of the typical dose-response effect between noise level and hearing loss as well (Rabinowitz et al, 2007). The use of HPDs can cause inaccuracy in individual noise exposure estimation. This may have resulted in an overestimation of hearing loss for HPD users at noise intensities exceeding 90 dBA, at which a higher percentage of usage is reported. For this reason, stratified analysis for subgroups of HPD users are performed. The interpretation of the results of the HPD users is difficult because data on the effectiveness of hearing protection and the consistency of wearing are unknown. But also for the non-users the results do not show the expected relationship of noise intensity and hearing loss (Figure 2.3).

Apparently, the variability between individual workers combined with confounding factors such as the use of hearing protection, differences in past exposure, slight TTS-effects, and the inaccuracy of the noise exposure estimations prevent us from making accurate predictions of the effects of noise intensity on hearing, even in a population of this large size.

Effects of hearing protection

Hearing protection may have its greatest effect at high ambient noise levels. Workers exposed to higher noise intensities are obliged to wear hearing protection and are more bothered by ambient noise, making them more consistent in wearing their protection (Rabinowitz et al, 2007). In lower ambient noise levels HPDs may interfere with communication, jeopardizing the consistency of usage (Suter, 2002). Current analysis shows that 84.4% of the employees exposed to noise levels exceeding 90 dBA indicated to use HPDs versus 53.6% of the employees exposed to noise levels between 80 and 90 dBA.

Regression analysis shows a positive association of hearing loss and HPD use; employees using HPDs had on average 1.4 dB higher $PTA_{3,4,6}$ values than non-users. Bauer et al. (1991) also found a positive association between of the usage of HPDs and hearing loss by analysing a very large population of workers exposed to occupational noise. This can be explained by the suggestion that workers with beginning hearing problems are better motivated to use HPDs more consistently than their colleagues without hearing problems. When workers are divided into highly exposed employees (> 90 dBA) and employees exposed to moderate noise levels (80-90 dBA), HPD usage only shows a significant association with hearing in the moderately exposed group (data not shown). HPD use does not contribute significantly to the multivariate regression model for $PTA_{3,4,6}$ in the highly exposed group, despite the assumption that these are more consistent users.

In this study, HPD usage was scored as a binary variable, while the actual consistency of usage would be a more suitable predictor. The individual fitting of HPDs, the consistency of HPD usage and exposure level during use and non-use are crucial elements in determining the actual noise dose (Seixas et al, 2005). In addition, HPD data are based on employees' self-report, which can be subject to reporting bias and social desirability (Griffin et al, 2009). These uncertainties can lead to misclassification, thereby overestimating HPD usage and underestimating the true effect of hearing protection (Davies et al, 2008). Unfortunately, data about the effectiveness of the HPDs and about the consistency of usage were unavailable.

Effects of noise exposure time

The relationship of hearing loss and exposure time, defined as years of employment in construction, is also explored. Exposure time is positively related to hearing threshold levels; longer exposure times are associated with higher $PTA_{3,4,6}$ values. This effect was about 0.09 dB loss in $PTA_{3,4,6}$ for each year of exposure, after adjustment for age, noise intensity, and other risk factors. This increase is similar as reported in ISO-1999, which predicts an average increase in median $PTA_{3,4,6}$ values of 1 dB/decade for exposure levels of 90 dBA. Also a review by Rösler (1994) reports the same amount of increase in age-corrected HTLs at 4 kHz, after the first 10 years of exposure.

When comparing the age-corrected $PTA_{3,4,6}$ values of the study population and the ISO-1999 predicted NIPTS as a function of exposure time, the greater inter-individual variation in the distribution of NIHL in exposed construction workers is remarkable. This suggests a high variation in factors influencing the susceptibility to hearing loss in each exposure year interval of the study group, such as HPD use, prior employment, non-occupational noise exposure, hearing disorders, and variability in noise intensity. However, the median values of both the noise-exposed workers and the ISO-1999 predictions have a similar slope, at least for exposure times between 10 and 40 years.

An interesting aspect is the relationship during the first 10 years of noise exposure. Construction workers employed for less than 10 years, show greater hearing losses than expected based on the interpolation of ISO-1999. In addition, observed hearing loss increases over the first 10 years of exposure at the same rate as in the following 10-40 years of exposure duration, where a pattern of strongly increasing thresholds would have been expected (ISO, 1990; Rösler, 1994; Prince, 2002). To investigate the role of job history in this group with short exposure duration, this relationship is determined only for construction workers younger than 30 years of age that reported no prior employment. This selection of 2,190 employees shows a similar pattern of median age-corrected $PTA_{3,4,6}$ values that is about 10 dB HL higher than predicted by ISO-1999.

A number of previous studies also found a discrepancy between ISO-1999 predictions and measured hearing loss during the first years of exposure. Analyses based on serial audiograms of railway workers showed that hearing thresholds exceed model predictions in the very beginning of noise exposure, showing age-corrected hearing loss at job entrance of 9 dB averaged over 2 and 4 kHz (Henderson & Saunders, 1998). Another study, monitoring a cohort of newly enrolled construction apprentices showed HTLs of 12.2 dB HL at 4 kHz at baseline (Seixas et al, 2004) without any change in audiometric hearing thresholds over the first 3 years of employment (Seixas et al, 2005). The reported hearing threshold levels at job entrance in these studies are all higher than 0 dB HL and correspond to the median age-corrected $PTA_{3,4,6}$ of 10.9 dB HL found here.

The ISO-1999 model depends on the interpolation of predicted hearing thresholds after 10 years of exposure and the assumed hearing thresholds of 0 dB HL at the beginning of employment. Our findings suggest that this may not correctly represent the true development of NIHL over this period of exposure. The interpolation of the ISO-1999 formula could either be less applicable to the population of interest, or the starting point of 0 dB HL is set too low, possibly due to the fact that the amount of early hearing damage in this population is underestimated.

NIHL in young employees

A Dutch survey of health-related and occupational problems among construction workers shows that 7.6% of construction workers younger than 25 years are diagnosed with NIHL (Arbouw, 2009). Reported prevalence of hearing loss among young adults entering the construction industry in literature is even higher, ranging from 14.4 to 16% (Seixas et al, 2005; Rabinowitz et al, 2006). This suggests that the starting point of 0 dB defined in ISO-1999 is set too low in this population, because NIHL is already present in workers even before employment. Possibly, this is caused by noise exposure in recreational settings, underlining that non-occupational noise is another complicating factor in the relationship of occupational noise exposure and hearing impairment. Neitzel et al. (2004) demonstrated that approximately one-third of apprentices in the construction industry experience equivalent noise levels higher than 80 dBA from recreational noise exposure, placing them at risk for NIHL even before considering occupational exposure. Effects of both occupational and non-occupational noise exposure will accumulate and exposure to non-occupational noise prevents workers to recover from occupational noise exposure. Since the current study was conducted during audiometric screening in an occupational health setting, no information concerning exposure to leisure noise is available. Information about non-occupational noise exposure and a baseline audiometric measurement would be highly advisable for medico-legal purposes.

Effects of confounding factors

The influence of other possible confounding factors must be considered when interpreting the presented relationships between hearing loss and noise exposure. Despite confounding factors such as job history and use of hearing protection, the multiple linear regression analysis still show a significant contribution of noise exposure to the regression model. Lifestyle factors, such as smoking, alcohol intake and hypertension, do not show a relationship with NIHL in this population. The multivariate model for $PTA_{3,4,6}$ only explains 41.1% of the variance in hearing threshold levels; hence, most of the variation is not explained by variables measured in this study. Other studies performing multiple regression analyses to examine the effect of noise exposure and hearing ability adjusted for several confounders, found smaller R^2 for their multivariate models of 30.6% (Agrawal et al, 2010) and 36% (Toppila et al, 2000).

Differences in the individual susceptibility to noise may be responsible for the large spread of individual threshold values. Several possible explanatory variables that hypothetically could be responsible for part of the variance, such as medical history, non-occupational noise exposure and drug usage, are not included in these analyses, because of a lack of information concerning these factors.

Conclusion

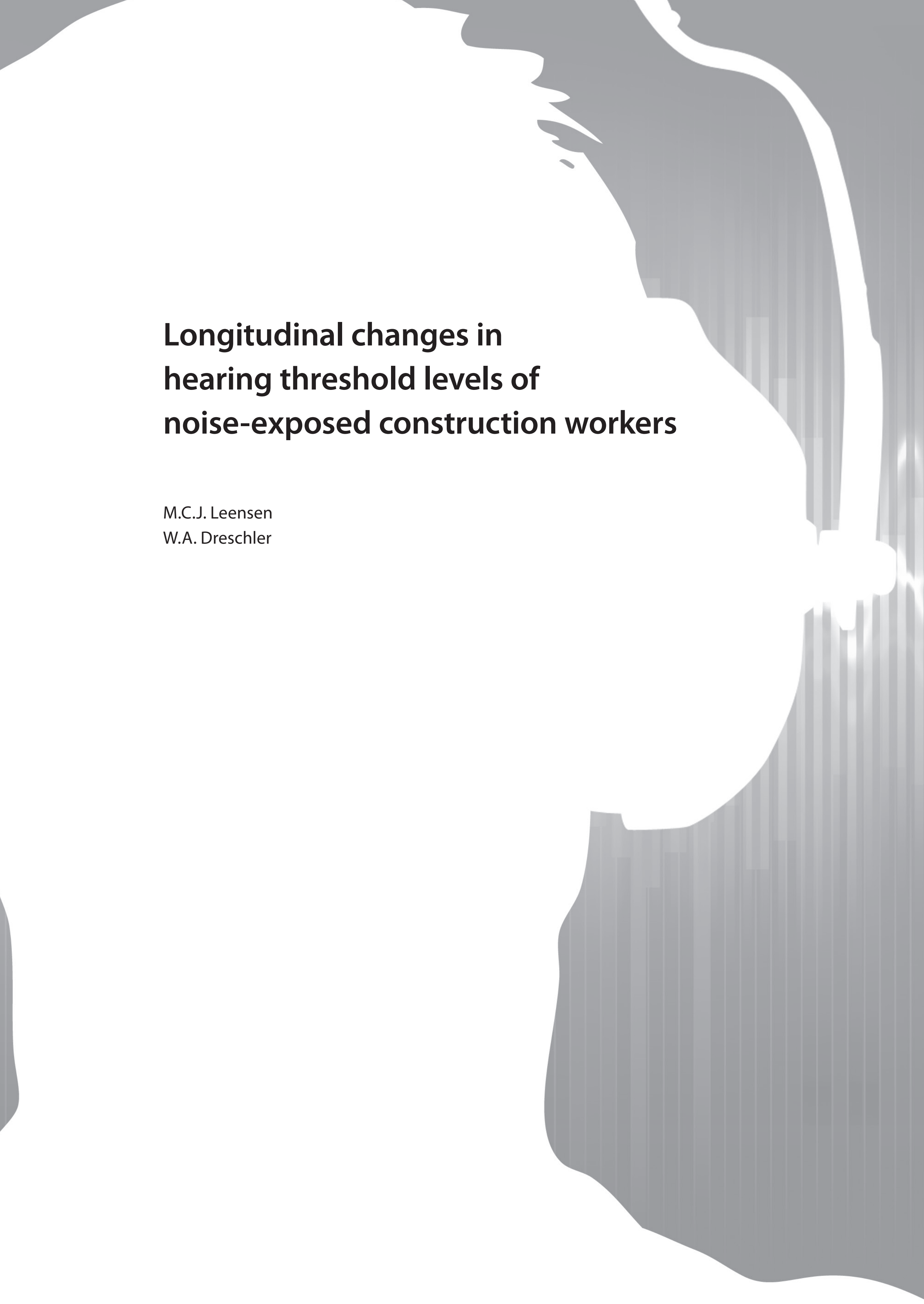
This analysis of a large audiometric dataset show that Dutch construction workers exhibit greater hearing losses than expected based solely on ageing. Accumulation of the inevitable age-related hearing loss may result in moderate to severe hearing impairment at retirement age.

Regression models show a great inter-individual variability in reported hearing loss, and only a weak relationship between noise level and hearing ability is found. At low noise exposure levels, hearing loss is much greater than predicted whereas at high levels hearing loss is less. This latter might be partly explained by the role of personal hearing protection, which is worn by a greater proportion of highly exposed workers than workers exposed to lower noise levels. Individual noise exposure level measurements can increase the accuracy of the noise intensity estimates and results in a more reliable estimate of this relationship.

Growth of hearing loss with progressing exposure time is in accordance with ISO-1999 predictions for exposure durations between 10 and 40 years. However, the interpolation described in the ISO-1999 model that predicts hearing loss developed during the first 10 years of exposure is not consistent with our data and seems to be inapplicable in this population. Our hypothesis is that pre-existing hearing loss from non-occupational noise exposure is the most important explanation for this inconsistency.

In a follow-up study, personal dosimetry and extensive information on job history should be taken into account estimating noise exposure levels. In addition, serial audiometry with a baseline measurement at job entrance should be performed and more detailed information should be collected about factors influencing hearing ability, such as, non-occupational noise exposure, medical history and details of hearing protector usage.

3



Longitudinal changes in hearing threshold levels of noise-exposed construction workers

M.C.J. Leensen
W.A. Dreschler

Abstract

Purpose: Longitudinal analysis of audiometric data of a population of noise-exposed workers provides insight in the development of noise-induced hearing loss (NIHL) over a period of 4 years, as a function of noise exposure, and age.

Methods: Over a period of approximately 4 years after the measurements reported in Chapter 2, 17,930 construction workers of this baseline cohort had one or more follow-up assessments. Their pure-tone audiometric thresholds obtained during these periodic occupational health examinations were available for analysis. Linear mixed models were fitted to explore the relationship between the annual rate of change in hearing and noise intensity, exposure duration, and age. The audiometric data of a subset of 3,111 workers who were tested on three occasions, were used to investigate the pattern of hearing loss development.

Results: The mean annual rate of change in this study population was about 0.56 dB/yrand this became larger with increasing noise intensity and increasing age. The duration of noise exposure did not affect the annual shift in hearing loss. During the first decade of noise exposure, mean rate of change again deviated from ISO-1999 predictions, in that hearing thresholds improved. The change in hearing over of three measurements showed a concave development of hearing loss as a function of time, corresponding to NIHL development.

Discussion: The deviation from ISO-1999 predictions observed in Chapter 2 is probably the result of the higher average normal-hearing levels in survey data. Because hearing threshold levels obtained at follow-up were better than those obtained at baseline, no statement can be made about the NIHL development during the first decade of exposure. This improvement in hearing threshold levels is likely the results of measurement variation in occupational screening audiometry, rather than an actual improvement in hearing ability.

Introduction

Noise is one of the most prevalent occupational hazards. Despite the widespread recognition of the impact of noise on hearing, occupational noise exposure remains a significant problem, especially in the construction industry (Suter, 2002), where the majority of the workforce is exposed to daily noise levels exceeding 80 dBA (Neitzel et al, 2011). As a result, noise -induced hearing loss (NIHL) is the most commonly reported occupational disease in the Netherlands (Van der Molen & Lenderink, 2012). Averaged over the past five years, 39% of occupational disease reports concerned NIHL, the majority of which derived from the construction industry (Van der Molen & Lenderink, 2012).

Indeed, the cross-sectional data analysis reported in Chapter 2 showed that noise-exposed construction workers had greater hearing losses compared to the reference population reported in ISO-1999 annex A (1990), as well as to their non-noise-exposed colleagues.

The ISO-1999 standard combines data from numerous cross-sectional studies into a widely used model to predict hearing loss for a noise-exposed population. This model assumes that a subject's hearing threshold level (HTL) is composed of two additive elements: an age-related component estimating the age-related hearing loss (ARHL) in annex A, and an estimation of the noise-induced permanent threshold shift (NIPTS) resulting from on-the-job noise exposure.

However, the relationships of hearing threshold and noise exposure found in Chapter 2 deviates from the relationship described in ISO-1999 in two important aspects. First, there was only a weak relationship between noise intensity and hearing threshold levels. When the daily noise exposure level rose from 80 dBA towards 96 dBA only a minor increase in hearing loss was shown (0.18 dB increase per dB increase in noise level). The duration of noise exposure seemed a better predictor than noise exposure level, probably because of limited accuracy of noise exposure estimates (Seixas et al, 2004; Seixas et al, 2012) and the confounding effect of hearing protection usage (Rabinowitz et al, 2007).

Second, despite the stronger relationship of hearing loss and exposure time, it only corresponded to ISO-1999 predictions for durations between 10 and 40 years, whereas the observed thresholds in the first decade of exposure were higher than predicted by ISO. ISO-1999 presents an algorithm, to calculate NIPTS for exposure durations between 10 and 40 years. However, previous research showed that most NIHL arises during the first 10 to 15 years of noise exposure (Tayloret al, 1965; Rösler 1994; Prince et al, 2002). ISO-1999 designs this steep increase in HTL by an extrapolation of 0 dB NIPTS at the start of noise exposure to the NIPTS predicted after 10 years. Instead of this progressive increase in hearing loss during the first decade of exposure, the retrospective analysis in Chapter 2 showed an increase in HTLs with increasing exposure duration, which was similar to the relationship found for longer exposure durations. Hearing loss was higher than ISO-1999 predicted, and more importantly, workers employed for less than one year showed average age-corrected hearing losses of about 10 dB HL. Other studies focusing on the first effects of occupational noise exposure also showed deviations from the ISO-1999 interpolation, in that their observed elevation of baseline hearing thresholds was similar to the findings in Chapter 2 (Henderson & Saunders, 1998; Seixas et al, 2004). This poses the question whether the interpolation proposed by ISO-1999 for the first decade of noise exposure is applicable to occupationally exposed employees, or whether employees may have some pre-existing hearing loss when entering the workforce, probably due to recreational noise exposure.

The data analysis described in Chapter 2 was based on cross-sectional data collection, and so estimations of hearing loss over time were done across subgroups of the total study population, rather than obtained individually. Considering the pattern of hearing loss development during the first decade of exposure, the preferred approach to study the development of early losses is using longitudinal studies, especially since this may follow a nonlinear history (Johnson, 1991). Longitudinal analyses combining baseline data with follow-up measurements of the same study population could give more insight into the development of hearing loss over time.

However, the ability to observe small threshold shifts over time requires that the measurement procedure is sufficiently sensitive to detect relatively small changes. Small threshold shifts for an individual employee cannot easily be distinguished from normal test-retest variability of standard pure tone audiometry (Royster & Royster, 1986), which is about 5 dB (Dobie, 1983; Hall & Lutman, 1999; Helleman & Dreschler, 2012). Nevertheless, on group level pure-tone audiometry can establish overall hearing trends for noise-exposed employees by longitudinal analysis of a large audiometric database (Royster & Royster, 1986). Analysing repeated measurements over time within the same individuals may reduce variability in threshold determinations and by averaging over large groups, small changes can be identified.

Of the original group of 29,644 construction workers that were studied by cross-sectional analysis in Chapter 2, 17,930 performed one or more follow-up audiograms in the 4-year period following baseline assessment. Industrial data provide an immediate source of practical knowledge concerning effects of noise on hearing, and longitudinal analysis of hearing threshold levels over time can provide insight in development of NIHL.

Aim of this study was to describe the change in hearing threshold levels of noise-exposed workers as measured during regular periodic audiometric screening over time, and to estimate the typical rate of change in hearing sensitivity per year. The relationship of this rate of change in hearing with both occupational noise exposure and age is examined, and compared to ISO-1999 predictions. Particular interest is in workers exposed for less than 10 years in order to establish the amount of hearing loss growth during the first decade of exposure. In addition, the association of demographic or work-related variables with the development of NIHL is studied, because this may identify specific risk groups. Finally, the audiometric data of the employees having more than one follow-up measurement are investigated, to analyse the course of hearing loss development.

Methods

The study of longitudinal changes in hearing threshold levels was based on data collected by Arbouw, the Dutch national institute on occupational health and safety in the construction industry. These data were extracted from medical records of periodic occupational health examinations (POHE) that were performed as part of regular occupational healthcare. This POHE consisted of an extensive self-administered questionnaire and a physical examination, including standardized audiometric testing. POHEs are offered to all construction employees, irrespective of occupational exposure to noise. Every employee is invited to participate once every four (age < 40) or two (age ≥ 40) years. Participation is completely voluntary.

Data collection

The starting point for the data collection in this study was the dataset used for the cross-sectional data analysis described in Chapter 2. That study population, referred to as the baseline cohort, consisted of 29,216 employees examined between 1 November 2005 and 20 July 2006. All additional records of follow-up POHEs of this baseline population performed until July 2010, as well as data from their baseline records, constituted the current dataset of investigation. In total, 22,575 follow-up records were available for analyses.

Of the these records, 4,645 were from the same individuals who had two follow-up examinations during the measurement period. The three measurements available for this subset were kept in a separate dataset, in order to investigate the pattern of hearing loss development over three measurement occasions. For the main analyses, the baseline data and the most recent measurement of this subset were kept, and thus the final dataset consisted of two measurements of 17,930 unique subjects.

Audiometric measurement

The core of the data collection was formed by the hearing threshold levels as obtained by regular screening audiometry. Pure-tone audiometry was assessed in accordance with ISO-8253.1 (2010). Audiometers were annually calibrated according to ISO-389.1 (1998). POHEs were usually conducted during the work shift and at workplaces. If this was possible a mobile unit equipped with a soundproof booth was used. Pure tone air-conduction thresholds were determined at frequencies 0.5, 1, 2, 3, 4, 6, and 8 kHz in both ears, in 5-dB steps ranging from -15 dB HL to a maximum of 90 dB HL. A HTL of 90 dB HL was the upper limit of the test equipment and a hearing threshold level was marked as 95 dB HL if the participant did not respond to this maximum sound signal. Because of this ceiling effect, only hearing threshold recordings of 90 dB HL and lower were preserved in this analysis.

Questionnaire

Prior to the physical examination, the POHE participants completed an extensive self-administered questionnaire. Relevant demographic, occupational, and health-related data were extracted from these questionnaires. This included information regarding job title, use of hearing protection devices (HPDs) (yes/no), the number of years employed in both the construction industry and the current occupation, presence of hearing complaints, and whether employees were troubled by noise at work. In addition, cigarette smoking status, alcohol intake and blood pressure were recorded. Hypertension was defined as systolic blood pressure ≥ 140 mmHg combined with diastolic blood pressure ≥ 90 mmHg (De Moraes Marchiori, 2006).

Noise exposure estimation

To estimate daily noise exposure, workers were classified by the time weighted average (TWA) noise exposure levels estimated for standardized job titles. These TWA exposure levels were extracted from a database of Arbouw (Arbouw, 1998) that reported data of measurements of TWA noise levels based on personal dosimetry sampling for several job titles. Exposure levels for remaining job titles were based on sound level measurements during specified activities and on group data recorded in previous POHEs. For more information on these noise exposure level estimations, see Chapter 2. All noise exposure levels were expressed as equivalent 8-h, A-weighted sound-pressure levels $L_{A, eq(8h)}$ calculated using an exchange rate of 3 dB. The reported years of employment in construction industry were used to estimate the duration of noise exposure. The correspondence between reported years employed in construction and years worked in the current job was used to determine whether an employee has had a change in job history. If the number of years employed in construction sector exceeded the number of years on current job, it was assumed that the former job has had equivalent exposure levels.

The workers that were employed in non-noise exposed jobs could function as a reference group. Since this study focuses on the change in hearing threshold levels over time, the 1,077 subjects that reported no occupational noise exposure during the intermediate measurement period were defined as the reference population.

Exclusion criteria

In some cases a medical record could not be used for analysis. Whole employee records were removed for the following reasons:

- Insufficient follow-up period; the interval between measurements should be at least 1 year. In total 475 employees that had their follow-up examination within 11 months after baseline were excluded from analysis.
- Incorrect data collection; 410 workers were omitted for having either demographic discrepancies or missing hearing threshold levels. Similar exclusion criteria as in the baseline cohort were applied (see Chapter 2).
- Lack of correspondence between successive datasets; after merging medical records of each individual, 2,623 cases showed discrepancies between repeated measurements of variables of noise exposure could not be used for current analysis.
- Audiometric discrepancies; 2,160 audiograms that did not correspond with signs of NIHL or presbyacusis, or demonstrated changes (either positive or negative) that showed major deviations from expected values were also excluded.

In the appendix, more details on the specific exclusion criteria are described.
As a result, 12,269 subjects were considered to have reliable data and were kept for further analysis.

Statistical analyses

The data were analysed using SPSS (version 19.0) and R software (R Foundation 2008, from http://www.R-project.org).

Linear mixed effect models can be used to fit longitudinal data in which the number and spacing of observations vary among participants. So these models were fitted to current data to assess the longitudinal changes in this study sample, in both ears and across all frequencies, while accounting for the effects of repeated measurements within each individual. Fixed effects in these models were fitted to estimate of the average intercept and the effect of different factors and covariates on hearing threshold levels or change in hearing loss. Random effects accounted for individual variation in individual thresholds, ear, and differences in thresholds among the frequencies. Additionally, these random effects accounted for the autocorrelation due to repeated measurements within each individual and allowed for unbalanced data due to missing values.

First step in data analysis was that the average hearing threshold levels of the total study population collected at two measurement occasions were examined using mixed effects modeling. For this general analysis, longitudinal change was represented by a fixed linear effect of 'time', and the audiometric configuration was represented by the term ‘frequency’. A term representing the tested ‘ear’ also was included. In addition, two-way interactions among these fixed factors were incorporated. After this first analysis, the change in HTLs, and the effect of different parameters on this change were examined, using the rate of change in dB/year as dependent variable. By dividing the difference in HTLs by time between baseline and follow-up measurements in years, effects of different interval periods were eliminated, reducing the amount of parameters in the model. To further reduce the number of parameters, analyses focused on the change in hearing loss in the pure-tone average of the noise-sensitive frequencies 3, 4, and 6 kHz ($PTA_{3,4,6}$). Again, these longitudinal analyses of changes in hearing were conducted using mixed effects modeling, with variation between subjects, ears within subject, and measured frequency treated as random effects. The predictors of primary interest were, besides frequency and ear, baseline age, noise intensity, noise exposure duration, and HPD usage. Adjustments were made for covariates thought to be correlated with either HTLs and occupational noise exposure; baseline hearing status ($PTA_{3,4,6}$), change in job history, duration of the follow-up interval, smoking status, hearing complaints, and noise disturbance at work.

Only factors and interaction terms that showed a significant contribution to the fitted model, tested with conditional F-tests at the 0,05 level, were investigated for significant coefficients at each level. When coefficients proved to be significant, the term was retained in the model. The results of the models are displayed as the estimated effects for the fixed factors and interaction terms retained in the model, and coefficients and corresponding 99% confidence levels are presented for each term. In case of an interaction between two variables, the difference between a certain condition and the reference is obtained by summing up the coefficients obtained for each term contributing to that interaction.

Results

Characteristics of study population

In total, hearing threshold levels of 12,269 male construction employees were collected. This population can be divided into a large group of noise-exposed employees (n = 11,192) and an internal reference group that was not exposed to noise during the measurement interval (n = 1,077). Mean age of both groups was similar ($p = 0.095$), but the distribution over age groups differed slightly ($p < 0.001$) (Table 3.1). Noise-exposed workers were on average 3.8 years longer employed in the construction industry ($p < 0.001$). Based on their job title, the majority of the exposed group, 75.5%, was estimated to work in average daily noise levels of 90 dBA or higher. Of the exposed employees, 76.9% reported to use hearing protection, 21.8% had complaints of worsened hearing, and 39.0% is bothered by noise during work. Baseline hearing of the reference group was slightly better than hearing levels observed in the exposed workers ($p < 0.001$) in both ears (Table 3.1).

Hearing threshold levels

The mean HTLs of the study population at both measurement occasions are shown in Figure 3.1, for both ears separately. Mean hearing levels are plotted against the HTLs of the reference group and the predicted hearing levels by ISO-1999 annex A, based on individual median age-related hearing loss calculated for each employee in the noise-exposed group.

Linear mixed effect models were run with random effects for 'subject' and 'ear', and fixed effects for 'ear' (left, right), 'frequency' (0.5 – 8 kHz), and 'measurement time' (baseline, follow-up). Two-way interactions between the fixed effects were also included in the model.

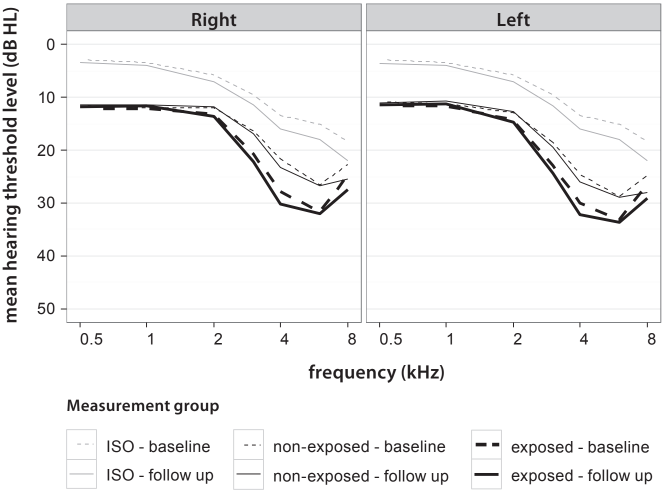


Figure 3.1. Mean HTLS of the noise-exposed workers, the non-exposed references and ISO-1999 predictions, for both baseline (dashed line) and follow-up (solid line) measurements.

Table 3.1. Demographic, work related and hearing loss factors, by subject group.

		Reference (n=1077)	Exposed (n=11.192)
Baseline age, yrs (mean ± SD)		45.4 (10.5)	44.9 (10.8)
Distribution of baseline age (%)	< 25	*4.1*	*9.6*
	25 – 34	*13.5*	*5.1*
	35 – 44	*28.2*	*26.3*
	45 – 54	*30.3*	*41.5*
	55 – 65	*24.0*	*17.5*
Years in construction (mean ± SD)		21.7 (12.4)	25.5 (12.0)
Estimated $L_{eq, 8h}$, (dBA) (%)	< 80	*100.0*	*0.0*
	80-89	-	*24.5*
	≥ 90	-	*75.5*
HPD usage (%)	yes	*7.0*	*76.9*
Job history (%)	Never changed	*22.2*	*47.7*
	Recently changed	*25.5*	*8.6*
	Changed	*52.3*	*43.7*
Baseline $PTA_{3,4,6}$ value (mean ± SD)	Left	24.0 (15.0)	28.6 (16.3)
	Right	21.6 (13.6)	26.8 (15.8)
Baseline hearing status $PTA_{3,4,6}$ > 20 dB HL (%)	Left	*50.0*	*60.9*
	Right	*40.9*	*55.9*
Do you experience hearing complaints? (%)	Yes	*16.4*	*21.8*
Are you bothered by noise at work? (%)	Yes	*4.5*	*39.0*
Smoking status (%)	Never	*44.0*	*33.5*
	Ex	*36.8*	*36.8*
	Current	*19.2*	*29.7*
Hypertension (%)	Yes	*20.1*	*20.9*
Interval period (mean ± SD)		3.2 (0.8)	3.0 (0.9)

Italic values represent percentages.

All fixed factors in the model showed significant effects, and the coefficients and corresponding 99% confidence intervals for all terms of the full model are displayed in Table 3.2. The main effect of ‘frequency’ (F[1,141468] = 10764.24, p < 0.001) indicated that hearing threshold levels differed over the frequencies at which they had been obtained. Mean baseline HTLs in the right ear ranged from 12.2 dB HL at 0.5 kHz to 31.3 dB HL at 6 kHz, and were highest in the higher frequency region. Also, 'frequency' showed a significant interaction with both 'measurement time' (F[6,141468] = 68.48, $p < 0.001$) and 'ear' (F[6,165054] = 377.75, $p < 0.001$) indicating that the mean HTL difference between both ears or measurements differed between the tested frequencies (Table 3.2).

Table 3.2. Coefficients and 99% confidence intervals for the different terms in the linear mixed model predicting hearing threshold levels of the study population.

Model terms	Coefficient	99% CI
HTL right at 0.5 kHz	12.18	11.83 – 12.54
HTL right at 1 kHz	12.19	11.83 – 12.55
HTL right at 2 kHz	13.30	12.94 – 13.66
HTL right at 3 kHz	20.45	20.09 – 20.81
HTL right at 4 kHz	27.46	27.11 – 27.82
HTL right at 6 kHz	31.32	30.96 – 31.68
HTL right at 8 kHz	24.73	24.37 – 25.08
Left * 0.5 kHz	-0.54	-0.89 – -0.19
Left * 1 kHz	-0.01	-0.51 – 0.48
Left * 2 kHz	1.49	1.00 – 1.99
Left * 3 kHz	2.64	2.14 – 3.13
Left * 4 kHz	2.58	2.08 – 3.08
Left * 6 kHz	2.00	1.51 – 2.50
Left * 8 kHz	2.12	1.62 – 2.62
Follow-up * 0.5 kHz	-0.28	-0.44 – -0.11
Follow-up * 1 kHz	-0.18	-0.41 – -0.05
Follow-up * 2 kHz	0.53	0.30 – 0.76
Follow-up * 3 kHz	1.66	1.42 – 1.89
Follow-up * 4 kHz	2.54	2.31 – 2.77
Follow-up * 6 kHz	0.71	0.48 – 0.95
Follow-up * 8 kHz	3.00	2.77 – 3.24

The coefficients reflect mean baseline HTLs of the right ear, the difference in baseline HTLs in the left ear relative to the right ear, and the difference in HTLs obtained in the follow-up measurement relative to baseline, for each frequency.

‘Measurement time’ showed a significant main effect (F[1,165054] = 19.09, $p < 0.001$), indicating that the thresholds obtained at follow-up showed poorer hearing at 2 to 8 kHz. However, at 0.5 and 1 kHz a small but significant improvement of hearing levels was observed (Table 3.2). These effects are also shown in Figure 3.1. The main effect of ‘ear’ (F[1,11323] = 15.47, $p < 0.001$) showed that across the frequencies measured, hearing sensitivity was slightly poorer in the left ear than in the right, with differences ranging from 1.49 to 2.64 dB HL.

The focus of this study is on change in hearing loss over time. Thus, in order to reduce the number of variables in the model, the successive analyses concerned the rate of change in dB per year, rather than absolute hearing threshold levels. A linear mixed effect model with random effects for ‘subject’ and ‘ear’, and fixed effects for ‘ear’ and ‘frequency’ showed that the rate of change differed across test frequencies (F[6,141474] = 357.55, $p < 0.001$); hearing was significantly worsened at higher frequencies, varying from 0.19 dB/year at 2 kHz to 0.99 dB/year at 8 kHz (Table 3.3). The negative coefficient obtained for a change in hearing at 0.5 kHz reflected an improvement in hearing of 0.08 dB/year at this frequency. There was no significant effect of ‘ear’ (F[1,11322] = 3.33, $p = 0.068$), indicating that, although baseline HTLs were different, the rate of change was similar in left and right ears.

Table 3.3. Coefficients and 99% confidence intervals for different terms in the linear mixed model predicting the annual rate of change in HTLs of the study population.

Model terms	Coefficient	99% CI
HTL change right at 0.5 kHz	-0.08	-0.16 – -0.01
HTL change right at 1 kHz	-0.05	-0.13 – 0.03
HTL change right at 2 kHz	0.19	0.11 – 0.27
HTL change right at 3 kHz	0.58	0.51 – 0.66
HTL change right at 4 kHz	0.87	0.79 – 0.95
HTL change right at 6 kHz	0.38	0.30 – 0.46
HTL change right at 8 kHz	0.99	0.91 – 1.07
HTL change left	0.03	-0.01 – 0.08

The coefficients reflect mean annual change in HTL for each frequency, and the overall difference in the left ear relative to the right. Since this term is not significant, mean annual changes displayed are similar in the left ear.

Relationship of noise exposure and rate of hearing loss

NIHL affects the high-frequency region, so the greatest change in HTLs of noise-exposed employees was expected in this region (Table 3.3). In order to investigate the effect of age, noise exposure, and covariates on hearing loss development, the rate of hearing loss is defined as the annual rate of change in the pure tone average of hearing threshold level at 3, 4 and 6 kHz ($PTA_{3,4,6}$). The mean rate of hearing loss observed for the total study population was 0.54 dB/year (SD = 3.04).

The relationship between annual rate of change in hearing and noise exposure, was investigated by fitting a linear mixed effect model for fixed effects of 'noise intensity' and 'exposure duration', and random effects for 'subject' and 'ear'. Since hearing thresholds deteriorate with increasing age, the effect of 'baseline age' was also investigated as a covariate in this model. All three parameters showed a significant bivariate relationship with $PTA_{3,4,6}$. The complete mixed model showed both a positive association of annual shift in hearing with noise intensity ($F[1,12253] = 11.51$, $p < 0.001$) and with baseline age ($F[1,12253] = 123.73$, $p < 0.001$). The main effect of 'exposure duration' did not significantly contribute to the model ($F[1,12253] = 0.004$, $p = 0.946$). This variable was highly correlated with baseline age, which already explained most of the variance associated with age and/or duration and change in hearing. There were no significant interaction terms between the three fixed factors. The coefficients of the model are shown in Table 3.4. A positive coefficient means a deterioration in hearing ability, a negative coefficient indicates an improvement in hearing thresholds. The intercept value of -1.08 indicated that an improvement in $PTA_{3,4,6}$ was observed for workers in the reference condition that the intercept represented. This reference condition concerned workers of 16 years old, exposed to daily noise levels not exceeding 80 dBA and employed in construction for less than 1 year at baseline. The positive coefficients for noise intensity and baselineage showed that the deterioration in $PTA_{3,4,6}$ became larger with increasing noise exposure level and increasing age; with every dB increase in intensity of the noise exposure above 80 dBA, the change in

Table 3.4. Coefficients and 99% confidence intervals for the different parameters of noise and age in the linear mixed model predicting the annual rate of change in $PTA_{3,4,6}$.

Model terms	Coefficient	99% CI
Intercept	-1.08	-1.338 – -0.822
Noise level	0.024	0.006 – 0.043
Exposure duration	0.000	-0.009 – 0.010
Age	0.048	0.037 – 0.059

$PTA_{3,4,6}$ increased with 0.024 dB/yr, and with every year increase in baseline age exceeding 16 years, the change in $PTA_{3,4,6}$ was 0.048 dB greater. Overall, these model coefficients meant that, for example, worker aged 45 years who was exposed to a daily noise level of 90 dBA, would show an average annual deterioration in $PTA_{3,4,6}$ of 0.55 dB/yr.

Comparison to the ISO-1999 model

TTo gain more insight into the relationship between noise exposure and hearing threshold changes, the impact of both parameters was investigated further. These analyses concerned the annual rate of change in $PTA_{3,4,6}$ as a function of either noise intensity or noise exposure duration. These observed relationships were compared to ISO-1999 predictions for threshold changes due to NIHL.

For exposure times between 10 and 40 years the median value of expected NIPTS could be calculated. For exposure times shorter than ten years, median expected NIPTS values were interpolated from the value of NIPTS for ten years. For each participant predicted median NIPTS was calculated, both at baseline and at follow-up, based on their noise exposure history. The same was done for ARHL, and both components of hearing loss were added according to the formula described in ISO-1999. The annual rate of change in hearing predicted was assessed by subtracting the predicted hearing loss at baseline from the prediction at follow-up, divided by the duration of the measurement period in years. The relationship between predicted

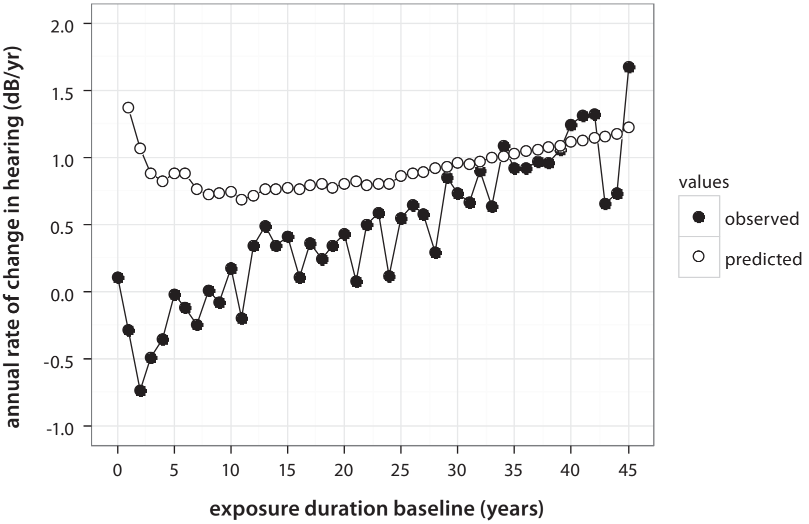


Figure 3.2. Observed versus ISO-1999 predicted annual rate of change in $PTA_{3,4,6}$ as a function of noise exposure duration at baseline

and observed annual rate of hearing loss as a function of exposure duration is shown in Figure 3.2. Again, a positive change indicates a deterioration of hearing ability, a negative change indicates an improvement in hearing threshold level.

In general, the growth of the hearing loss predicted by ISO-1999 was dominated by NIHL in the first years of exposure and reduced with increasing noise exposure duration; the predicted rate of NIPTS was highest for the shortest exposure duration, ranging from 1.9 dB at start of exposure to 0.4 after 10 years[1]. In the consecutive period of 10-40 years of exposure, the yearly growth rate due to noise exposure was only low and the increase in hearing thresholds was dominated by the ageing effect. The observed rate of hearing loss showed a quite different pattern. In general, hearing thresholds showed an improvement in the workers that were exposed to noise for the shortest duration (<10 yrs). For longer durations, indeed a deterioration in hearing was observed, and this rate of hearing loss tended to increase with increasing exposure duration due to effects of aging rather than effects of NIHL. For exposure durations exceeding 30 years the observed rate of hearing loss was reasonably consistent with ISO-1999 predictions. For workers exposed to noise for less than 30 years ISO-1999 tended to overestimate the degree of hearing loss increase (Figure 3.2).

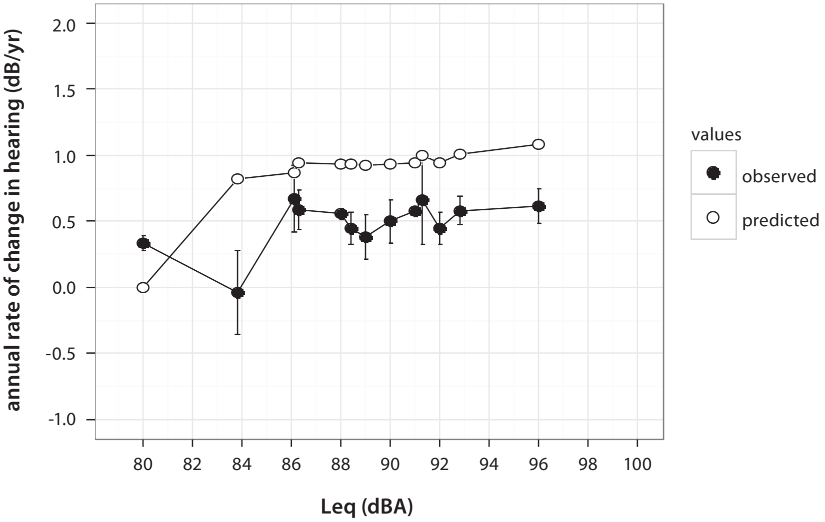


Figure 3.3. Observed versus ISO-1999 predicted annual rate of change in $PTA_{3,4,6}$ as a function of noise exposure level during measurement interval. Error bars represent one SE

---

[1] These values concern NIPTS only, therefore deviate from the values displayed in Figure 3.2 that reflect total predicted hearing loss based on both NIPTS and ARHL.

Similarly, $PTA_{3,4,6}$ values as function of daily noise exposure level were examined (Figure 3.3). $PTA_{3,4,6}$ values were almost evenly distributed over the range of noise intensities, except for the small group of participants exposed to average daily noise levels of 84 dBA. Both curves showed a similar pattern of increasing rate of hearing loss, although only slightly, with higher noise exposure level. This corresponds to the findings obtained in the linear mixed model described above.

Effects of covariates

The data collection also provided information about various demographic and work-related variables that could interact with NIHL development, such as the use of hearing protection. These variables may affect the degree of NIHL, as they may be associated with hearing damage or may increase a participant's susceptibility to noise. To investigate the relationship of these variables with change in hearing loss change a linear mixed effect model was fitted containing all these variables (see Table 3.1) as fixed factors. After initially fitting this full model, factors with non-significant terms as assessed by conditional F-tests, were eliminated from the model. The final model contained 7 fixed effects, as well as a random effect for 'subject'. Coefficients are presented in Table 3.5.

Table 3.5. Coefficients and 99% confidence intervals of the total model predicting the annual rate of change $PTA_{3,4,6,}$ containing all significant covariates.

Model terms	Coefficient	99% CI
Intercept	-0.481	-0.796 – -0.166
Ear: left	0.144	0.079 – 0.208
Baseline age	0.079	0.073 – 0.086
Noise level	0.029	0.008 – 0.049
HPD use: yes	0.216	0.061 – 0.371
Baseline hearing	-0.055	-0.059 – -0.051
Hearing complaints: yes	1.007	0.846 – 1.167
Interval	-0.226	-0.295 – -0.156

The variables that remained significant in the model, in addition to the shown effects of age (F[1,11636] = 990.90, $p < 0.001$) and noise intensity (F[1,11636] = 12.39, $p = 0.004$), included tested ear, use of hearing protection, baseline hearing level, presence of complaints about hearing and interval duration. Noise exposure duration did not show a significant contribution to this multivariable model either (F[1,11636] =2.54, p = 0.111).

This multifactorial model showed that the annual rate of change in hearing loss was 0.14 dB greater in the left than in de right ear (F[1,11636] = 28.71, p < 0.001). Also, baseline hearing level, expressed in $PTA_{3,4,6}$ showed a significant effect (F[1,11636] = 1166.79, p < 0.001). The use of hearing protection showed a positive association with change in $PTA_{3,4,6}$ (F[1,11636] = 12.50, p = 0.003), indicating that employees using hearing protection showed a 0.21 dB greater annual change in $PTA_{3,4,6}$ than those who did not. Participants having subjective complaints about poor hearing showed a change in hearing level that was significantly larger (0.99 dB) than that of participants without hearing complaints (F[1,12162] = 250.56, p < 0.001). There was a strong association between rate of hearing loss and duration of the intermediate measurement interval (F[1,12162] = 9.75, p < 0.001); the negative coefficient of -0.23 dB indicated that the annual rate of change in hearing loss became smaller with increasing interval duration.

Pattern of hearing loss development

Finally, the subgroup with three audiograms was analysed to investigate the pattern of hearing loss development. Hearing loss deteriorates over time, due to both exposure to noise and aging effects. In the majority of the employees that were tested twice, the extent to which noise and the extent to which ageing were responsible for this reduction in hearing sensitivity over time could not be established. In case of three measurement occasions, a distinction between both causes of hearing loss could be made based on the pattern of hearing loss development, albeit only at group level rather than individually. The ISO-1999 model showed that NIHL steeply increases during the first decade of noise exposure, followed by a slowing rate of growth with prolonged exposure to noise. This should result in a logarithmic progression, or concave form of hearing loss growth over time rather than a linear relationship. Presbyacusis on the other hand is known to be a progressive hearing loss over time, which is manifested as a convex rate of growth, especially in the higher-frequencies.

The complete linear mixed effect model containing all significant covariates showed that interval duration was negatively associated with annual rate of hearing loss in $PTA_{3,4,6}$ (Table 3.5). This indicated that the rate of hearing loss became smaller with increasing interval length, which corresponded to a concave course of hearing loss development over time. In order to verify this pattern of development, the rate of hearing loss as a function of interval level was investigated in the subgroup of 3,111 workers that were tested at three occasions. To do so, individual baseline $PTA_{3,4,6}$ was subtracted from the PTA-values obtained at both follow-up measurements to obtain the difference in hearing relative to baseline. Then a linear interpolation between baseline $PTA_{3,4,6}$, which was set at 0 dB, and the difference in $PTA_{3,4,6}$ obtained at the last follow-up measurement was fitted. When hearing loss develops linearly over time, the differences obtained at the intermediate follow-up measurement should fall onto this linear interpolation line. Figure 3.4 shows this linear interpolation against mean shifts in hearing as a function of interval length (presented as percentage of total interval time, in bins) and the average difference in $PTA_{3,4,6}$ established at the intermediate measurement.

A paired Student's t test comparing the observed difference at the intermediate measurement occasion and the shift predicted by linear interpolation showed significant differences; observed differences in $PTA_{3,4,6}$ were significantly higher ($p < 0.001$) in both ears than those based on linear interpolations (0.80 dB in the right and 0.94 dB in the left ear). This demonstrated the concave course of hearing loss growth that corresponds to NIHL predictions.

Discussion

The aim of this study was to describe the change in hearing threshold levels in a large group of noise-exposed male construction workers, as monitored by regular periodic audiometric screening over a period of approximately 4 years. Overall, a small average deterioration of hearing threshold levels was observed, ranging from 0.5 to 3.0 dB HL depending on tested frequency. Although baseline HTLs of left ears were slightly poorer than those of the right ears, hearing loss development in both ears was similar.

The annual rate of change in hearing over time ranged from 0.2 to 1.0, and was highest at the high frequencies that are sensitive to noise. Analysis of of change in hearing over time in a subgroup with three audiometric assessments showed a concave pattern of hearing loss development, which corresponds to NIPTS development. So, the pattern of hearing loss development indicates that the high-frequency hearing loss observed in this population is mainly attributable to noise exposure, rather than to presbyacusis that also begins to play a significant role in these middle-aged noise-exposed workers.

The average annual rate of hearing loss in $PTA_{3,4,6}$ was 0.56 dB/yr for the total population. This is lower than the annual shift in total HTL of 0.94 dB/yr predicted by ISO-1999 (1990). This can also be observed from Figures 3.2 and 3.3, displaying ISO-1999 predictions for the total change in hearing loss that are higher than the observed rates of change in hearing thresholds. The finding that noise-exposed workers did not lose hearing ability as fast as expected, was observed in previous studies as well (Seixas et al, 2005; Clark & Bohl, 2005; Rabinowitz et al, 2011). It might indicate a smaller NIHL development in this study, due to beneficial effects of hearing

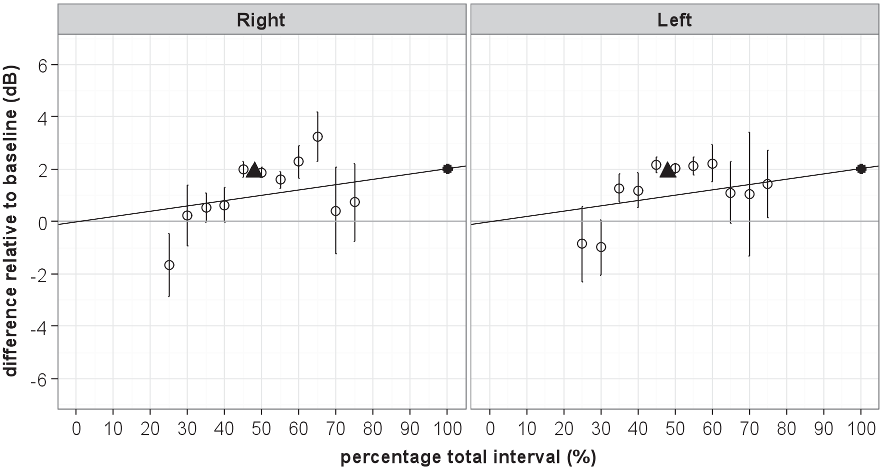


Figure 3.4. Difference in intermediate $PTA_{3,4,6}$ relative to baseline as a function of interval duration, displayed as the percentage of the total interval duration. The solid line represent the linear interpolation between the baseline $PTA_{3,4,6}$ and $PTA_{3,4,6}$ obtained at the second follow-up test. The black square represents mean $PTA_{3,4,6}$ at second follow-up, the black triangle represents the mean $PTA_{3,4,6}$ at intermediate follow-up measurement.

conservation interventions that result in less NIHL. In addition, the noise exposure levels used in this study were only rough estimations of actual exposure levels based on job titles, which might have introduced differences between observed and predicted hearing loss (Rabinowitz et al, 2007).

Nevertheless, the observed rate of change in hearing was also smaller than predicted for age-related hearing loss alone, which was on average 0.86 dB/yr for the total study population. This indicates that the total change in hearing is less than the sum of the effects predicted for noise exposure and age. Albera et al. (2010) observed that the progression of presbyacusis in noise-exposed listeners with NIHL was less than predicted for non-exposed individuals according to ISO-1999. The cochlear structures already damaged by exposure to noise, cannot be significantly damaged by age-related effects anymore (Albera et al, 2010). The results of the total linear mixed effect model predicting annual rate of change in hearing demonstrates this as well; baseline hearing was negatively associated with the degree of hearing loss development, indicating that subjects with higher PTA-values, thus more hearing loss, showed a reduced increase in hearing loss compared to normal-hearing subjects.

Change in hearing and noise exposure

When looking at the relationship between annual rate of change in hearing and noise exposure, hearing loss develops faster with increasing noise exposure and increasing age, which was expected from the ISO-1999 model predictions. However, fitted coefficients were small and exposure duration, as assessed at baseline, did not significantly affect hearing loss development when adjusted for age and noise level. This also corresponds to ISO-1999 predictions, which show that the rate of NIHL in exposed workers decelerates after the first 10-15 years of noise exposure (ISO-1999, 1990; Rösler, 1994). Effects of covariates were also assessed using a linear mixed effect model. The variables available in the data collection showed similar effects on the relationship between the annual shift in hearing and noise and age as was found for absolute hearing in Chapter 2, except for job change and noise nuisance during work. Participants having complaints about their hearing at follow-up show a larger increase in hearing loss over the measurement period than their colleagues without complaints. Workers that indicated to have used hearing protection during the measurement period showed larger annual shifts in hearing than those who did not use hearing protection. Although this contradicts an expected protective effect of HPDs, it corresponds to the positive association of using HPDs and hearing loss that was observed in Chapter 2. As was observed there, in the current study workers reported to use HPDs were exposed to higher noise intensities than those who report not to use HPD. Moreover, the binary variable of self-reported HPD usage is much less informative than data on actual consistency of usage, which would be a more accurate predictor of hearing loss.

In addition, tested ear, baseline hearing, and interval duration showed a significant effect on the rate of change in hearing. When adjusting for all other significant covariates, the left ear showed a slightly larger change in hearing loss than the right ear. This may be related to the significant effect of baseline hearing status, which was negatively associated with annual shift in hearing. Finally, interval duration negatively affected change in hearing; the annual rate of hearing loss decreased for increasing intervals, which corresponds to the concave pattern of hearing loss development caused by exposure to noise.

Development of hearing loss during the first decade of exposure

Although duration of noise exposure showed no significant effect on rate of hearing loss when adjusted for age, noise level, and available covariates, their relation is of interest, particularly for workers with baseline exposure for less than 10 years. The cross-sectional data of Chapter 2 showed a strong deviation from ISO-1999 predictions; mean age-corrected $PTA_{3,4,6}$ of the noise-exposed workers was 10 dB HL at the beginning of employment, which increased slightly with increasing exposure duration. This was in contrast to the predicted steep increase in hearing loss from 0 dB HL at the begin of employment. Hearing loss at the start of employment was also found by others (Seixas et al, 2004; Rabinowitz et al, 2006; Seixas et al, 2012) and might be the result of a pre-existing hearing loss when entering the workforce, due to previous educational, occupational, and recreational noise exposure. However, the finding could not be explained by the available cross-sectional data in Chapter 2, and longitudinal analysis of follow-up data of this subgroup was thought to enlighten this deviation.

Yet, it is known that screening audiometry applied in a survey, as was the case during the POHEs in this study, yield poorer hearing threshold levels than laboratory methods (Dobie, 1983; Schlauch & Carney, 2012). Data from a public health survey in the USA conducted between 1935 and 1936 (Glorig, 1956) were used to derive the reference thresholds described in the first standard defining average normal hearing by ASA (1951). The currently used ISO standard of audiometric zero (ISO-389.1, 1998) is derived from data obtained in several laboratory studies. Differences between both standards are known to be about 10 dB HL in favor of ISO reference levels. These differences reflect the differences in survey and clinical audiometry. Using the clinically obtained ISO-398.1 reference as audiometric zero leads to mean normal-hearing thresholds obtained from group survey data that fall at values near 10 dB HL (Schlauch & Carney, 2012).

Actually, this is seen in the audiometric data obtained during POHEs that are presented here, as well as in Chapter 2. All mean or median low frequency HTLs presented in these studies are around 10 dB HL (see Figures 2.1 & 3.1), indicating that reliable measurements up to 0 dB HL could not be established in an occupational audiometric survey. The observation in Chapter 2 of a 10 dB HL loss at the start of employment was suggested to be a result of pre-existing hearing loss when entering the workforce. Although this theory still may be valid, the limited ability of screening audiometry to accurately assess normal hearing threshold levels up to lower values than 10 dB HL might be an alternative explanation for this finding. In that case, the age-corrected $PTA_{3,4,6}$ value of 10 dB HL reflects the average normal hearing threshold in survey observations rather than pre-existing hearing loss.

Although survey methods yield poorer average normal thresholds levels than laboratory methods, useful conclusions about trends in hearing loss over time can still be drawn from group survey data (Royster & Royster, 1986). Therefore these longitudinal analyses, instead of cross-sectional evaluation, were used to investigate the development in hearing loss during the first decade of exposure. Unfortunately the workers exposed to occupational noise for a period of 10 years or less showed a negative mean rate of hearing loss (Figure 3.2), suggesting an improvement in hearing ability instead of the noise-induced deterioration that was expected (ISO-1999, 1990; Rösler, 1994). This finding is rather unfortunate, because it blurs any detrimental effects of noise exposure on hearing during the first years of exposure.

More so, the improvement in hearing is also reflected in both models predicting the annual rate of change in $PTA_{3,4,6}$, which have a negative intercept, indicating that subjects in the reference condition experience an improvement in hearing. To gain more insight in this finding, average hearing threshold levels of baseline and follow-up measurements were plotted, for 5 age groups (<25, 25-34, 35-44, 45-54 and ≥55 years) separately in Figure 3.5.

The youngest age groups show better HTLs at follow-up compared to baseline at the majority of the tested frequencies. For the group with a mean age of 40 years, hearing thresholds of both measurements seem similar, whereas the older age groups show the expected deterioration in hearing at follow-up at 2 kHz and higher. However, a small but significant improvement at 0.5 kHz was shown in all age groups except for the oldest workers. Clearly, some degree of HTL reduction was expected in all age groups in this study population, either due to progressive NIHL in the shorter exposed young workers or to presbyacusis in the older groups. Whereas the absence of a significant decrease in hearing ability of the younger workers would indicate that hearing conservation effectively prevented the development of NIHL, an average improvement in hearing across a group of workers is highly unexpected. The most probable explanation for such a change would be alterations in test equipment or

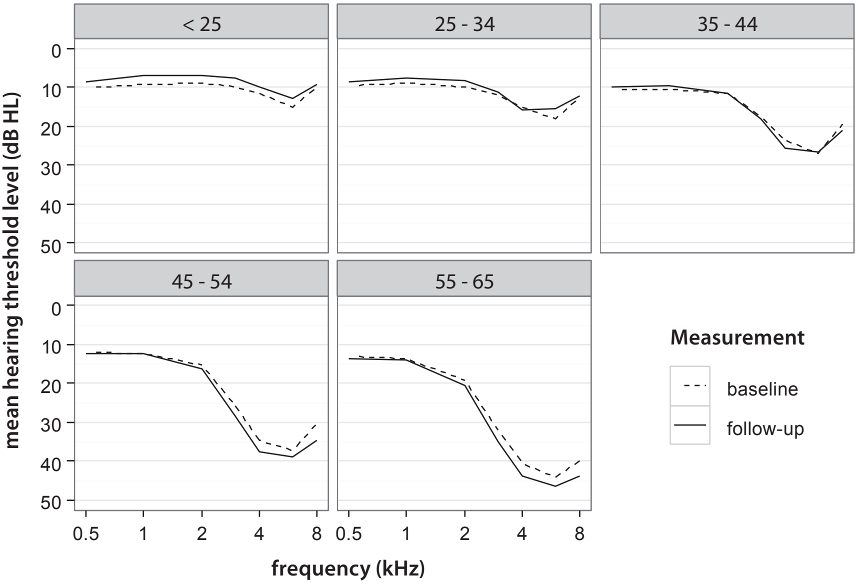


Figure 3.5. Mean hearing threshold levels obtained at baseline (dashed lines) and follow-up (solid lines), separated for five age groups.

measurement procedures. However, standardized audiometry was conducted according ISO-8253.1 (2010) and no systematic changes in this standard or in equipment and test characteristics could be identified during this follow-up period.

Behavioral audiometric thresholds vary somewhat from one test to the next, because of tester and patient experience and motivation (Schlauch & Carney, 2012). Clinical audiometry reports small test-retest variability, varying from 2.1 to 6.8 dB depending on frequency (Hétu, 1979; Dobie, 1983; Hall & Lutman, 1999), which slightly increases with increasing interval length. When audiometric testing is applied in occupational screening this variability increases even more due to various sources of systematic and random errors (Hétu, 1979). The control of most of these sources is specified in occupational standards for adequate screening (ISO-8253.1, 2010). However, in the practice of industrial screening these requirements cannot always be met. Given the fact that there have been no changes in the procedure for calibration, and that there has not been an systematic change in the type of audiometers and/or the method of audiometry, the most important factors that could have influenced the results in the current study are:

- Influence of background noise levels in the testing room; because audiometry requires determination of the lowest signal level that a person can hear, ambient noise in the audiometric test environment should be under the maximum permissible ambient noise levels specified by ISO- 6189 (1983). Test rooms calibrated according to this standard will make it possible to test to 0 dB HL for persons whose hearing is that sensitive (Franks, 2001). Nevertheless, these levels are rarely achieved without an audiometric soundproof booth, which is not always available in occupational assessments. In that case, tests are performed in a quiet room, introducing possible interference of background noise. If the availability of a test booth and/or the quality of the sound isolation improved over the years, this is the most likely explanation for the improvement of hearing ability, especially for subjects with hearing in the normal range. Also, the use of different types of supra-aural headphones might have introduced differences in the amount of background noise levels caused by headphone attenuation (Franks, 2001).

- Residual TTS at the time of testing; hearing screening is regularly performed during a working day. Exposure to noise prior to audiometric testing could result in temporary threshold shifts in hearing. Consequently, employees need to be advised to have a period without noise exposure of 14-16 hours preceding the hearing test (May, 2000; Franks 2001; NVAB, 2006). In practice, this is difficult to accomplish, and temporary effects of noise exposure, either occupational or non-occupational, on HTLs cannot be ruled out completely, hence the observed improvement could reflect a reduction in the degree of TTS (Seixas et al, 2005; Rabinowitz et al, 2011). There is a constant effort to better meet the criteria for a noise-free period and if this has been successful in the past years, this is a second potential explanation for the improvement of hearing ability, especially for the high frequency thresholds in noise-exposed participants.

- Familiarization with the examination procedure; it is possible that familiarity with the examination procedure might lead to an improvement in performance (Hétu, 1979). Royster et al. (1980) observed an improvement of 0 to -1 dB/yr in HTLs at 3, 4 and 6 kHz with respect to baseline, over the first 3 to 4 annual audiometric tests. They consider this attributable to a learning effect.

In addition, other factors such as differences in earphone placement, partial or complete obstruction of the external auditory canal by cerumen, interfering signals from the test equipment, bias introduced by the tester or the examination procedure, the instruction and the presence of tinnitus influence test variability (Hétu, 1979). All above-mentioned sources of error may have influenced the obtained thresholds to some extent, and the improvement in HTLs indicates that these might have been more prominently present during the baseline assessments than during follow-up. However, specifications of test conditions were not available in current data collection. Consequently, above-mentioned suggestions can only be offered as likely and not a certain explanations.

This makes it also unclear whether the causes of the improvement in HTLs are restricted to examinations of the younger workers only, or that these affect the entire cohort. Because the young workers do not show any aging effects yet, measurement variation is reflected more clearly in this subset than in older workers who additionally show some degree of age-related hearing loss (Figure 3.5). The consistent improvement in HTL at 0.5 kHz in all but the oldest age groups, as well as the observed rate of change that is smaller than predicted, indicate however, that measurement variability concerned the entire study population, underestimating the rate of hearing deterioration due to NIHL.

Quality of the survey data

Data from audiometric survey of the entire Dutch workforce of the construction industry has the advantage of its large size. Despite the rather high measurement variability, longitudinal analysis on group level can demonstrate trends in hearing loss over time. In addition, by analysing this large amount of data, a judgment on the quality of the collected data can be given.

First of all, by merging baseline and follow-up data sets, inconsistencies between data from multiple measurements of the same individual were revealed. In 15% of the cases there was a lack in correspondence between baseline and follow-up data concerning demographic variables, such as gender and date of birth, or work-related variables, such as job title or employment years that are used to estimate noise exposure. Because correct data could not be recovered, these cases were excluded from further analysis. Most of this data derived from the self-administered questionnaire, hence may be the result from recall bias of individual workers completing it. Also typographical errors when entering questionnaire responses may have induced these deviations.

Another 15% of the cases in this study was excluded because any of their audiometric assessments suggested evidence of hearing loss due to other than noise-induced or age-related causes, since including these cases would disturb the assessment of the relationship between noise exposure and hearing loss. Several criteria were defined to exclude cases with flat and/or conductive losses that do not result from exposure to noise. In screening audiometry, only air conduction thresholds can be reliably obtained, omitting bone conduction and thereby the possibility of correcting for the conductive component of the hearing losses. Instead, more information about the otologic and medical history in the data collection would be helpful in interpreting audiometric abnormalities.

In addition, cases showing a large unilateral difference between baseline and follow-up, either deteriorations or improvements, were excluded. In general NIHL develops more of less bilaterally, and only small difference between ears are expected. The excluded cases showed an unexpectedly large change in hearing ability over the 4-year time period, in only one ear. This reflects bias in audiometric measurements rather than an actual change in HTLs.

Like many surveys conducted in ‘real world’ environments, some information in this study is poorly quantified or absent. Analyses as those conducted in this study would give more accurate results when information of otologic and medical history, exposure to non-occupational noise, individual noise dosimetry, actual attenuation from HPDs and the consistency of their usage, and test conditions was available.

Conclusion

Over an interval period of 4 years, an overall deterioration of hearing threshold levels of construction workers was established on group level. The annual rate of change in hearing loss was positively associated with both age and noise intensity. Analysis of the pattern of hearing loss development indicates that the observed change in $PTA_{3,4,6}$ was in correspondence with predicted NIHL development.

Current longitudinal analyses could provide only limited relevant information on the development of hearing loss during the first decade of noise exposure; instead of a sheer deterioration in hearing, an improvement of hearing threshold levels was found. Our hypothesis is that this was rather the result of measurement variability in screening audiometry than an actual improvement of hearing ability. In addition, average HTLs reflecting normal-hearing, such as those of the youngest workers and low-frequency HTLs of the total study population, are 10 dB HL. This increased value for normal hearing in this audiometric survey is a likely explanation for the hearing loss present during the first decade of noise exposure observed in Chapter 2, although some degree of pre-existing hearing loss could not be ruled out completely.

These analyses showed that a large data collection of audiometric survey data can be used to assess group effects in hearing over time. Inconsistencies in data and measurement factors affecting the stability of the database showed that small shifts in individual hearing threshold levels cannot easily be distinguished from normal measurement variability. Additional data collection and a better specification of the test conditions and procedures might improve this.

Appendix: details on exclusion criteria used

In some cases a medical record could not be used for analysis. Of the 17,390 subjects that were examined twice, 5,128 were excluded from further analysis. In addition, of the subset of 4,645 subjects with three measurements, 1,434 cases were omitted. Reasons for exclusion were briefly mentioned in the methods section. Below, the specific definitions and reasoning for the used exclusion criteria are described.

Insufficient follow-up period

Although POHEs are offered once every two or four years, depending on employee's age, some medical records were collected with a different frequency. To ensure that the interval between baseline and follow-up measurements was sufficiently long to establish a change in hearing loss, this period should be at least one year.

Incorrect data collection

To make sure analysis were performed on actual audiometric thresholds, and that accurate data on noise exposure was available, similar exclusion criteria were defined as those used for the baseline cohort in Chapter 2. 222 subjects of current study population had no recorded audiometric data and 84 participants showed HTLs exceeding 90 dB HL at either one or more frequencies measured in both ears (referred to as code '95'). In addition, 526 subjects showed missing or immeasurable HTLs exceeding 90 dB HL in one ear and thresholds of 90 dB HL or better at all frequencies in the contralateral ear. For these subjects, only the contralateral ear was preserved in the dataset, and 240 left and 286 right ears were excluded from analyses. Finally, 81 female workers were discarded because of their concentration in non-noise-exposed jobs.

In addition, criteria were defined to check for incorrect of missing data regarding noise exposure estimations. 415 workers had insufficient noise exposure data missing either information on job title or duration of employment. In case the data of these workers were available in the baseline data collection, missing follow-up data can be adopted from the baseline set after merging both databases, so these subjects were not excluded yet.

Lack of correspondence between successive datasets

The merge of the baseline and follow-up data provided an opportunity to control the quality of the data, by checking the data obtained during two examinations for correspondence. When there was no correspondence between data, subjects were excluded from analysis since it could not be revealed which of the two data collections contained accurate data.

First, date of birth was compared, and in 23 subjects different values were reported. These cases were excluded from the dataset. In addition, factors important for noise exposure estimation, such as reported job title and years employed in construction were also compared. Reported job title was used to estimate the workers' daily noise exposure levels. For 4,178 worker there was no correspondence between reported job titles, and, more importantly, estimated noise exposure intensity deviated for 1,762 subjects. According to the information in the medical records, 453 of these workers recently changed their jobs. Correct daily noise exposure could thus not be salvaged for the remaining 1,309 worker, hence they were not included in analyses.

The reported amount of years worked in construction, which defines the duration of noise exposures, is also checked for correct correspondence. The difference in reported years is calculated, accounting for the interval between measurements. 1,314 records reported different data for employment duration, showing a deviation between both measurements that exceeded five years. Because correct data for exposure duration could not be recovered, these records were omitted.

In total, 15.5% of the data collection is excluded based on the lack of correspondence between both datasets.

Audiometric discrepancies

Participants were excluded from present study if an audiometric assessment at any measurement occasion suggested evidence of hearing loss due to other than noise- or age-related causes. The following exclusion criteria were defined:

- Diagnosed hearing loss due to other etiologies than noise and age; each medical record reported the diagnosis of an otological disease if present. Eighty participants were diagnosed with hearing loss due to another cause than noise or aging. Since the focus of this study was on the development of hearing loss caused by noise exposure, these subjects were excluded.

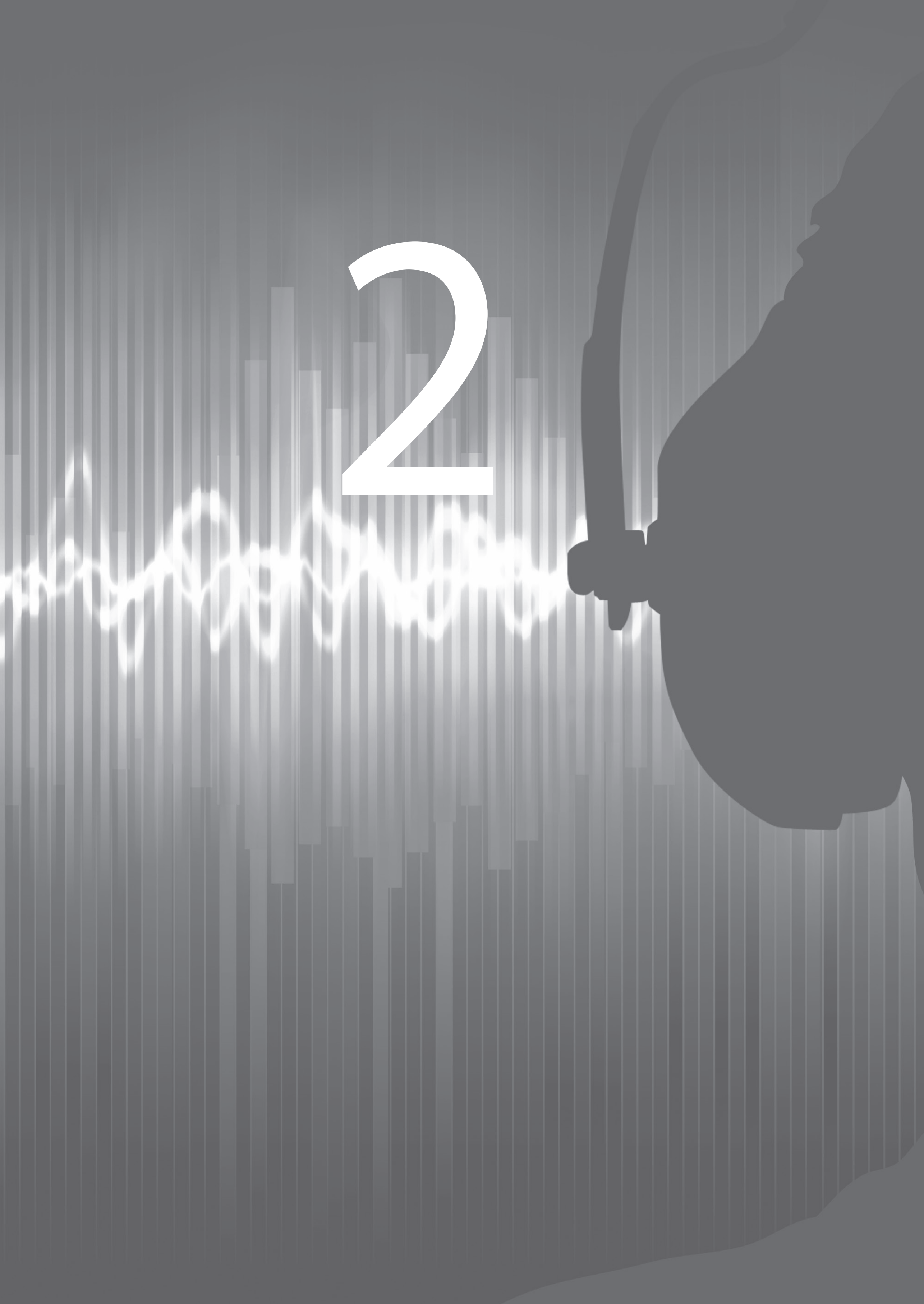
- Audiometric configuration; based on previous research regarding NIHL, the study population was divided into subgroups according audiometric configuration (Jansen et al, 2008; Helleman et al, 2010). Five groups and a rest group were defined: normal hearing, subnormal hearing, mild notch, profound notch, and sloping audiogram. Normal hearing was defined as having every threshold at 20 dB HL or better, and the subnormal hearing group showed a flat loss with every threshold at 30 dB HL or better. The notching audiograms indicating NIHL had an elevation in hearing threshold at 3, 4, or 6 kHz when compared to the average of 0.5, 1, and 2 kHz and the better threshold of 6 and 8 kHz, which was small in the mild notch group and larger in the profound notch group. Finally, the sloping audiogram was defined to have similar thresholds at 3 and 4 kHz as the notcht groups, but without showing an improvement at the higher frequencies, indicating age-related high-frequency hearing loss. 215 participants having both audiograms defined as 'rest' showed audiometric configurations likely to correspond to causes of hearing loss beyond the scope of this article. Therefore, these subjects were excluded from analysis. In addition, data of 794 ears defined as 'rest' were discarded.

- Conductive hearing loss; NIHL is sensorineural, hence no conductive loss was expected as a result of noise exposure. Since bone-conduction thresholds cannot be adequately assessed in a hearing screening program, an alternative criterion concerning conductive hearing loss was defined using ISO-1999 predictions; low-frequency hearing loss caused by age or noise exposure was expected not to exceed 40 dB HL. So 50 subjects, and additionally 31 ears, having a pure-tone average of 0.5 and 1 kHz > 40 dB HL were considered to have conductive losses, and were excluded.

- Large unilateral change in hearing ability compared to the change in the contralateral ear; a rough analysis of the differences in hearing thresholds over the 4-year measurement period showed that there were both very large deteriorations as well as improvements in hearing ability. So, a confidence criterion should be composed to define the limits of reliable differences. To do so, the observed change in one ear is compared to the observed change in the contralateral ear. Since NIHL is mostly symmetrical, these differences should be more or less similar. A confidence interval of change was calculated by the median difference ± 3 * standard deviation of the difference, and rounded to 5 dB intervals. This way an interval of change of -25 to 25 dB HL was defined for the lower frequencies up to 4 kHz, and an interval of change of -45 to 45 dB HL for 6 and 8 kHz. Participants having differences in HTL change between ears that lie outside this interval were considered outliers. Based on this criterion, 1,783 subjects were excluded from the dataset.

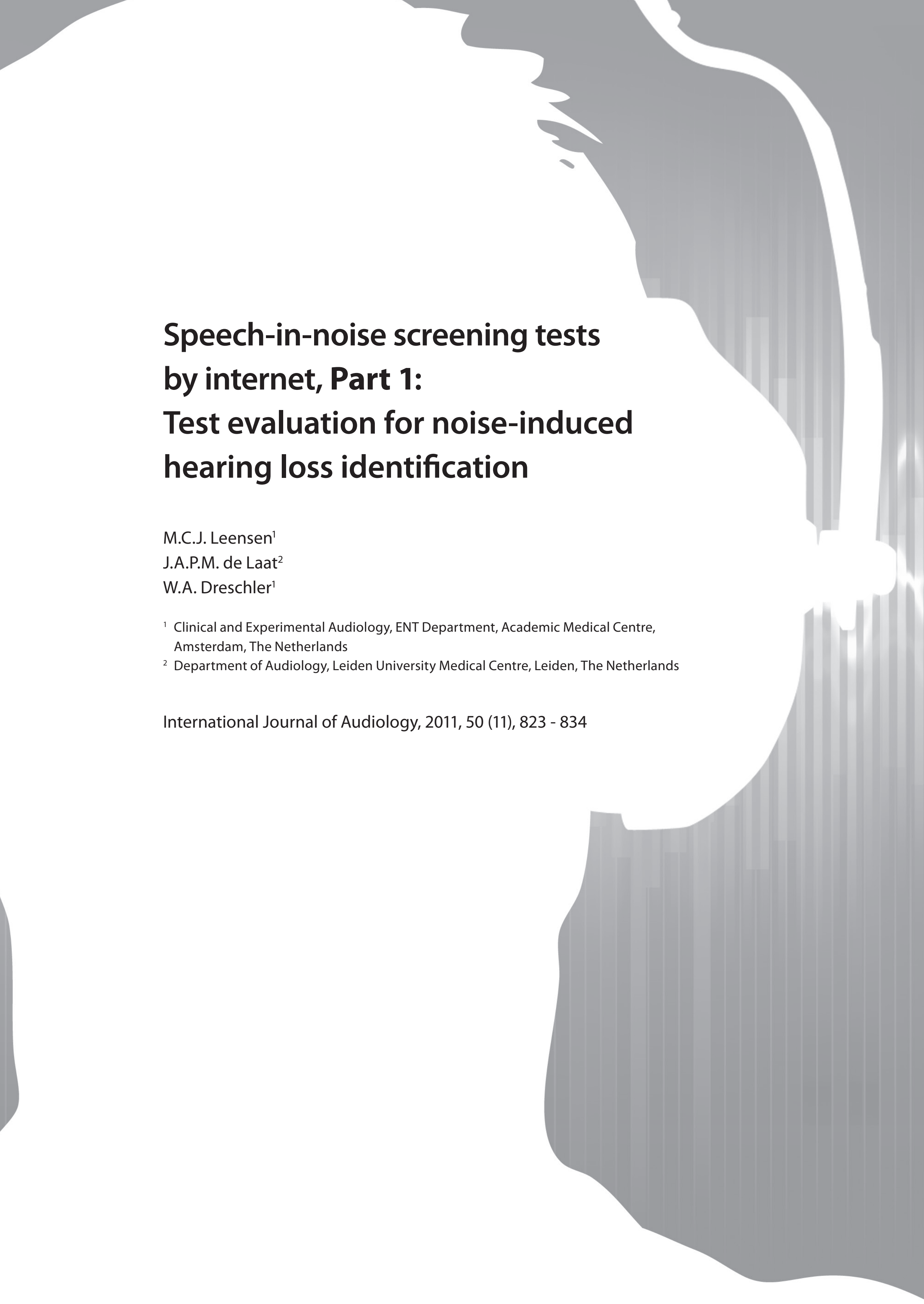
- Large change in low frequency hearing thresholds; after the exclusion of the above-mentioned cases, still very large differences in hearing thresholds existed, which were much greater than expected to occur over a 4-year period. Low frequency hearing thresholds are affected by noise and age only in a minor extent. So, in order to reduce the large unreliable differences observed, ears that showed a change in HTLs at 0.5 or 1 kHz that exceeded the confidence interval of change of -25 to 25 dB HL for change in hearing were also excluded from analysis. 32 participants that showed this in both ears were excluded, and another 133 ears were discarded for this reason as well.

2



Development of an alternative approach using speech-in-noise testing

4



Speech-in-noise screening tests by internet, Part 1: Test evaluation for noise-induced hearing loss identification

M.C.J. Leensen[1]
J.A.P.M. de Laat[2]
W.A. Dreschler[1]

[1] Clinical and Experimental Audiology, ENT Department, Academic Medical Centre, Amsterdam, The Netherlands
[2] Department of Audiology, Leiden University Medical Centre, Leiden, The Netherlands

International Journal of Audiology, 2011, 50 (11), 823 - 834

Abstract

Objective: In the Netherlands three internet-based self-screening tests for hearing have been developed; the National Hearing Test (NHT), Earcheck (EC), and Occupational Earcheck (OEC). These tests are adaptive speech-in-noise tests using either digit triplets or monosyllables, presented in stationary speech-shaped noise. These tests can be highly valuable in increasing the awareness and prevention of noise-induced hearing loss (NIHL). This study evaluates these online speech-in-noise tests and investigates their potential to detect NIHL.

Design: In a multi-centre study the results of the three online screening tests are compared to pure-tone audiometry and to the Dutch sentence SRT test (Plomp & Mimpen, 1979a), which is considered the clinical standard.

Study sample: In total, 49 normal-hearing listeners and 49 patients with different degrees of NIHL participated.

Results: The online tests show good reliability, but there is much overlap in outcomes between normal-hearing listeners and participants with early NIHL. In addition, rather low correlations of the tests' results with both the Dutch sentence SRT test and pure-tone thresholds are found. These findings result in rather low test sensitivity: 54% (NHT) and 51% (EC), or low specificity: 49% (OEC).

Conclusions: The online screening tests in their current form are unsuitable to be used for early NIHL screening purposes.

Introduction

Noise-induced hearing loss (NIHL) is a significant social and public-health problem. In present society large groups of individuals are frequently exposed to high sound levels, either during leisure time or in occupational settings. Currently, NIHL is the most reported occupational disease in the Netherlands (Van der Molen et al, 2010). Occupational NIHL is generally detected by pure-tone air conduction audiometry. In the Netherlands, this is offered at least once every four years to all employees exposed to daily noise levels exceeding 85 dBA, and occurs on a voluntary basis.

Besides noise at workplaces, exposure to loud sounds is increasingly encountered during recreational activities. Concern is growing that over-exposure to amplified music, due to the use of personal music players or discotheque attendance, will cause NIHL in adolescents and young adults. Although evidence supporting a relationship between exposure to leisure noise and hearing damage in young people remains ambiguous (Meyer-Bisch, 1996; Mostafapour et al, 1998; Niskar et al, 2001; Biassioni et al, 2005; Shah et al, 2009; Zhao et al, 2010), any exposure to noise of significant intensity and/or duration is known to increase the risk of hearing damage. Considering the noise levels and the time spent listening to personal music players, approximately 5 to 10% of the listeners are estimated to be at risk of developing permanent hearing loss after five or more years of exposure (SCENIHR, 2008). Furthermore, the damage from chronic exposure to loud music is cumulative, so a slight hearing loss in adolescence can eventually become a substantial one in adulthood (Chung et al, 2005; SCENIHR, 2008), especially among those with higher susceptibility to noise (Biassoni et al, 2005) or those who are employed in a job with significant noise exposure.

Because of the gradual development of NIHL, persons with mild high-frequency hearing loss are often unaware of their impairment until the hearing loss reaches a certain degree (Vogel et al, 2009). Since hearing damage is irreversible, it is of major importance that it is recognized as early as possible. The earlier NIHL is detected, the earlier precautionary measures can be undertaken to prevent more impairing, permanent, hearing damage (Meyer-Bisch, 1996). One of these measures is changing young people’s personal behaviour related to the length of time and sound level at which music is played (Vogel et al, 2007). However, a Delphi study by Vogel et al. (2009) showed that adolescents first must become aware that they personally are at risk for hearing loss due to listening to high-volume music, before the promotion of protective behaviours would be useful.

An objective hearing test can be of great help in the early detection and prevention of NIHL. Access to an easily administered hearing screening test can raise awareness of possible hearing problems and reduce the risk of hearing loss after exposure to noise. Feedback of individual hearing status stimulates persons to seek audiological help

(Smits et al, 2004; 2006a) or to change their (music listening) behaviour in order to prevent NIHL.

The first noticeable disability caused by noise-induced hearing loss is often a reduced ability to understand speech in a noisy environment. Therefore, the currently available Dutch internet-based speech-in-noise tests can be considered for NIHL screening purposes. The first speech-in-noise self-test for the Dutch language was developed and validated by Smits et al. (2004). This screening test, referred to as the 'National Hearing Test' (NHT), presents digit triplets in noise, hence it can be easily administered by telephone. Subsequently, an internet version of this test was generated, in collaboration with the Dutch National Hearing Foundation (NHF) (Smits et al, 2006a). The NHT measures the ability to understand speech in noise by determining the speech reception threshold (SRT), i.e. the signal-to-noise ratio corresponding to 50% intelligibility. The test result is presented to the participant, accompanied by a recommendation for follow-up referral, if required. For reasons of comparison, the bandwidth of the online test materials is limited to 0.3 - 3.4 kHz, to mimic the telephone network frequency band (Smits et al, 2006a). Because NIHL predominantly affects the high frequency region, an internet-based speech-in-noise test that also includes the higher frequencies was generated: 'Earcheck' (EC). This test, following a similar procedure as NHT but presenting nine different CVC words in a broadband noise, was developed by LUMC Leiden and NHF (Albrecht et al, 2005). The test specifically aimed at young persons between 12 and 24 years old, to raise awareness about the risks of exposure to loud music in this population. A third speech-in-noise test was developed by LUMC and NHF, specifically applicable in commercial enterprises, to monitor the hearing ability of employees in noisy occupations: 'Occupational Earcheck' (OEC) (Ellis et al, 2006). The procedure for this test is similar to that of Earcheck, but it was designed to have better precision by increasing the number of stimuli and by consecutive monaural testing of both ears. Also, it uses a different set of CVC words, containing matching vowels and more high-frequency consonants (Ellis et al, 2006; Kuipers, 2007).

These speech-in-noise self-tests are considered to be suitable for screening purposes. Because of their adaptive nature, they can be implemented as quick and fully automated tests (Jansen et al, 2010), measuring over a range that includes both normal and impaired hearing (Soli & Wong, 2008). The screening tests are developed to be performed in an at-home or private situation. Consequently, respondents make use of a variety of computer equipment and transducers, and both presentation level and the level of ambient noise in the room are unknown. However, measuring the ratio of speech intensity and level of masking noise makes the test relatively independent from the absolute presentation level (Plomp, 1986). Since both speech and noise with identical spectra are played through the same playback device the signal-to-noise ratio in each frequency band is not sensitive to subtle differences in the transfer characteristics between different devices, making the test robust against possible transmission losses and variations in equipment (Smits et al, 2004; Culling et al, 2005). Finally, the tests are relatively robust against background noise (Culling et al, 2005), require little or no calibration (Jansen et al, 2010) and are insensitive to conductive hearing losses, provided that the presentation level is at the most comfortable loudness level.

When measuring the speech reception threshold in noise, higher SNRs are required for subjects with NIHL than for those without hearing loss (Chung & Mack, 1979; Smoorenburg et al, 1982; Bosman & Smoorenburg, 1995). However, the correlation between SRT in noise and the pure-tone audiogram is only modest (Smoorenburg et al, 1982). For subjects with NIHL the highest correlation coefficient was found when comparing their results of the Dutch sentence SRT test to their pure-tone average of 2 and 4 kHz (r = 0.72, Smoorenburg, 1990; Smoorenburg, 1992). When their SRT was obtained using CVC syllables this correlation was even lower (r = 0.59; Bosman & Smoorenburg, 1995).

Although subjects with a sensorineural hearing loss are more adversely affected by noise than normal-hearing subjects, studies examining patients with high-frequency hearing loss starting above 2 kHz showed similar word recognition in continuous noise as for normal-hearing listeners (Pekkarinen et al, 1990; Phillips et al, 1994). For sentence recognition in noise only small differences up to 0.5 dB were found (Festen & Smits, 2007, as reported in Rhebergen et al, 2010a). Festen & Smits (2007) used the speech intelligibility index (SII) model to predict that a sharp, infinitely deep notch of 2/3 octave width at 2 or 4 kHz would result in an increase in sentence SRT of only 2 dB. Word recognition requires only a small amount of information (Quist-Hanssen et al, 1979) and is often based on vowel recognition only (Smoorenburg, 1992). The online speech-in-noise tests contain speech materials consisting of a closed set of a small number of CVC syllables, most of which contain a unique vowel. Vowels contain speech information in the low and mid-frequencies, up to approximately 2.5 kHz. Subjects with NIHL often still exhibit normal or near-normal hearing threshold levels at these frequencies, so they may benefit from this preserved hearing for the understanding of CVC words in noise (Quist-Hanssen et al, 1979). Consequently, results of the speech-in-noise screening tests for subjects with NIHL are expected to deviate only slightly from normal performance, suggesting that the applicability of the speech-in-noise tests to detect NIHL might be low.

This study was designed to examine the performance of normal-hearing listeners and patients with different degrees of NIHL on the three Dutch online speech-in-noise tests, in order to investigate their potential to discover NIHL. For this purpose, a multi-centre study was conducted, in which the different online screening tests were compared to pure-tone thresholds and the Dutch sentence SRT test developed by Plomp and Mimpen (1979a). The test-retest reliability and validity of the three online tests are evaluated and the sensitivity of the three tests for NIHL is determined.

Methods

Participants

The number of normal-hearing listeners and hearing-impaired participants to be tested was based on a power analysis. This showed that a sample size of 44 in each group had 80% power to detect a difference between means of 2.0 dB, using a two group t-test with a 0.05 two-sided significance level and a common standard deviation of 3.3 (Kuipers, 2006; Jongmans et al, 2008). In this calculation the following assumptions were used: the normal-hearing (NH) group and the hearing-impaired (HI) group are expected to result in different outcome categories, so the difference between groups will be at least an interval width apart. A difference of 2.0 dB, corresponding to the smallest interval width found in Occupational Earcheck, was thus considered relevant. We anticipated that only 90% of included patients will have valid measurements, therefore in total 100 participants were included, 50 in each group.

Hence, data were collected from 50 participants with normal-hearing and 50 hearing-impaired subjects, measured at three different audiology departments; LUMC Leiden, UMCN St Radboud Nijmegen and AMC Amsterdam. Only participants who were native speakers of the Dutch language were included. The normal-hearing group (33 female, 17 male) consisted of college undergraduates, recruited from the universities allied to the three university hospitals, complemented by a small number of lab workers. On the day of testing, all subjects had pure-tone thresholds of 15 dB HL or better at the octave frequencies from 0.125 to 8 kHz (including 3 and 6 kHz), except for one who was excluded from further analysis, leaving 49 subjects in this group. Their ages ranged from 18 to 50 years, with a mean age of 27.0 years (SD = 8.4 years).

Subjects in the hearing-impaired group (3 female, 47 male) were patients of one of the three ENT departments who had recently received audiological evaluation. In addition, a small number of subjects that participated in previous research concerning occupational noise exposure and hearing loss of AMC Amsterdam completed the group. Patients with normal or near-normal low-frequency hearing (pure-tone thresholds at 0.125 to 1 kHz of 20 dB HL or better) and high-frequency hearing loss (one or more pure-tone thresholds at 2 to 6 kHz greater than 25 dB HL) were selected. The included subjects had a history of noise exposure, although it is impossible to prove a direct relationship between the hearing loss and the exposure to noise. Exclusion criterion was an air-bone gap greater than 15 dB in the tested ear. One patient did not meet the defined criteria and was excluded from further analysis, leaving 49 HI subjects for analysis with a mean age of 56.1 years (range 36-72 years, SD = 8.6 years).

A Student's t test showed that the hearing-impaired patients were significantly older than the normal-hearing students ($p < 0.001$). Across the three centres only small variation in the participants' age was observed and the NH participants tested in LUMC turned out significantly younger than the NH subjects tested in the UMCN. For more details on sub-group demographics see Table 4.1.

Table 4.1. Demographics of the population per centre, displayed for both the normal-hearing (NH) and hearing-impaired (HI) group.

		n	Sex (M) n (%)	Age mean (SD)	Right ear tested n (%)	SRT$_q$ mean (SD)	SRT$_n$ mean (SD)
NH	*AMC*	22	8 (36%)	26.5 (6.0)	14 (64%)	26.8 (2.3)	-5.6 (1.2)
	LUMC	12	2 (17%)	21.9 (4.4)	8 (67%)	31.2 (2.6)	-5.7 (1.0)
	UMCN	15	6 (40%)	32.1 (11.1)	9 (60%)	34.6 (2.2)	-5.1 (1.6)
	Total	**49**	**16 (33%)**	**27.0 (8.4)**	**31 (63%)**	**30.3 (4.1)**	**-5.4 (1.3)**
HI	*AMC*	10	9 (90%)	49.0 (7.8)	4 (40%)	30.7 (3.1)	-3.6 (1.5)
	LUMC	25	24 (96%)	58.9 (8.9)	12 (48%)	44.5 (8.1)	-0.4 (2.9)
	UMCN	14	14 (100%)	56.2 (6.5)	8 (57%)	42.0 (5.4)	-1.0 (3.1)
	Total	**49**	**47 (96%)**	**56.1 (8.6)**	**25 (51%)**	**41.0 (8.4)**	**-1.2 (2.9)**

The number of males and the number of participants whose right ear was tested are displayed, percentages are in parenthesis. Also, SRTs measured by Dutch sentence SRT test in quiet (SRT_q*) and in stationary noise (*SRT_n*) are shown, averaged over test and retest sessions.*

Audiology departments: LUMC Leiden, UMCN St Radboud Nijmegen, and AMC Amsterdam.

The group of hearing-impaired listeners was divided into two groups of participants having either a narrow audiometric dip (HI-ND), corresponding with early NIHL, or a broad dip (HI-BD) corresponding with more severe hearing loss. Distinction was made based on whether or not their hearing threshold level at 2 kHz was affected; when hearing threshold at 2 kHz was more than 15 dB HL poorer than the pure-tone average of the lower frequencies 0.5 and 1 kHz, the patient was classified as having a broad dip (n = 24, mean age 58.5 years, SD = 10.2 years). If not, the participant was classified as having a narrow dip (n = 25, mean age 54.1 years, SD = 7.6 years).

A power analysis, using the input as described above, shows that a sample size of 15 in each group is required to have 80% power to detect a difference between the three subject groups with a one-way ANOVA. For each of the three subject groups, mean audiometric hearing thresholds of the ears selected for monaural testing are displayed in Figure 4.1. See 'procedure' for more details on the criteria used to select the tested ear.

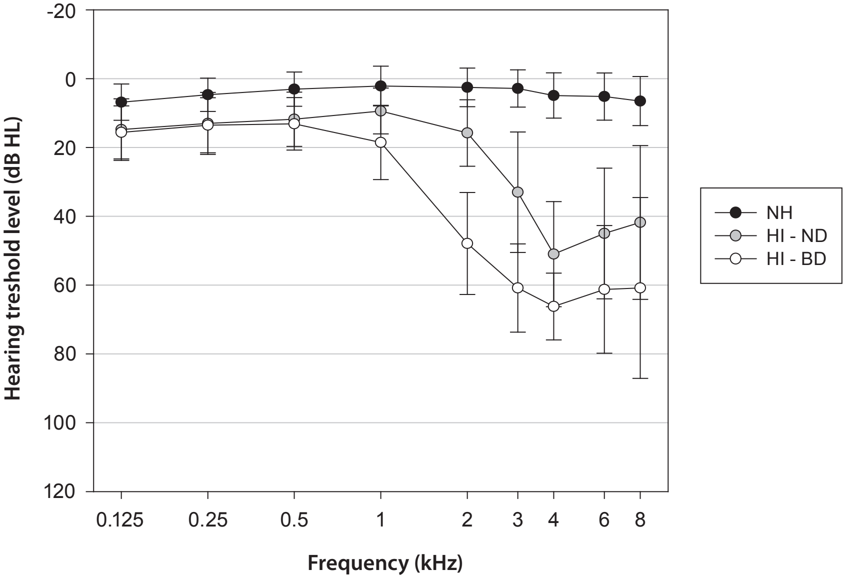


Figure 4.1. Audiometric thresholds of the ear selected for monaural testing, averaged for each of the three subject groups. Error bars represent one SD.

Procedure

All participants signed informed consent forms before starting the experiment. The experimental protocol and all procedures in this study were approved by the ethics committee of the University of Amsterdam (approval number: 08/049).

At the beginning of the experiment, a pure-tone audiogram was recorded at the octave frequencies of 0.125 - 8 kHz, including also 3 and 6 kHz. In addition, bone conduction was measured at 0.25, 0.5, 1, 2 and 4 kHz. The different speech-in-noise tests will be compared in case of monaural signal presentation, since OEC is designed to measure both ears separately. The test ear was chosen based on audiogram configuration. For the normal-hearing listeners, this was either the subject's best ear, or, in the case of no difference, the right ear. For the NIHL subjects, the ear showing the most pronounced audiometric dip was selected.

Following pure-tone threshold testing, participants performed the Dutch sentence SRT test developed by Plomp and Mimpen (1979a). This test applies open sentences, has a very high test-retest reliability of 0.9 dB, and is used in most clinics in the Netherlands. Therefore, this Dutch sentence SRT test is considered as the clinical (or 'gold') standard. The results of the online tests will be compared relative to the performance on this clinical standard, in order to assess these tests' validity.

This Dutch sentence SRT test was first performed in quiet, using lists 1 and 2. The presentation level of all speech-in-noise tests needs to be approximately 20 dB higher than this SRT obtained in quiet (SRT_q), to ensure that both speech and noise are well above threshold. The noise level for all consecutive speech-in-noise tests was fixed at 65 dBA or at SRT_q + 20 dBA in cases of highly elevated SRT in quiet. Next, two lists in stationary noise were conducted. The speech stimuli recorded by a female speaker were used, and the order of the lists in noise was counterbalanced.

After finishing the Dutch sentence SRT test lists, the subject performed the three different online speech-in-noise tests. A retest was conducted, with an intermediate period of approximately 45 minutes. The sequence of test conditions was counterbalanced according to a digram-balanced Latin square, to avoid order effects. All test outcomes will be described by the term 'speech reception threshold' (SRT). These outcomes concern speech-in-noise test results, and therefore this SRT is defined as the SNR (in dB) required to correctly recognize 50% of the presented speech stimuli rather than an absolute threshold level.

Equipment and set-up

All audiometric testing and the majority of the speech-in-noise tests were carried out in a sound-insulated booth. However, no internet access was available in the audiologic booth in LUMC and online speech-in-noise measurements were carried out in a quiet room, without interfering noise[2]. Pure-tone audiometry was administered using a Decos (AMC, LUMC) or Interacoustics (UMCN) clinical audiometer and TDH-39 headphones. Calibration of hearing levels was done according to ISO-389.1 (1998).

For the speech-in-noise tests, signals were presented via a standard soundcard (Gina 24/96) on a PC at a sample frequency of 44.1 kHz and were fed through a TDT

---

[2] In the case of the measurements in LUMC in the quiet room, ambient noise levels were monitored during the experimental sessions. These levels, in 1/3 octave bands, were compared to ambient noise exposure limits defined in ISO-6189 (1983), concerning measurements of pure-tone thresholds for screening purposes down to 0 dB HL. Noise levels exceeded these limits at only one frequency, by only 5 dB. Ambient noise levels are assumed to be sufficiently reduced by the (circum-aural) HDA-headphones, and were thus considered to be of no influence on performing the supra-threshold speech-in-noise tests.

headphone buffer (HB6) and a TDT programmable attenuator (PA4). In the UMCN, signals were fed through an AC-40 clinical audiometer. To ensure standardized and controlled testing conditions, participants received the signals via headphones, at a fixed noise level. The Dutch sentence SRT test is presented trough TDH-39 headphones. After finishing this test, the subjects performed the different online speech-in-noise tests, using Sennheiser HDA-200 headphones in an otherwise identical test set-up. The participant was seated in front of a computer touch screen to enter test responses. The noise levels of each test were calibrated with a B&K sound level meter 2260 and a B&K type 4153 artificial ear, with the use of a flat-plate adaptor.

Speech-in-noise test stimuli

Dutch sentence SRT test

The Dutch sentence SRT test (Plomp & Mimpen, 1979a) comprises ten lists of thirteen short meaningful Dutch sentences containing eight or nine syllables each. These sentences are presented either in quiet or in masking noise with a spectrum that matches the long-term average spectrum of the speech. A simple up-down procedure was used to estimate the SRT. Noise level was fixed at a minimum of 65 dBA and the SNR was varied adaptively by changing the speech level. The first sentence was presented at a level below SRT that was gradually increased with 4 dB steps until the sentence was reproduced entirely correct. The level of each consecutive sentence depended on the accuracy of the response to the previous sentence; this was decreased by 2 dB after a correct response and increased by 2 dB after an incorrect response. An errorless reproduction of the entire sentence was required for a correct response. The SRT is calculated as the average SNR over sentence 5 to 13, plus the SNR of a virtual 14th sentence determined from the response to the previous sentence.

Online speech-in-noise tests

The three different online speech-in-noise tests, National Hearing Test, Earcheck and Occupational Earcheck, are based on the intelligibility of speech in stationary masking noise. For each test, the spectrum of the noise is matched to the long-term average spectrum of the speech material used and the root mean square (RMS) level of the noise was scaled to match that of the speech. Speech and noise files are stored in MP3 format and a Macromedia flash player (Macromedia Inc., San Francisco, USA) web application is used to mathematically mix the SNRs of the speech and noise files.

Normally the online tests are performed at a level that is most comfortable and loud enough for the respondent; prior to performing the test, a stimulus is presented without noise and participants are instructed to adjust volume to a level where the stimulus is clearly understood. Next, test instructions are presented on screen. Respondents are recommended to perform the test by headphones, but PC loudspeakers can also be used for diotic testing. Then the speech stimuli are presented once for familiarization of the stimuli and their corresponding buttons, and then the test starts. Immediately after finishing the test, the test result is shown; the SRT results are classified into categories of hearing ability, e.g. ‘good’, ‘insufficient’ or ‘poor’ hearing, which is presented to the respondent accompanied by appropriate advise. For reasons of time and standardization, all this is eliminated from the experimental set-up and the test is performed at a fixed presentation level, starts immediately after entering the participant’s ID code, and test results are presented as SRT values in dB.

All tests are performed according to the up-down procedure with the noise level fixed at a minimum of 65 dBA and the presentation level of speech stimuli varying adaptively with a 2 dB step size, as described above. The signal-to-noise level of the first presentation was fixed at 0 dB. The SNRs presented ranged from –14 to +4 dB. Participants were instructed to listen and enter their response using the buttons on the computer screen. Although these online speech-in-noise tests roughly follow the same test principles, there are some differences between the three versions regarding the procedure and stimuli used, which are described below and summarized in Table 4.2.

National Hearing Test

The National Hearing test (NHT) was developed by Smits et al (2004) as an automatic speech-in-noise test to be performed by telephone, or by internet (www.hoortest.nl) (Smits et al, 2006a). Dutch monosyllabic digits are used to construct a set of 80 different digit triplets. A series of 23 triplets is randomly chosen from this set, and SRT is calculated by averaging the SNRs of triplets 5 to 23 and the virtual 24th one, based on the last response. A response is considered correct only when all three digits are identified correctly. The test results are categorized into in three classes: ‘good’ hearing (SRT ≤ -5.5 dB), ‘insufficient’ hearing (-5.5 dB < SRT < -2.8 dB) and ‘poor’ hearing (SRT ≥ -2.8 dB). For reasons of comparison to the telephone version of the test, the signals of the internet version simulated the characteristics of the telephone network and were bandwidth limited to 0.3 - 3.4 kHz.

Earcheck

The Earcheck (EC; ‘Oorcheck’ in Dutch) is a screening test that specifically aims at young persons aged between 12 and 24 years (www.oorcheck.nl). The recorded test files have no bandwidth limitations using both broadband speech and noise, covering the full bandwidth up to 16 kHz. Speech material consists of nine different Dutch CVC syllables, randomly presented three times each. The words are chosen from the Dutch wordlist used for diagnostic speech audiometry (Bosman, 1989), with a phonemic distribution representative for the Dutch language. Consequently, the nine words all contain unique vowels (rat /rɑt/, thumb /dœym/, goat /xɛit/, chicken /kɪp/, fire /vyr/, lion /lew/, cat /pus/, saw /zax/, and wheel /wil/).

Table 4.2. Test characteristics of the three different online speech-in-noise screening tests.

	National Hearing Test	Earcheck	Occupational Earcheck
Speech material	Digit triplets (monosyllables)	9 CVC words	9 CVC words, paired vowels
Speaker	female	female	female
Response	forced choice from 'telephone' pad	forced choice from 9 pictures option: not understood	forced choice from 9 pictures option: not understood
No. of stimuli	23	27	35
Noise	stationary LTAS, until 3.4 kHz	stationary LTAS, full band	stationary LTAS, full band
Start SNR	0 dB	0 dB	SNR after 1st mistake
Result categories: Good	SRT ≤ -5.5	SRT ≤ -10	SRT ≤ -10
Moderate	-	-10 < SRT ≤ -7	-10 < SRT ≤ -8
Insufficient	-5.5 < SRT ≤ -2.8	-7 < SRT ≤ -4	-8 < SRT ≤ -6
Poor	SRT > -2.8	SRT > -4	-6 < SRT ≤ -4
Very poor	-	-	SRT > -4

LTAS: long-term average spectrum, of the specific speech material of each test

On screen, nine response buttons containing a written representation of the words and a corresponding picture are shown. A tenth button saying 'not recognized', is added to prevent respondents from guessing. The SRT is calculated by averaging the SNRs of presentations 8 to 27. The test results are classified in four categories of hearing status. The category 'good' hearing corresponds to a SRT of -10 dB or less, and a 'moderate' hearing status corresponds to a SRT that is between -10 dB and -7 dB. SRT results lying between -7 dB and -4 dB are categorized as 'insufficient' hearing status and a SRT result of -4 dB or higher corresponds to 'poor' hearing (Albrecht et al, 2005). The cut-off values defined to classify the SRT results into these different hearing status categories derived from a validation experiment (Martens et al, 2005).

Occupational Earcheck

The Occupational Earcheck (OEC; 'Bedrijfsoocheck' in Dutch) is a speech-in-noise screening test, developed for use in occupational hearing conservation (www.bedrijfsoorcheck.nl). The test is similar to EC in that it has no bandwidth limitations and that the speech material comprises nine Dutch CVC syllables, represented by nine response buttons and a tenth one labelled ‘not recognized’. However, the speech stimuli are different; although the words are chosen from the Dutch wordlist used for diagnostic speech audiometry as well (Bosman, 1989), the speech set is specifically selected to contain a higher proportion of high-frequency consonants and to include only five different vowels (bed /bɛt/, knife /mɛs/, bag /tɑs/, pan /pɑn/, cat /pus/, book /buk/, sock /sɔk/, sun /zɔn/, arrow /pɛil/). The words are randomly presented four times each, resulting in a total set of 35 stimuli (Ellis et al, 2006).

The first presentation is at a signal-to-noise ratio of 0 dB, and with every correct response to the subsequent stimulus speech level is attenuated by 2 dB. The actual test starts at the SNR of the first incorrect response, so the starting level is set individually. The SRT is calculated by averaging the SNRs of stimuli 6 to 35 and the test results are classified into five, smaller, categories: ‘good’ hearing (SRT ≤ -10 dB), ‘moderate’ hearing (-10 dB < SRT ≤ -8 dB), ‘insufficient’ hearing (-8 dB < SRT ≤ -6 dB), ‘poor’ hearing (-6 < SRT ≤ -4 dB) and an extra fifth category ‘very poor’ hearing (SRT > -4 dB) (Kuipers, 2007).

The performance of the three subject groups on these online tests will be compared to assess the sensitivity of the tests for NIHL detection. The validity of the online test results will be assessed by comparing the performance on the screening tests to the performance on Dutch sentence SRT test and to pure-tone thresholds. In addition, online test reliability is assessed and compared between the different tests, by analysing the test and retest results.

Results

In total, 98 subjects completed the experiment. The number of participants in each centre, demographic characteristics and outcomes of the Dutch sentence SRT test were shown in Table 4.1 for both the normal-hearing and the hearing-impaired listeners. A one-way ANOVA showed small significant differences between the NH participants of the three centres for sentence SRT_q (F[2,95] = 100.60, p < 0.001). This is probably associated with significant differences in hearing thresholds; normal-hearing subjects tested in the AMC had better threshold levels at the lower frequencies (0.125 – 1 kHz) than subjects tested in the other centres (repeated measures ANOVA for all frequencies F[1,2] = 8.16, p = 0.001).

There are no significant centre differences in the normal-hearing Dutch sentence SRT results in noise (F[2,95] = 2.47, p = 0.096). The SRTs obtained are in excellent agreement with the normal-hearing average of –5.5 dB reported by Plomp and Mimpen (1979a). Since there is no difference in sentence SRT in noise between the normal-hearing measurements in the three centres, the participants tested in the different centres may be considered as belonging to the same population. Consequently, no further distinction between centres is made and all participants are analysed together within each of the three subject groups.

Mean SRTs and reliability of the online screening tests

The three internet-based speech-in-noise tests were evaluated by performing them twice. The mean test and retest results of the three subject groups are shown in Table 4.3. Test and retest differences and correlations are analysed in terms of test reliability, which is compared across the three online tests. These analyses are conducted using data of all 98 participants together.

Table 4.3. Mean SRTs (SD) for the three subject groups, and test-retest characteristics of the three online speech-in-noise tests calculated for the total group of participants.

Test	Subject group	Test 1 *mean (SD)*	Test 2 *mean (SD)*	Mean Δ test-retest	SEM	ICC test-retest	Slope (%/dB)	Se for NIHL (%)	Sp for NIHL (%)
EC	*NH*	-12.0 (1.7)	-12.6 (1.5)	0.82 (p<0.001)	1.24	0.75	13.0	51	90
	HI-ND	-10.7 (3.0)	-11.4 (2.3)						
	HI-BD	-8.4 (2.4)	-9.6 (2.3)						
OEC	*NH*	-9.6 (1.4)	-9.7 (1.7)	0.45 (p=0.015)	1.26	0.68	11.0	92	49
	HI-ND	-7.6 (2.0)	-7.9 (2.0)						
	HI-BD	-5.9 (2.7)	-7.2 (1.7)						
NHT	*NH*	-7.5 (1.2)	-7.9 (0.8)	0.29 (p=0.097)	1.20	0.79	13.4	53	94
	HI-ND	-6.3 (1.7)	-6.7 (2.3)						
	HI-BD	-2.9 (2.5)	-2.8 (2.0)						

Displayed ICC's are Two way Mixed model, Type Consistency, Single measures ICC's. Also, sensitivity (Se) and specificity (Sp) for the detection of NIHL for each of the three online speech-in-noise tests are included.

Learning effect

First the difference in SRT over the two test sessions is calculated. This difference estimates a systematic change in performance due to learning or fatigue. Averaged across all participants, the performance on the second test was better than on the first, for all online tests. Nevertheless these differences between the SRTs for the test pairs were small, ranging from 0.29 dB (NHT) to 0.82 dB (EC). A paired-sample t-test revealed that the test-retest difference was significant for EC ($p < 0.001$) and OEC ($p = 0.015$), suggesting a small learning effect for these tests. These effects however, are of a magnitude that is smaller than inter-individual differences and thus are not particularly relevant for clinical purposes. NHT did not show a significant learning effect ($p = 0.097$).

Test-retest reliability

Reliability can be defined as the consistency of a test's results across series of observations. A measure to express this is the standard error of measurement (SEM), that is calculated dividing the within subject standard deviation of the differences by √2. A speech-in-noise test can only differentiate between subjects with different degrees of hearing loss if the SEM is small. The test-retest reliability of all three tests turns out to be comparable; SEM values ranged from 1.20 dB for NHT to 1.26 dB for OEC. In addition, test reliability can be expressed by the intraclass correlation coefficient (ICC), which is also comparable across the three tests (0.68 – 0.79). In correspondence with the slightly lower SEM, the ICC is highest for the NHT (Table 4.3).

Intelligibility functions

Another important test characteristic is the slope of its psychometric function. Steeper intelligibility functions result in more precise speech-in-noise tests, with greater discriminative power. The steepness of the performance intensity (PI) function indicates the rate at which speech information becomes intelligible with increasing signal-to-noise ratio. Although the adaptive procedure only yields SRTs, PI functions can be estimated based on a fit of the proportions correct at different presentation levels. The SNR of every presentation was corrected by the individual SRT for that test, in order to correct for inter-individual differences. Then for each SNR the proportion correct was calculated. For the different speech-in-noise tests, the intelligibility scores per corrected SNR were obtained from pooled data of both test sessions. A test-specific psychometric function was fitted to the data, using the following logistic regression function:

$$P(\mathrm{SNR}) = \gamma + (1-\gamma)\,\frac{1}{1+e^{[-(\mathrm{SNR}\;-\mathrm{SRT})\cdot 4s]}} \qquad \textit{(Equation 4.1)}$$

where p is the proportion correct at a given signal-to-noise ratio, γ is guess level, and s represents the slope of the psychometric function at SRT. In the case of EC and OEC that both use nine different stimuli guess level γ is set at 0.11. The NHT is based on correct triplet recognition and thus its guess level is around zero. The mean intelligibility functions obtained for the EC, OEC and NHT are shown in Figure 4.2, in separate plots for NH and HI listeners. There was no further distinction of the hearing impaired subjects in ND or BD groups; to assure a sufficient amount of measurements in each group, intelligibility functions are based on pooled data of all 49 HI participants.

The three online tests show similar slopes, the exact slopes are reported in Table 4.3. For all tests, slopes are slightly steeper for normal-hearing subjects than for hearing-impaired listeners. The slope of the OEC curve is somewhat shallower than the slope of the other two online tests, despite the specific test properties chosen to improve the test's precision (e.g. selection of speech material, higher number of stimuli). The PI functions of the Dutch sentence SRT test are displayed for comparison. These functions were much steeper, with slopes of 30%/dB in the normal-hearing group and 28.8%/dB in the hearing-impaired listeners.

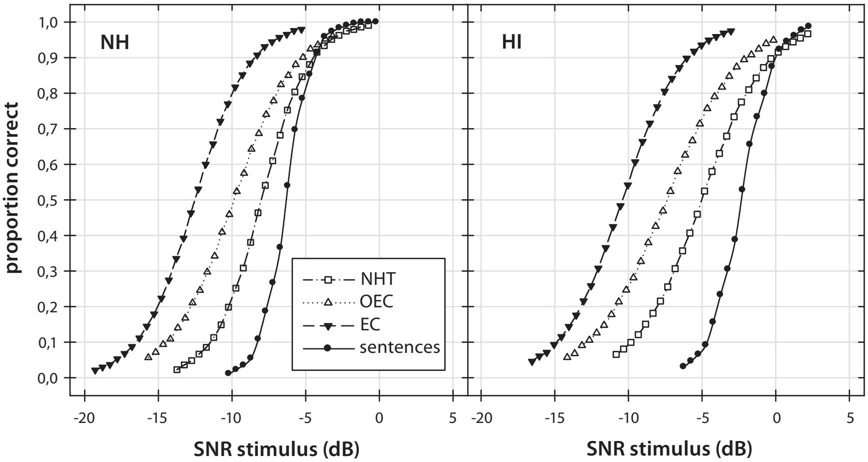


Figure 4.2. PI functions of the three different online tests and the Dutch sentence SRT test, showing proportion correct as function of the signal-to-noise ratio of the presentations. The presented SNR is relative to the individual SRT and then corrected to the average test SRT.
To assure a sufficient amount of measurements in the HI group, no further distinction of the hearing impaired subjects in ND or BD groups was made, and intelligibility functions are based on pooled data of all 49 HI participants.

Relationship between SRTs results and the pure-tone audiogram

In the following analyses only results of the first test session were taken into account, as this is representative for real-life test performance. Retest results showed a slightly lower (thus better) SRT due to training effect, and thus are not representative for respondents performing the test only once.

First, mutual Pearson correlation coefficients are calculated, using results of the total study population, in order to establish the amount of association between the different tests. All correlations are significant ($p < 0.001$), and are shown in Table 4.4.

Test validity relates to the correlation between the test's results and other, accurate, measures of the same behaviour. In order to assess the validity of the online speech-in-noise tests, their results are compared to the results of the Dutch sentence SRT test.

Table 4.4. Bivariate correlation coefficients (Pearson's r) for different speech-in-noise test outcomes with SRT results and PTA-values..

	EC	OEC	NHT	$PTA_{0.5,1,2,4}$	$PTA_{2,4}$	$PTA_{3,4,6}$
EC	-	0.61	0.60	0.66	0.64	0.62
OEC	0.61	-	0.64	0.69	0.67	0.66
NHT	0.60	0.64	-	0.72	0.74	0.69
Sentence SRT	0.65	0.76	0.77	0.82	0.82	0.80

All online tests are significantly correlated with this test ($p < 0.001$). The correlation coefficients show moderate association (Table 4.4), and are quite similar for the three tests, ranging from 0.65 (EC) to 0.77 (NHT). In Figure 4.3, bivariate scatterplots are given, presenting the 50% points of the three online screening tests versus results of the Dutch sentence SRT test. In each panel, datum points of the listeners with normal-hearing are in closer proximity to one another, whereas the datum points from the listeners with hearing loss show a wider distribution, reflecting a larger inter-subject variability. Excluding all data of the normal-hearing participants yielded correlation coefficients of the same magnitude as those presented in Table 4.4.

Correlations of the online test results and pure-tone thresholds are analysed in the same way. For this purpose, three pure-tone averages (PTA) are calculated, concerning the average over the frequencies 0.5, 1, 2 and 4 kHz that are important for speech intelligibility ($PTA_{0.5,1,2,4}$); the average of the thresholds at 2 and 4 kHz, which yielded the strongest correlation with results of the Dutch sentence SRT test in studies of Smoorenburg (1990; 1992) ($PTA_{2,4}$); and the PTA of the higher, noise-sensitive, frequencies 3, 4 and 6 kHz ($PTA_{3,4,6}$). Correlation coefficients between these pure-tone averages and online test results are statistically significant ($p < 0.001$) yet lower than the correlation coefficients found comparing PTA-values to the Dutch sentence SRT (Table 4.4). The correlation coefficients for the three tests do not differ much, but r is highest for NHT and lowest for EC. Correlation is highest when SRTs are compared to $PTA_{0.5,1,2,4}$ and becomes slightly lower when online test results are compared to $PTA_{3,4,6}$. In Figure 4.3, also bivariate scatterplots of the SRT results of the three online tests versus $PTA_{0.5,1,2,4}$ are displayed.

Group differences between tests

Finally, the results for NH listeners and both groups of HI participants with different degrees of NIHL are compared. This is done in order to establish the amount of separation in recognition performance measured with each test and to obtain the tests' sensitivity to identify subjects with NIHL. Again, only the speech reception thresholds of the first test were used for this evaluation.

A one-way ANOVA of the individual SRT results of each testis performed to investigate differences between the subject groups. This shows that the main effect of listener group is significantly different (EC: $F[2,95] = 20.56$, $p < 0.001$, OEC: $F[2,95] = 32.04$, $p < 0.001$, NHT: $F[2,95] = 60.01$, $p < 0.001$). Post-hoc t-tests with Bonferroni corrections for multiple comparisons show small but significant differences between all three subject groups, except for the EC; normal-hearing EC results do not differ from EC results of the hearing-impaired subjects with a narrow dip ($p = 0.079$).

Boxplots representing the obtained SRT results for each subject group are displayed in Figure 4.4. The boxes represent the inter-quartile range, the median is demonstrated by the vertical line inside the box and error bars represent the 5th and 95th percentile of the results in each group. The dashed lines in Figure 4.4 correspond to hearing status categories as defined for the different tests. The overlap between normal-hearing listeners and subjects with narrow audiometric dips in the EC results is evident, since a high proportion of the mildly impaired participants are classified as having 'good' hearing. Although the group differences are statistically significant, NHT results are distributed similarly. OEC displays a somewhat better separation between the normal-hearing and hearing-impaired listeners. This test correctly classified the majority of the hearing-impaired listeners, while only half of the normal-hearing subjects fall into the 'good' hearing category.

Sensitivity and specificity for NIHL

In order to investigate how well the different speech-in-noise tests discriminate between NH and HI respondents, SRT results and the cut-off values defined for each test were used to calculate their sensitivity and specificity. Test sensitivity refers to the percentage of HI participants classified correctly as having hearing ability worse than

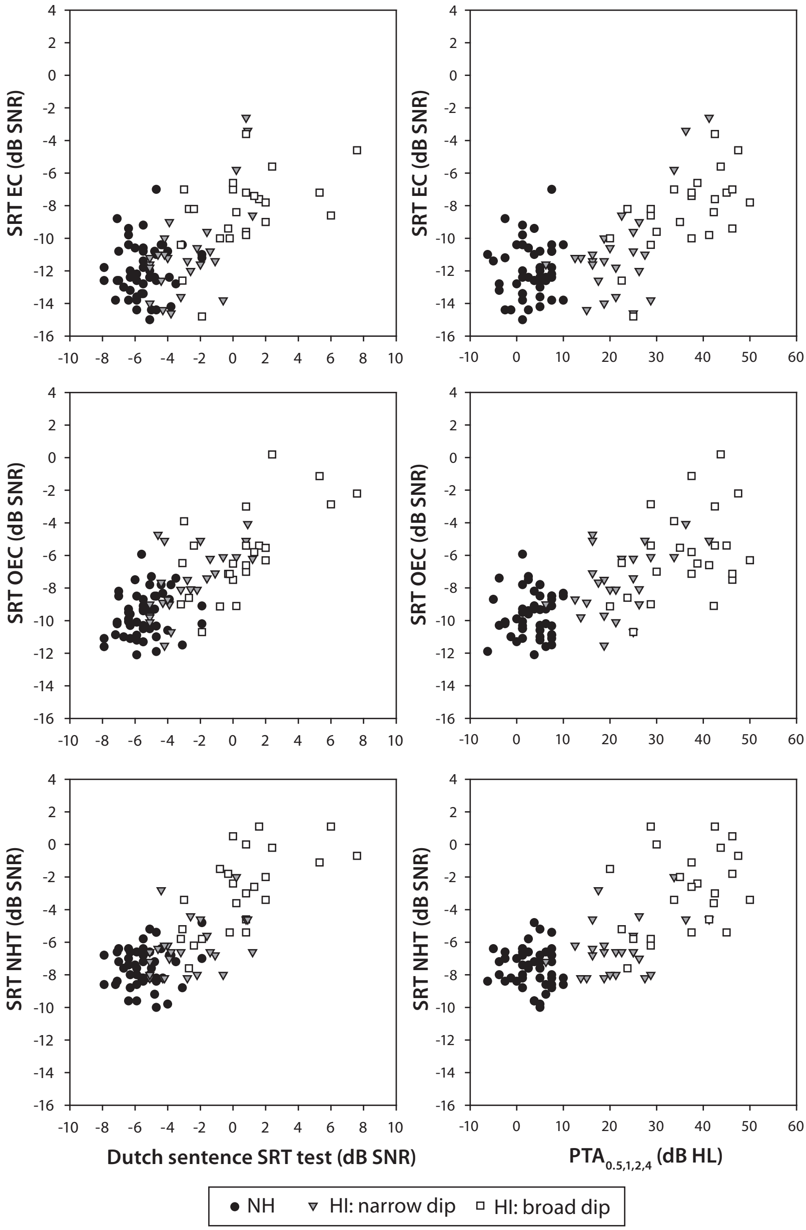


Figure 4.3. Bivariate scatterplots of SRT measured with each of the online speech-in-noise tests against sentence SRT in noise and PTA$_{0.5,1,2,4}$, separated for the three subject groups; NH listeners (black), HI subjects with narrow noise dip (grey) and HI subjects with broad noise dips (white).

‘good’; test specificity refers to the number of normal-hearing subjects correctly classified as having ‘good’ hearing. The results for each of the three screening tests are shown in Table 4.3. The highest test sensitivity for NIHL was found for OEC, almost all HI participants were classified as ‘moderate’, ‘insufficient’ or ‘poor’. Yet, this accounted for the majority of the normal-hearing group as well (Figure 4.4), resulting in a very low specificity of 49%, meaning that half of the NH subjects is also classified as being hearing impaired. The two other tests show, in spite of a high specificity of 94% for NHT and 90 % for EC, low sensitivity values of 55% and 51% respectively. This means that a large number of the subjects with NIHL, especially those with narrow dips, will be classified as being normal-hearing when performing NHT or EC.

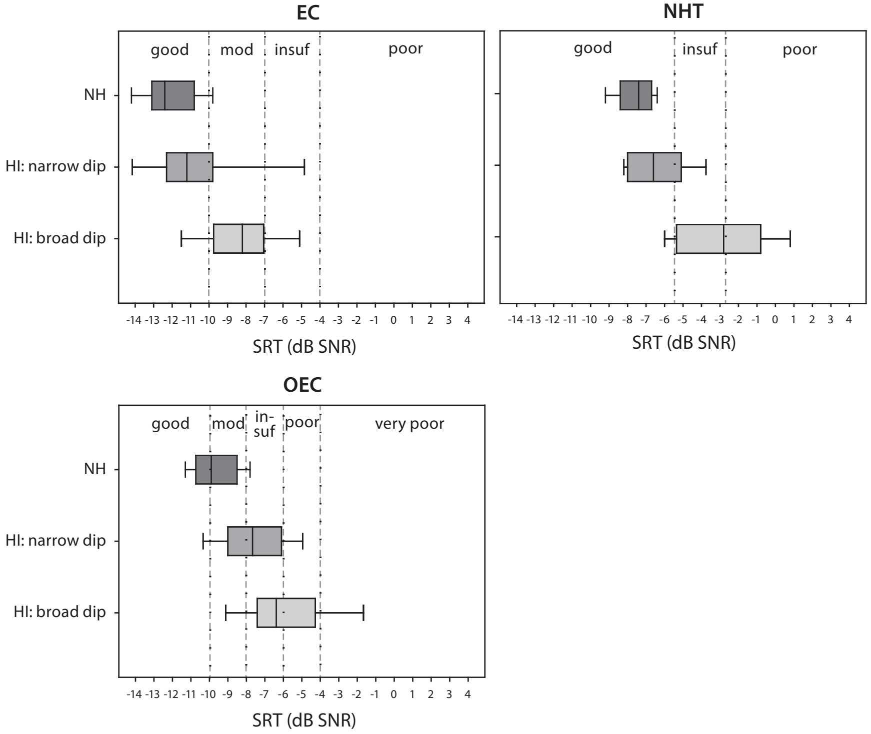


Figure 4.4. Boxplots of SRTs of the different online speech-in-noise tests, for NH listeners (dark grey), HI listeners with a narrow dip (gray) and HI listeners with a broad dip (light grey). Error bars show 5th and 95th percentile of SRT values. Vertical dashed lines indicate the outcome categories of each test, ranging from good, on the left, through moderate (mod) and insufficient (insuf), to (very) poor, on the right of each plot.

Discussion

Noise-induced hearing loss is a highly prevalent public-health problem that is irreversible yet preventable. Increased public awareness about NIHL and early detection of hearing loss in groups at risk can help to prevent the development of NIHL. Since monitoring by pure-tone audiometry, as applied in an occupational setting, is not easily accessible for a broad population, alternative ways of monitoring, e.g. speech-in-noise testing, appear attractive, especially when such tests can be conducted through the internet. A speech-in-noise test has the advantage that it is independent of presentation level and less sensitive to background noise, making it suitable for use as a self-test at a remote test site (e.g. at home), in less well-controlled conditions.

An existing Dutch online screening test, the National Hearing Test (Smits et al, 2004; 2006a), has been adapted to serve as a screening instrument in either a population of occupationally exposed workers (Occupational Earcheck) or a young population exposed to leisure noise (Earcheck). In both populations noise-induced hearing loss is expected to be prevalent, and the screening tests could be of great significance in the identification of this NIHL. However, this study was conducted to investigate the value of these tests in discovering early noise-induced hearing loss and shows that their sensitivity to detect NIHL is only low.

Age effects

The normal-hearing and hearing-impaired subjects tested in this study are not well matched with respect to their age; the HI listeners were on average older than the NH students. Since speech-in-noise recognition becomes poorer with increasing age, especially when age is above 50 years (Plomp & Mimpen, 1979b), the two subject groups differ in more ways than just hearing status. The HI subjects measured in this study had a mean age of 56 years that, according to the data presented by Plomp & Mimpen (1979b), corresponds to a median SRT of only 1 dB poorer than the median SRT for the young NH subjects. Moreover, Van Rooij & Plomp (1992) showed that speech recognition in the elderly was influenced mainly by the progressive high-frequency hearing loss that develops with age, rather than by a decrement in cognitive performance. This made us assume that the difference in age did not strongly affect our results, and this assumption is supported by the fact that the differences in SRT between the subject groups are actually smaller than expected despite the older HI subjects.

Test reliability

The online speech-in-noise screening tests did show reliable results in this population of normal-hearing listeners and patients with NIHL. The test-retest reliability is mainly determined by the slope of the psychometric function; steeper slopes mean that a small change in SNR would result in a large change in performance. The slopes of intelligibility functions of both EC (13.0%/dB) and NHT (13.4%/dB) turned out to be comparable and slightly higher than the slope obtained for OEC (11.0%/dB). These slopes are considerably steeper than the slopes of the CVC intelligibility functions reported by Bosman & Smoorenburg (1995), who found a slope of 8.1%/dB for normal-hearing listeners and a similar slope of 7.8%/dB for a group of 20 subjects with NIHL, probably because their test used an open response format. However, slopes reported for similar speech-in-noise screening tests are somewhat steeper; Smits et al. (2004) obtained psychometric functions of NHT results of 10 normal-hearing subjects with slopes of 16%/dB using headphones and 20%/dB using telephone. The French digit triplet test by telephone also reached a higher slope, of 17.1%/dB (Jansen et al, 2010).

Comparing the test and retest results yielded high intraclass correlation coefficients of 0.68 – 0.79 and standard errors of measurement that are relatively small: around 1.2 dB. The reliability of the National Hearing Test is examined earlier by a series of studies (Smits et al, 2004; Smits & Houtgast, 2005; Smits & Houtgast, 2007). These investigations reported measurement errors ranging from 0.9 to 1.1 dB obtained in a normal-hearing population, which are slightly better than the SEM found in this study.

In addition, a significant test-retest difference was obtained for OEC and EC, suggesting a small learning effect. However, these effects were only 0.5 dB and 0.8 dB respectively and since this is smaller than the measurement error they can be considered as not clinically relevant. These results indicate that the three tests are reliable in measuring speech reception in noise. Reliability measures showed comparable results over the three online tests, but all results turned out to be slightly poorer for OEC. This is unexpected, since this test was specifically designed to have a higher precision.

Test validity

Previous studies reported moderately high correlations coefficients of 0.72 – 0.84 between results of the Dutch sentence SRT test and pure-tone averages for normal-hearing and hearing-impaired listeners (Smoorenburg, 1992; Bosman & Smoorenburg, 1995; Smits et al, 2004). The findings of this study show similar correlations; r is 0.82 comparing the results of the Dutch sentence SRT test with both $PTA_{0.5,1,2,4}$ and $PTA_{2,4}$. The online test results showed lower correlation coefficients ranging from 0.66 for EC results to 0.72 for NHT results when compared to $PTA_{0.5,1,2,4}$. This value for NHT is comparable to the correlation of 0.77 reported for both the French and the Dutch digit triplet in noise test and $PTA_{0.5,1,2,4}$ (Jansen et al, 2010; Smits et al, 2004). The latter one also found a correlation for NHT with the Dutch sentence SRT test of 0.87, that was higher than found in this study (r = 0.77). All correlations found between the SRTs and $PTA_{3,4,6}$ are slightly lower than found for the other PTA-values.

Results of the online tests

The online tests thus prove to yield reliable results, but are not strongly related to pure-tone thresholds and the Dutch sentence SRT results in this population. More importantly, the results of this evaluation study show that the speech-in-noise tests differentiated only to a limited degree between speech intelligibility of participants with noise-induced hearing loss and normal hearing. Normal-hearing participants do reach – on average - lower SRTs than hearing-impaired listeners but the differences between normal speech reception and SRTs of listeners with a narrow audiometric dip are only small or, in case of EC, even statistically insignificant. Consequently, the sensitivity of the tests to discover relatively mild high-frequency hearing losses appears to be rather low. This is reflected in the sensitivity and specificity results; NHT and EC yield a sensitivity of 55% and 51% respectively in this study population, meaning that both tests classify almost half of the listeners with NIHL incorrectly as normal-hearing listeners. On the other hand, the specificity of these two tests is high; 94% and 90% respectively. This is in contrast with the OEC that yield a high test sensitivity of 92% in this population but shows low specificity of 49%, incorrectly classifying the majority of the normal-hearing listeners. Apparently the additional fifth category had led to cut-off values that are too high to classify normal-hearing listeners as having a ‘good’ hearing.

The sensitivity and specificity of a test depend critically on the cut-off value that is used to distinguish between normal and impaired (hearing) performance. A change in the cut-off values used in EC and NHT will lead to a higher number of correctly classified hearing-impaired subjects, but this will go along with poorer specificity.

Current findings thus show that subjects with NIHL, especially with a narrow audiometric dip, will perform quite similar to normal-hearing participants on internet-based speech-in-noise tests, probably because they can benefit from their preserved hearing in the low and mid frequencies. This implies that the current tests for auditory screening and monitoring via internet, although proven to be a valuable tool for screening purposes in a general population, have a limited applicability for populations in which noise-induced hearing loss is prevalent, e.g. in occupational health care and in prevention programs for young persons. In fact, these results do not only concern respondents with NIHL, but may also be generalized to individuals with a high-frequency sensorineural hearing loss, regardless of the etiology.

Future research

Although low sensitivity for NIHL was expected, this is a disadvantageous finding, since early self-identification of hearing loss may result in increased awareness and appropriate audiological follow-up for those affected, thereby preventing NIHL. It would be very useful to investigate ways to improve sensitivity of these tests to discover high-frequency hearing loss. In order to do so, possible adaptations involving either the speech material or the masking noise used, might be considered. For the speech stimuli, words seems to be the best choice, since a self-test should be quick, easy and automated. We hypothesized that listeners with NIHL mainly rely on vowel recognition for their identification of CVC words in noise, especially in the small closed sets of stimuli that are presented. If this is true, one would expect a greater deviation between EC and OEC results, as the latter presents words containing high frequency consonants and only five different vowels, increasing the auditory similarity of the stimuli and forcing the listener to use high-frequency information to discriminate between the words. Although the difference in SRT results of OEC between NH listeners and HI subjects with a narrow dip is slightly greater than found for the EC, there is no overall difference between these tests.

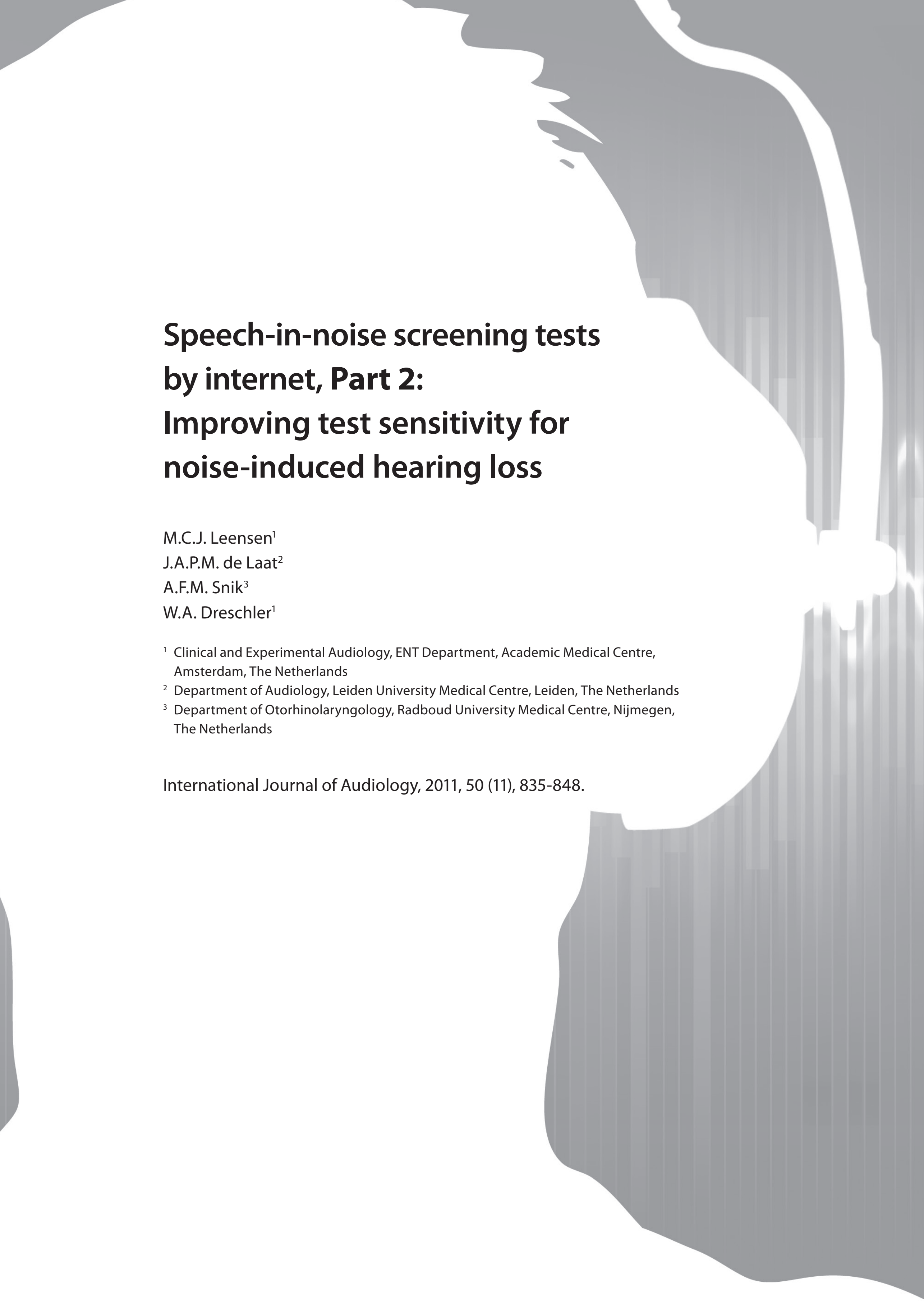
In addition, the use of different types of noise, such as filtered or interrupted noise, may discriminate better between normal and impaired performance, increasing the validity of the tests for the detection of NIHL. This will be considered in a follow-up study, in which we aim to make the tests more appropriate for screening on (early) NIHL, described in Chapter 5.

Conclusion

In this study three online Dutch speech-in-noise screening tests were evaluated in a population of normal-hearing listeners and participants with different degrees of noise-induced hearing loss, concerning their sensitivity for detecting NIHL. The tests showed reliable results, although correlations with Dutch sentence SRTs and pure-tone thresholds were moderate. SRT results of subjects with mild NIHL deviated only slightly from normal performance. Consequently, the sensitivity of the tests to discover high-frequency hearing loss is low, and the tests in their current form are not appropriate to be used for screening of NIHL in an early stage.

Since these online screening tests can play an important role in the prevention of NIHL possible adaptations leading to an improvement in test sensitivity for NIHL are to be investigated, in order to obtain a valid screening tool for high-frequency hearing loss detection. This will be described in Chapter 5.

5



Speech-in-noise screening tests by internet, Part 2: Improving test sensitivity for noise-induced hearing loss

M.C.J. Leensen[1]
J.A.P.M. de Laat[2]
A.F.M. Snik[3]
W.A. Dreschler[1]

[1] Clinical and Experimental Audiology, ENT Department, Academic Medical Centre, Amsterdam, The Netherlands
[2] Department of Audiology, Leiden University Medical Centre, Leiden, The Netherlands
[3] Department of Otorhinolaryngology, Radboud University Medical Centre, Nijmegen, The Netherlands

International Journal of Audiology, 2011, 50 (11), 835-848.

Abstract

Objective: An easily accessible screening test can be valuable in the prevention of noise-induced hearing loss (NIHL). The Dutch National Hearing Foundation developed 'Earcheck'; an internet-based speech-in-noise test, presenting CVC words in stationary broadband noise. However, its sensitivity to detect NIHL appeared to be low, 51% (Chapter 4). The aim of the current study is to examine ways to improve Earcheck's sensitivity for (early) NIHL using different forms of noise filtering.

Design: The test's stationary broadband masking noise is replaced by six alternatives, including noises that have been temporally modulated, spectrally filtered by high-pass or low-pass filters, and combinations of temporal modulation and spectral filtering.

Study Sample: In this multi-centre study, 49 normal-hearing and 49 subjects with different degrees of NIHL participated.

Results: Hearing-impaired subjects deviated more clearly from normal performance when executing the test with alternative masking noises, except for the high-pass filtered conditions. Earcheck with low-pass filtered noise made the best distinction between normal hearing and NIHL, without reducing test reliability. The use of this noise condition improved the sensitivity of Earcheck to 95%.

Conclusion: The use of low-pass filtered masking noise makes speech-in-noise tests more sensitive to detect NIHL in an early stage.

Introduction

Despite the fact that noise-induced hearing loss (NIHL) is preventable, it is still a highly prevalent public-health problem in modern society. NIHL is not only the most reported occupational disease in the Netherlands (Van der Molen, 2010), but it is also a growing concern in the general public, due to the increasing exposure to recreational noise. Young people especially are considered to be at risk of developing NIHL, exposing themselves to potentially damaging loud music when attending discotheques and live concerts or when listening to personal music players. Noise levels during these recreational activities are high and often exceed the occupational limit of 80 dBA set for an 8-hour working day, defined in the European Directive 2003/10/EC (EPC, 2003). Vogel et al. (2010) estimated that more than half of the 1512 adolescents participating in their study exceeded this occupational standard by listening to high volume music. Although research concerning the prevalence of hearing loss caused by leisure noise in youngsters demonstrated inconsistent results (Meyer-Bisch, 1996; Mostafapour et al, 1998, Niskar et al, 2001; Biassioni et al, 2005; Shah et al, 2009; Zhao et al, 2010; Shargorodsky et al, 2010), the reported average sound levels of these activities, ranging from 80 dBA to 115 dBA (SCENIHR, 2008), are high enough to pose a risk to hearing. This is particularly true for individuals being exposed for longer periods, and for young people involved in multiple noisy recreational activities or additionally exposed to occupational noise, resulting in cumulative effects that may lead to an increased prevalence of hearing loss (Torre III, 2008).

Noise-induced hearing loss develops gradually and is often unnoticed until the damage is substantial and severe enough to be measured (Shah et al, 2009). Therefore, the risk of hearing loss is easily underestimated (Vogel et al, 2008). Furthermore there is a great deal of misconception and unawareness among youngsters about the impact of hearing loss and the effect of overexposure to loud music in general (Chung et al, 2005; Vogel et al, 2008; Vogel et al, 2009; Shah et al, 2009). Adolescents first must become aware that listening to high-volume music may cause hearing damage and that they personally are at risk for hearing loss before the promotion of protective behaviours is useful (Vogel et al, 2009). In addition, self-experienced symptoms after recreational noise exposure might lead to greater awareness (Widen et al, 2009), which can change personal listening behaviour in order to protect hearing. Moreover, if hearing deterioration can be shown at an early stage actions can be taken to prevent further hearing loss (Meyer-Bish, 1996).

An objective hearing screening test that can detect hearing loss in an earlier stage can be of great help in preventing NIHL and raises awareness of possible hearing problems after music exposure (Koopman et al, 2008). Since subjects with NIHL often complain about a reduced ability to understand speech in noisy situations, a speech-in-noise test seems a suitable measure to detect this kind of hearing loss.

In The Netherlands an internet-based speech-in-noise test was implemented, as a screening tool for adolescents exposed to leisure noise. This test, “Earcheck” (Oorcheck in Dutch, www.oorcheck.nl), has been developed by the Dutch National Hearing Foundation and the LUMC Leiden (Albrecht et al, 2005). The test principles are derived from the National Hearing Test (Smits et al, 2004; 2006a). This test is bandwidth limited, whereas Earcheck covers the full bandwidth up to 16 kHz. Earcheck is incorporated in a special educational website aiming at adolescents and young adults in the age range 12 - 24 years, facilitating early NIHL identification and increased awareness about the risks of noise exposure.

The test presents a closed set of nine different CVC words against a background of stationary speech-shaped noise. The test uses an adaptive up-down procedure corresponding to the one described by Plomp and Mimpen (1979a), to asses the speech reception threshold (SRT), i.e. the signal-to-noise ratio (SNR) required to recognize 50% of the speech correctly. A total of 27 stimuli are randomly presented, and the arithmetic average of the SNRs of the last 20 presentations results in the SRT. The Earcheck outcomes are classified into four categories of hearing status, accompanied by an appropriate advice for referral. This self-screening test is easy to administer and takes about three minutes to perform.

Speech-in-noise tests, such as Earcheck measure the speech reception threshold in a stationary noise with the same long-term average spectrum as the speech material used. Because this makes the test independent of absolute presentation level (Plomp, 1986) and of variations in equipment used (Smits et al, 2004; Culling et al, 2005), it is considered to be suitable for online screening purposes. Furthermore, the test is robust against background noise (Jansen et al, 2010), resulting in a test that is reliably administered in an at-home setting (Smits et al, 2004). However, the evaluation study described in Chapter 4 showed that the currently implemented Earcheck was not able to make a clear distinction between normal-hearing listeners and participants with different degrees of NIHL. Although Earcheck demonstrated fairly good test reliability, with a test-retest standard deviation (SD) of 1.2 dB and an intraclass correlation coefficient of 0.75, the test sensitivity for NIHL turned out to be rather low; only 51% compared to the results of the clinical audiogram. This means that half of the NIHL patients were (wrongly) classified as having normal hearing by Earcheck.

Subjects with NIHL exhibit poorer hearing thresholds in the higher frequencies, while thresholds in the lower frequency region remain (nearly) normal. These individuals could be benefitting from their intact low frequency hearing (Quist-Hanssen et al, 1979), by mainly relying on vowel recognition to identify CVC words in noise (Smoorenburg, 1992), especially if a closed set of stimuli is used. Consequently, Earcheck only demonstrated small differences between normal speech reception and the SRTs of listeners with early NIHL, resulting in low sensitivity to discover relatively mild high-frequency hearing losses (see Chapter 4). Since an adequate screening test can be of major importance in the prevention of NIHL, current study examines possible ways to improve Earcheck's sensitivity to discover (early) high-frequency hearing loss.

The sensitivity of a test is high when it clearly distinguishes between normal-hearing and hearing-impaired listeners (Theunissen et al, 2009). The difference in speech reception between normal-hearing and hearing-impaired listeners varies greatly depending on the nature of the interfering noise. Certain types of maskers may yield more information than a steady-state background. It is well known that listeners with normal hearing sensitivity perform much better when the masking noise is interrupted than when it is stationary (Festen & Plomp, 1990; Phillips et al, 1994; Stuart & Phillips, 1996; Bacon et al, 1998). They take advantage of the relatively high SNR in the silent periods of the interfering noise to extract speech information, in order to achieve higher performance than with a stationary masker. This is called masking release. Conversely, hearing-impaired listeners experience little or no benefit when going from stationary noise to fluctuating noise, even when the hearing loss is mild and more or less restricted to high audiometric frequencies (Phillips et al, 1994; Stuart & Phillips, 1996; Bacon et al, 1998; Versfeld & Dreschler, 2002). The SNR improvement during the gaps in the noise is limited by their elevated thresholds. In addition, they generally show reduced temporal resolution and degraded recovery from forward masking, preventing them from taking full advantage of dips in the masking noise. For sentence intelligibility, reported differences between normal-hearing and hearing-impaired listeners in interrupted noise are in the range of 7 to 15 dB compared to differences ranging from 2 to-5 dB in stationary noise (Peters et al, 1998). Also previous studies examining word recognition of subjects with NIHL in stationary and interrupted noise only demonstrated significant differences in performance relative to controls in interrupted noise conditions (Phillips et al, 1994; Stuart & Phillips, 1996). Spectral properties of the speech signal and the competing background noise affect the results of a speech-in-noise test as well. Normal-hearing speech reception in noise improved when spectral dips were added to the interfering noise, and this improvement increased as the width of these spectral dips increased (Peters et al, 1998). Hearing-impaired subjects showed a much smaller improvement, indicating reduced audibility for speech in noise with lower intensity. A spectrally filtered noise can also be used to improve discrimination between respondents with NIHL and normal-hearing listeners. Since NIHL affects the higher frequency region, a low-pass filtered masker would facilitate the use of high-frequency speech information, where limitations imposed by reduced audibility will impair speech intelligibility.

Considering the expected larger differences in SRT between hearing-impaired and normal-hearing listeners in time-modulated or spectrally filtered maskers compared to stationary noise (Festen & Plomp, 1990; Peters et al, 1998), SRT measurements in a modified noise could improve discrimination between hearing-impaired and normal performance, providing a more sensitive measure of hearing impairment. However, this only applies when the reliability of the tests using the modified masking noises remain unchanged or is at least equivalent to that of the original version (Smits & Houtgast, 2007).

The aim of this study is to improve an online speech-in-noise screening test for (early) NIHL. In order to do so, different forms of masking noise modification are investigated, by comparing the speech recognition performance of normal-hearing listeners and hearing-impaired participants with noise-induced hearing loss in these different noise conditions. In addition, the alternative test needs to be reliable and valid, so test-retest results are evaluated and performance on the different speech-in-noise tests is compared to performance on the Dutch sentence SRT test (Plomp & Mimpen, 1979a) and to pure-tone thresholds.

Methods

Participants

The same groups of normal-hearing and hearing-impaired listeners as described in Chapter 4 participated in the current study. Participants were tested at three different audiology departments; LUMC Leiden, UMCN St Radboud Nijmegen and AMC Amsterdam. There were no differences between the subjects tested at the different centres. All subjects were native speakers of the Dutch language.

The normal-hearing (NH) group consisted of 49 listeners (mean age 27.0 years, SD = 8.5 years; 16 male, 33 female), with pure-tone thresholds of 15 dB HL or better across octave frequencies from 0.125 to 8 kHz, including 3 and 6 kHz. The 49 hearing-impaired subjects (mean age 56.3 years, SD = 9.4 years; 47 male, 2 female) were patients of one of the three ENT departments who had recently received audiological evaluations. The inclusion criterion was a combination of one or more pure-tone thresholds greater than 25 dB HL at 2 to 6 kHz and thresholds of 20 dB HL or better at 0.125 to 1 kHz. Also the included subjects had a history of noise exposure, although it is impossible to prove a direct relationship between this exposure to noise and the hearing loss measured. Even though the exact cause of notch-shaped hearing loss remains unknown, the included audiogram configurations are characteristic for NIHL and the results are assumed to be applicable to a NIHL population. In addition, results may be generalized to individuals with high-frequency sensorineural hearing loss due to another cause. Patients with an air-bone gap greater than 15 dB in the tested ear were excluded.

The hearing-impaired participants were divided into subgroups having either a narrow audiometric dip (HI-ND, n = 25, mean age 54.1 years, SD = 7.6 years), corresponding with early NIHL, or a broad dip (HI-BD, n = 24, mean age 58.5 years, SD = 10.2 years), corresponding with more severe hearing loss. Distinction was made based on whether or not their hearing threshold at 2 kHz was affected; when hearing threshold at 2 kHz exceeded the pure-tone average of 0.5 and 1 kHz by more than 15 dB, the patient was classified as having a broad dip. For each group, mean audiometric hearing thresholds of the ears selected for monaural testing are displayed in Figure 5.1.

A power analysis showed that the sample size of 49 in each group will have 84% power to detect a difference in means of 2.0 dB, using a two group t-test with a 0.05 two-sided significance level and assuming a standard deviation of 3.1 dB (Jongmans et al, 2008). Details on power calculation and demographics are reported in Chapter 4. All participants signed informed consent forms before starting the experiment. This study was approved by the ethics committee of the University of Amsterdam.

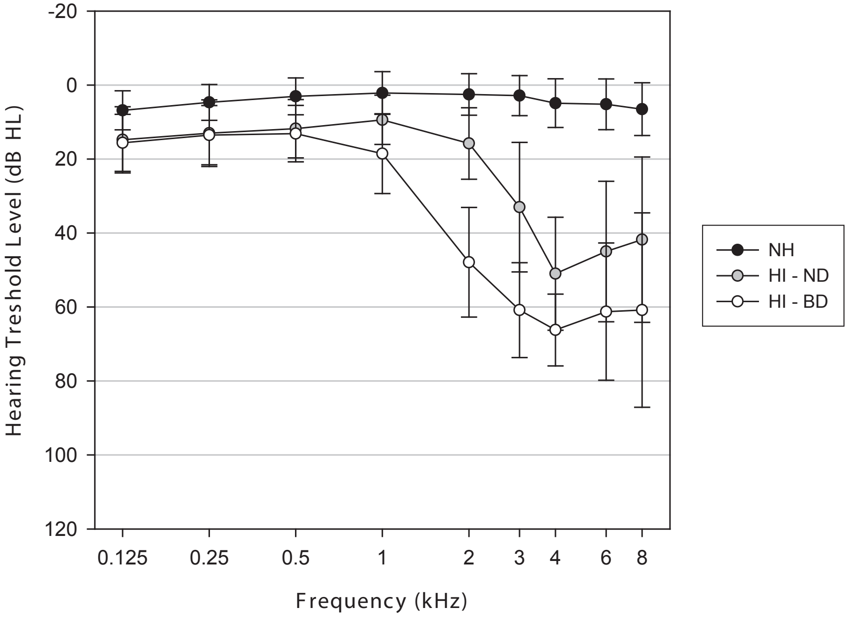


Figure 5.1. Audiometric thresholds of the ear selected for monaural testing, averaged for each of the three subject groups. Error bars represent one SD.

Speech-in-noise tests

Dutch sentence SRT test

Plomp and Mimpen (1979a) developed a speech-in-noise test that consists of 10 lists of 13 short everyday sentences, spoken by a female speaker and presented in a stationary interfering noise with the long term average speech spectrum. This test is considered as the clinical standard in the Netherlands, and the SRT obtained with this test was used as a reference value in this study to which performance on the alternatives of Earcheck is compared. Noise level was fixed and the SRT was measured adaptively according to the standard procedure of Plomp and Mimpen (1979a).

Earcheck

Earcheck (EC) is an online speech-in-noise screening test based on the intelligibility of nine different Dutch CVC words in stationary masking noise. These words were randomly presented three times each. On screen, nine response buttons containing a written representation of the words and a corresponding picture were shown. A tenth button saying ‘not recognized’, was added to prevent respondents from guessing, when the presented stimulus was not understood. Participants were instructed to listen carefully and enter their response using the buttons on the computer screen. The test was performed according to a simple up-down adaptive SRT-procedure with a 2 dB step size and fixed noise level. After an incorrect response, the signal-to-noise ratio (SNR) of the next presentation is increased by 2 dB and after a correct response SNR is decreased by 2 dB. The SRT was calculated as the average SNR over stimuli 7 to 27, and was defined as ‘good’ (SRT ≤ -10 dB), ‘moderate’ (-10 < SRT ≤ -7 dB), ‘insufficient’ (-7 < SRT ≤ -4 dB) and ‘poor’ hearing (SRT > -4 dB) (Albrecht et al, 2005).

All test results will be described by the term ‘speech reception threshold’ (SRT). For the purpose of this study, SRT is defined as the signal-to-noise ratio (in dB) that yields 50% intelligibility, rather than as absolute threshold level.

Stimuli

Speech

The speech material used was the closed set of nine different monosyllables comprising the speech stimuli of Earcheck. These CVC words were chosen from the Dutch wordlist used for diagnostic speech audiometry (Bosman, 1989), with a phonemic distribution representative for the Dutch language (Albrecht et al, 2005). Consequently, the nine words all contained unique vowels (thumb /dœym/, goat / γεit/, chicken /kIp/, rat /rαt/, fire /vyr/, lion /lew/, cat /pus/, saw /zaχ/ wheel / w¡l /).

Homogenizing the speech material

When using an adaptive procedure to assess speech intelligibility, it is important that the speech stimuli are of equal difficulty when heard in noise, to yield consistent and accurate results. One way to achieve this was to adjust the words in level with respect to an optimized perceptual homogeneity. These level corrections were derived from word-specific intelligibility functions, determined based on online test results of previously performed tests.

Earcheck was implemented online in April 2004. Test results were centrally stored, and data collection until December 2007 was available to determine word-specific intelligibility functions. Tests that resulted in a within-subject standard deviation of more than 2.5 dB were considered unreliable and were excluded (Martens et al, 2005). This resulted in a dataset of approximately 100,000 test results that were available for these analyses. Since the SRT measured with Earcheck is calculated by averaging the SNR of presentation 8 to 27, only these presentations were selected. Each word was presented at various signal-to-noise ratios during the adaptive procedure. In order to compensate for inter-individual differences in overall performance, relative SNRs were constructed by correcting the presentation level for the individual SRT. Since it was known whether the response at that SNR was correct or incorrect, the proportion correct could be calculated for each word at each relative SNR. Based on a fit of these proportions correct word-specific psychometric functions were estimated, using the following logistic regression function;

$$P(SNR) = \gamma + (1 - \gamma)\,\frac{1}{1 + e^{[-(SNR\ - SRT)\cdot 4s]}} \qquad \textit{(Equation 5.1)}$$

Where *P* is the proportion correct at a given relative signal-to-noise ratio, γ is guess level, and *s* represents the slope of the psychometric function at SRT. When using a closed-set of speech stimuli the guess rate is related to the number of alternatives (1/n), thus in this case γ is 0.11.

The relative SNRs at the 50% points for each intelligibility function resulting from this fitting procedure were used to adjust the RMS level of the particular word in order to achieve equal intelligibility. These level corrections were applied to the individual CVC words, meaning that the resulting dB-level of each word differs. To define the SNR in the measurements, the average speech level, i.e. the average level of all word-specific dB-levels, was used.

Masking noise

First a broadband stationary masking noise was constructed with a spectral shape similar to the long-term average spectrum of the homogenized word material. Then an experimental set of interfering noises was created by modulating and/or filtering this speech-shaped noise.

Speech-shaped noise

A stationary speech-shaped noise was generated by filtering a white noise, using a FIR filter. This filter was based on the long-term average speech spectrum of the concatenation of all test words, according to the methods described by Versfeld et al. (2000). The filtered noise was scaled to match the level of the speech material. This provides a reference condition against which SRTs in other types of noise can be compared.

Alternative noises

An experimental set of alternative Earcheck versions was created. In these tests, the homogenized words were presented in different background conditions, all of which were derived from the speech-shaped noise matching the long-term spectrum of the speech stimuli. Six different masking noises were created, either by spectrally filtering or temporally modulating this speech-shaped noise, or by a combination of this filtering and modulating. The appropriate parameters for the spectrally filtered noise conditions were determined using the speech intelligibility index (SII) according to ANSI S3.5 (1997). The SII model can predict the audibility of speech by calculating the proportion of total speech information that is available to the listener, as function of the SNR of the presentation and listeners' hearing threshold level.

The SII was calculated for several audiograms ranging from normal-hearing to severe noise-induced hearing loss and for various versions of Earcheck. These model-based predictions provided insight into the effects of different kinds of filtered masking noise on the SNR required for correct speech reception. Relevant parameters of filtered noise conditions, such as cut-off frequency, noise floor and filter shape (HP/LP/notch), are varied in order to predict their effects on the SRT. The noises that generated the largest differences between normal and impaired hearing ability and that resulted in a SRT that can be reliably measured at a remote test site, were chosen for the experiment.

We realize that this analysis was partly a first order approximation. The SII-model is validated for speech in stationary noise. Although an extended version is available for SII predictions in fluctuating noise (Rhebergen et al, 2006), this model cannot predict hearing-impaired speech reception and is not used in this study. However, modulation frequencies between 10-20 Hz are known to generate the lowest SRTs when using monosyllabic speech material, and several studies report 16 Hz as an optimum modulation rate (Festen & Plomp, 1990; Smits & Houtgast, 2007).

The following masking noises have been selected for the experiments:

1. Earcheck: a broadband stationary speech-shaped noise, as described above.
2. 16-Hz: a broadband interrupted noise, with a modulation depth of 15 dB.
3. LP: a low-pass filtered stationary noise, with a -15 dB noise floor

4. LPmod: a low-pass filtered stationary noise, combined with high-pass filtered 16-Hz modulated noise, with a modulation depth of 15 dB.
5. HP: a high-pass filtered stationary noise, with a -15 dB noise floor
6. HPmod: a high-pass filtered stationary noise, combined with low-pass filtered 16-Hz modulated noise, with a modulation depth of 15 dB.
7. NF: a broadband stationary noise, consisting of only the noise floor of -15 dB.

The characteristics of the filtered noise are specified in Table 5.1 and schematically illustrated in Figure 5.2. The spectrally filtered noises are digitally filtered with either a low-pass or a high-pass filter, employing a cut-off frequency of 1.4 kHz and a steep roll-off slope of more than 100 dB per octave. For all interrupted noises, the speech-shaped noise was modulated by a 16-Hz square wave, with 50% duty cycle. The final condition was a low-level broadband noise, referred to as the 'noise floor', created by attenuating the speech-shaped noise by 15 dB. In each alternative noise condition this noise was additionally present, to ensure that the noise floor was sufficiently high to mask potential ambient noise levels. In addition, the noise floor produced more or less equivalent masked thresholds for all subjects, minimizing differences in speech audibility among subjects.

Table 5.1. Characterization of the modified masking noises.

Name	SNR start	Filtering	Modulation	Noise floor
Earcheck	0	-	-	-
16 Hz	-10	-	16 Hz squarewave	-15 dB
LP	-10	LP (1.4 kHz)	-	-15 dB
LPmod	-10	LP (1.4 kHz)	16 Hz squarewave	-15 dB
HP	-10	HP (1.4 kHz)	-	-15 dB
HPmod	-10	HP (1.4 kHz)	16 Hz squarewave	-15 dB
NF	-15	-	-	-15 dB

The noises were generated such that the spectral part of the filtered noise that was included or the temporal part that was "on" was identical to the steady-state noise. Accordingly, the overall level of the modified noises was slightly reduced. No adjustments in level were made to compensate for this difference. This way, the benefits of removing parts of the background spectrum can be examined without the confounding effect of increases in the level of the remaining part of the spectrum.

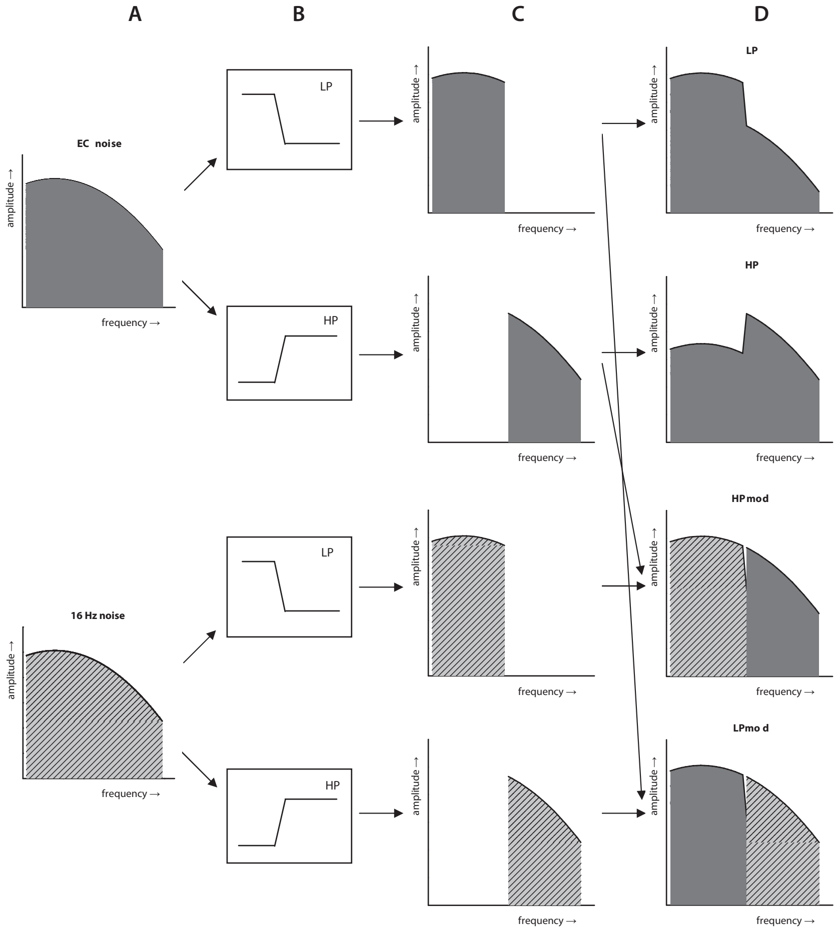


Figure 5.2. Schematic presentation of the creation of the different masking noise conditions. A: Spectral representation of the broadband masking noise indicated as the stationary EC noise (dark grey in upper section) and modulated 16-Hz noise (light grey shadowed in lower section). B: Schematic presentation of the filters; LP shows the low-pass filter, HP shows the high-pass filter. C: Representation of the filtered noise spectra of the stationary and modulated noise. D: Schematic representation of the modified masking noises. The upper section shows the stationary LP and HP conditions after combining the filtered results of C with the noisefloor of -15 dB. The lower section shows the LPmod and the HPmod conditions after combining the filtered results of the stationary filtered noise with a complementary modulated filtered noise as represented in C.

Constructing the tests

All processing was done using a 16-kHz sampling rate and the processed signals were converted to a 44.1 kHz rate. Speech and noise files were stored in MP3 format and a Macromedia Flash player (Macromedia Inc., San Francisco, USA) web application was used to mathematically mix the SNRs of the speech and noise files, according to an adaptive procedure with fixed noise level and variable speech level.

The constructed noises all had durations of 10 seconds, and they were recorded preceding each test for calibration purposes. Of each noise, a fragment of 2 seconds was randomly chosen to be used as test stimulus. Rise and fall times of 0.5 seconds were applied.

The SNRs of these modified tests ranged from -30 to -6 dB. In case subjects gave a correct response on -30 dB or an incorrect response on –6 dB the next stimulus was presented at the same SNR, due to ceiling effects. The starting level was fixed at 0 dB for the original Earcheck, and at lower SNRs for the modified masking noises (Table 5.1).

Procedure and set-up

Subjects were tested individually in a sound-insulated booth. At the beginning of the experiment, a pure-tone audiogram was recorded at the octave frequencies of 0.125-8 kHz and additionally at 3 and 6 kHz, using a Decos (AMC, LUMC) or Interacoustics (UMCN) clinical audiometer and TDH-39 headphones. In addition, bone conduction was measured at 0.25, 0.5, 1, 2 and 4 kHz. All consecutive speech-in-noise testing is done in case of monaural signal presentation. For the normal hearing listeners the tested ear was either the subject's best ear, or, in case of symmetric hearing, the right ear. For the NIHL subjects, the ear showing the most pronounced audiometric dip was selected, but in all cases it was checked that the asymmetry did not lead to cross hearing to the contralateral ear. Following audiometric threshold testing, participants performed the different speech-in-noise tests. Signals were played out via a standard soundcard (Gina 24/96) on a PC at a sample frequency of 44.1 kHz and were fed through a TDT headphone buffer (HB6) and a TDT programmable attenuator (PA4) via TDH-39 headphones. In the UMCN, signals were fed through the AC-40 audiometer.

First, the Dutch sentence SRT test was assessed in quiet (SRT_q), using list 1 and 2 as developed by Plomp and Mimpen (1979a). These measurements were used to set the masking noise level of all consecutive speech-in-noise tests. This noise level was fixed at 65 dBA, or at SRT_q + 20 dBA to ensure audibility in cases of highly elevated SRT in quiet where 65 dBA is not high enough above threshold. Next, two sentence lists in stationary noise were performed. The order of the lists in noise was counterbalanced. The participants received the sentences monaurally to the test ear and were instructed to repeat them as accurately as possible. A sentence was scored correct if all words in that sentence were repeated correctly.

After finishing the sentence tests, the subjects performed the various Earcheck tests with different masking noises. Again, signals were presented monaurally, using Sennheiser HDA-200 headphones in an otherwise identical test set-up. However, at the LUMC no internet access was available in the audiologic booth and the online Earcheck measurements had to be carried out in a quiet room. Ambient noise levels were monitored during all test sessions and are considered to have no effect of performing the supra-threshold speech-in-noise tests. See for more details Chapter 4. The participant was seated in front of a computer touch screen to enter the responses of the different versions of the Earcheck tests. The experiment was divided into two blocks, a test and a retest. Between these blocks subjects paused for approximately fifteen minutes resulting in an intermediate period of 45 minutes between each test and retest pair.

The sequence of test conditions was counterbalanced according to a Latin square method, to avoid learning effects and confounding effects of measurement condition order. The noise levels of each test were calibrated with a B&K type 2260 sound level meter and a B&K type 4153 artificial ear, with the use of a flat-plate adaptor.

Results

First, the SRT results obtained with Earcheck in the various masking conditions are analysed with respect to differences in masking noise and hearing ability. Second, the different Earcheck test and retest results and intelligibility functions are analysed to assess test-retest reliability. Third, correlations between word recognition in different masking noises and both performance on the Dutch sentence SRT test and pure-tone thresholds are analysed, to assess test validity. Finally, for the most discriminating test the sensitivity and specificity are calculated, as this will be the most appropriate candidate for a future NIHL screening test.

Effect of masker types on test results

The effects of the various masking noises employed in Earcheck are examined by analysing the performance of both normal-hearing subjects and participants with different degrees of NIHL. The average SRT results of these groups are displayed in Figure 5.3, for each masking noise condition. Only the speech reception thresholds of the first test are considered for this evaluation, because this is the most representative for people who will do the test only once, as this will be the case in normal practice.

The highest SRTs in each group are generated by the unmodulated Earcheck and lowest SRT values are found when only the noise floor is present. All other modified noise conditions yield SRT values that lie between these extremes. A repeated

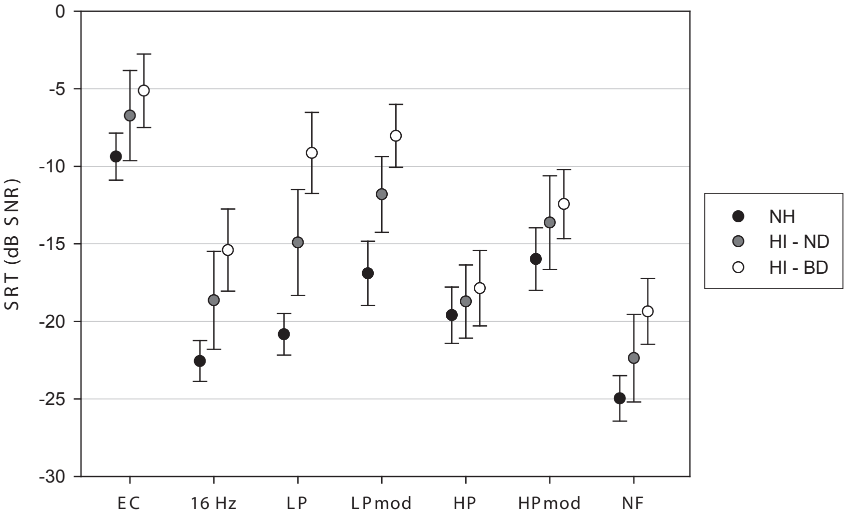


Figure 5.3. Mean SRT for each test, separated for the normal-hearing listeners (black symbols), the hearing-impaired with narrow dip (grey symbols) and the hearing-impaired with broad dip (white symbols). Error bars represent one SD.

measures analysis of variance shows a significant main effect of both 'test condition' (F[6,564] = 799.92, $p < 0.001$) and 'subject group' (F[2,94] = 122,78, $p < 0.001$). Also the interaction between 'test condition' and 'subject group' is significant (F[12,564] = 34,59, $p < 0.001$), indicating that the differences between the subject groups vary between the different test conditions.

Differences between subject groups

To further investigate these differences between subject groups for each test condition, test results are analysed using one-way ANOVA's and post-hoc t-tests with Bonferroni correction for multiple comparisons. As is presented in Table 5.2, the main effect of 'subject group' is significant in each test condition. Post-hoc t-tests show significant differences between nearly all subject groups for all tests, except for HP results; SRT results of the hearing-impaired subjects with a narrow dip do not differ from SRTs obtained by the two other subject groups. In addition, HPmod makes no significant distinction between the two hearing-impaired subjects groups.

All other tests result in significant differences between the three subject groups. While the results for both stationary noises show small differences across the groups, the SRT measured in the interrupted noise, either broadband or combined with low-pass filtering, increases when the respondent has more severe NIHL. However, the greatest differences between the subject groups are found when using the low-pass filtered masking noise.

Table 5.2. Results of a one-way ANOVA investigating the main effect of 'subject group' performed for each configuration of Earcheck and the mean differences (in dB) between the three subject groups for each test.

Test	F-value ANOVA	Δ NH/HI-ND *mean*	Δ NH/HI-BD *mean*	Δ HI-ND/HI-BD *mean*
Earcheck	32.7	-2.6	-4.2	-1.6*
16 Hz	85.4	-3.9	-7.2	-3.2
LP	162.0	-5.9	-11.7	-5.8
LPmod	93.5	-5.1	-8.9	-3.8
HP	5.6	-0.9#	-1.7*	-0.9#
HPmod	20.5	-2.3	-3.5	-1.2#
NF	61.8	-2.6	-5.6	-3.0

*# not significant, *significant at <0.05. Other differences were significant at $p < 0.001$. All p-values are corrected using Bonferroni correction for multiple comparisons.*

Reliability of the online screening tests

It is the goal of this study to find an Earcheck condition that discriminates better between NH and HI listeners. A great spread in test results and large differences in SRT values of these groups are important test parameters in providing a sensitive measure of disability. However, this only applies when the reliability of the improved test is comparable to that of the original version. To evaluate the reliability of each test condition, test and retest results are investigated. The mean test and retest results of the three subject groups are displayed in Table 5.3, with relevant test characteristics of the various Earcheck configurations with respect to test-retest reliability also presented. These characteristics are obtained using results of the total population of 98 participants.

In the two LP noise conditions, some of the hearing-impaired participants have several consecutive incorrect responses at -6 dB SNR, resulting in an invalid SRT. Although this results in a test outcome that still clearly deviated from normal-hearing performance indicating poor performance (see Figure 5.3), these data are omitted when determining the test-retest reliability characteristics.

Table 5.3. Mean SRT values and test-retest characteristics for the various configurations of Earcheck.

Test	Subject group	Test 1 *mean (SD)*	Test 2 *mean (SD)*	Mean test-retest difference[1]	SEM[1]	ICC[1] test-retest	SD_{inter}	Slope (%/dB)
Earcheck	*NH*	-9.4 (1.5)	-11.2 (1.3)	1.62 (p<0.001)	1.17	0.84	2.79	11.5
	HI-ND	-6.7 (2.9)	-8.6 (2.7)					
	HI-BD	-5.1 (2.4)	-6.1 (2.8)					
16 Hz	*NH*	-22.6 (1.3)	-23.2 (1.3)	0.70 (p=0.001)	1.45	0.86	3.74	11.2
	HI-ND	-18.6 (3.2)	-19.6 (4.0)					
	HI-BD	-15.4 (2.6)	-16.1 (2.8)					
LP	*NH*	-20.8 (1.3)	-21.3 (1.7)	0.71 (p<0.001)	1.25	0.93	5.38	11.2
	HI-ND	-14.9 (3.4)	-16.3 (3.4)					
	HI-BD	-9.1 (2.6)	-9.2 (2.8)					
LPmod	*NH*	-16.9 (2.1)	-17.8 (1.8)	0.48 (p=0.028)	1.39	0.87	4.31	10.6
	HI-ND	-11.8 (2.4)	-11.8 (2.8)					
	HI-BD	-8.0 (2.0)	-8.1 (1.7)					
HP	*NH*	-19.6 (1.8)	-21.0 (2.2)	0.73 (p=0.009)	1.92	0.41	2.22	8.6
	HI-ND	-18.7 (2.4)	-18.9 (2.7)					
	HI-BD	-17.9 (2.4)	-18.1 (2.7)					
HPmod	*NH*	-16.0 (2.0)	-16.5 (2.2)	0.46 (p=0.073)	1.77	0.60	2.79	10.0
	HI-ND	-13.6 (3.0)	-13.9 (2.8)					
	HI-BD	-12.4 (2.2)	-12.9 (2.1)					
NF	*NH*	-25.0 (1.5)	-25.8 (1.5)	0.73 (p<0.001)	1.28	0.83	3.07	12.2
	HI-ND	-22.4 (2.8)	-22.7 (2.7)					
	HI-BD	-19.4 (2.1)	-20.3 (2.9)					

[1] Test-retest parameters are calculated excluding the subjects with an invalid SRT for LP (n = 9) or LPmod (n = 13) conditions . Displayed ICCs are Two- way Mixed model, Type Consistency, Single measures ICCs.

Learning effect

One measure of reliability is the difference between test and retest results, estimating a possible systematic change in performance due to learning or fatigue. Table 5.3 shows that the performance in the second test is better than in the first for all conditions, although the differences are small (< 1 dB). Paired-sample t-tests reveal that test-retest differences are significant for all tests, except for HPmod ($p = 0.073$). Although this suggests a learning effect for these tests, the effects are of a magnitude that is smaller than inter-individual differences and thus are not considered relevant for clinical applicability.

Test-retest reliability

Reliability can also be defined as the consistency of a test's results across series of observations. This is expressed as the standard error of measurement (SEM), calculated by dividing the standard deviation of the differences by $\sqrt{2}$ (Weir, 1995). This SEM is smallest for the original Earcheck, 1.17 dB (Table 5.3). The SEM values of the modified test conditions are all larger, although this deviation is only small for the LP and NF conditions (1.25 dB and 1.28 dB respectively). Both tests using high-pass filtered masking noise show the highest SEMs.

Speech intelligibility measurements should differentiate between subjects with different degrees of hearing loss. Therefore results should show small SEM and large inter-individual standard deviations (SD_{inter}) (Wagener & Brand, 2005). Table 5.3 shows that the spread in SRT values over subjects is higher than the SEM for each test condition. SD_{inter} was higher for the low-pass filtered and interrupted noises than for the stationary noise.

In addition, test reliability can be expressed by the intraclass correlation coefficient (ICC) of the test and retest results (Table 5.3). These ICCs are comparable for both stationary broadband Earchecks (EC: 0.84 and NF: 0.83). The ICC's for the modified test conditions are larger, except for both high-pass filtered test conditions that yield rather low ICCs. The highest value for ICC was found for LP noise (0.93).

Intelligibility functions

Another important test characteristic is the slope of the psychometric function, which is a good indicator of the precision of the test. Test-specific intelligibility functions are determined based on a fit of the proportions correct at different presentation levels, for the pooled data of both test sessions. The SNR of every presentation is corrected by the individual SRT for that test and for each relative SNR the proportion correct is calculated. This data is fitted with the logistic regression formula (Equation 5.1), with guess rate 0.11. These intelligibility functions are shown in Figure 5.4. The results for the two groups of HI subjects were merged, in order to obtain a sufficient number of data points per condition.

The horizontal arrangement of the 50% points of the psychometric functions represents the differences in SRTs obtained in the various testing conditions. The online tests showed relatively similar slopes (Table 5.3). The stationary broadband Earcheck has a slope of 11.5%/dB, and the alternative test conditions yield comparable slopes, ranging from 10.0 to 12.2%/dB; at 8.6%/dB only the slope of HP is slightly shallower than that of the other tests. For all tests, slopes are slightly steeper for normal-hearing than for hearing-impaired listeners, but these differences were very small.

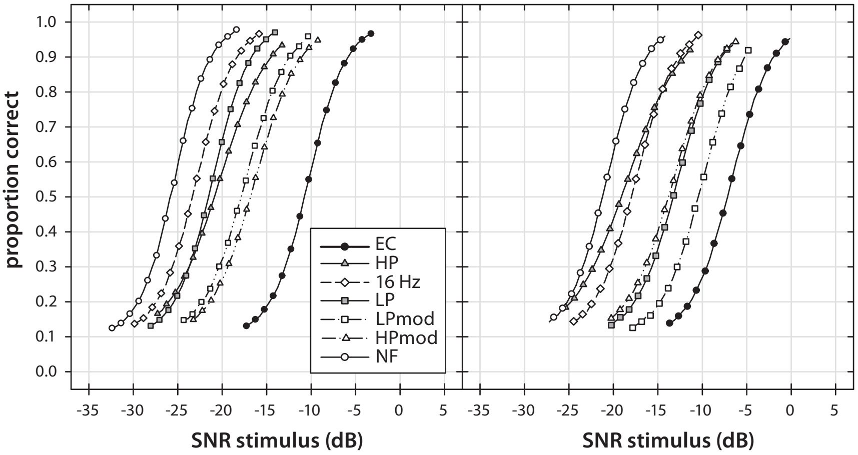


Figure 5.4. Performance-intensity functions of Earcheck employing the different masking noises for both the normal-hearing group (left) and the hearing-impaired listeners (right), showing proportion correct as function of the signal-to-noise ratio of the presentations. The presented SNR is relative to the individual SRTs and then corrected to the average test SRT. To assure a sufficient amount of measurements in the HI group, no distinction between hearing-impaired subjects in ND or BD groups was made, and intelligibility functions are based on the pooled data of all 49 HI participants.

Relationship between SRT results and the pure-tone audiogram

Test validity relates to the correlation between the test's results and other parameters of auditory functioning. Therefore performance on the alternative tests is compared to performance on the Dutch sentence SRT test (Plomp & Mimpen, 1979), the clinical standard in the Netherlands. In addition, the relationships of the different SRT results and pure-tone thresholds are assessed.

In these analyses only results of the first test session are taken into account. This is representative for real-life test performance, since retest results show slightly better SRTs due to training effect. First, mutual Pearson correlation coefficients between the different Earcheck variants are calculated, in order to establish the amount of association between the different test conditions (Table 5.4). The correlations between online test results and the Dutch sentences SRT test are all statistically significant ($p < 0.001$) and show different degrees of association (Table 5.4). The SRTs for both HP conditions show only weak correlation with the sentence test outcomes. Both test versions using broadband stationary noise show moderate correlation; for the unmodulated Earcheck r is 0.71. The SRTs obtained in interrupted noise and low-pass filtered background show reasonably high correlations (around 0.80).

Table 5.4. Bivariate correlation coefficients (Pearson's r) for SRT in noise of different tests with SRT results and PTA-values. All correlation coefficients are significant at $p < 0.001$.

	EC	16 Hz	LP	LPmod	HP	HPmod	NF
EC	-						
16 Hz	0.77	-					
LP	0.76	0.88	-				
LPmod	0.76	0.86	0.91	-			
HP	0.39	0.42	0.40	0.40	-		
HPmod	0.62	0.63	0.56	0.64	0.56	-	
NF	0.77	0.80	0.83	0.81	0.44	0.62	-
Sentence SRT	0.71	0.80	0.80	0.79	0.40	0.60	0.79
$PTA_{0.5,1,2,4}$	0.75	0.84	0.91	0.88	0.35	0.57	0.82
$PTA_{3,4,6}$	0.76	0.86	0.92	0.88	0.34	0.57	0.81

Earcheck results obtained in the different masking noise conditions are compared with pure-tone thresholds. Table 5.4 displays correlation coefficients of SRTs and the pure-tone averages $PTA_{0.5,1,2,4}$ (important for speech intelligibility) and $PTA_{3,4,6}$ (noise-sensitive frequencies). All correlation coefficients between SRT results and $PTA_{0.5,1,2,4}$ are statistically significant ($p < 0.001$), although correlations vary from weak (0.35 for HP) to high (0.91 for LP). The coefficients for $PTA_{3,4,6}$ are about the same as those found for $PTA_{0.5,1,2,4}$. Again, the highest correlation is obtained comparing hearing thresholds with SRT in low-pass filtered noise.

In addition, the correlation of SRT results and pure-tone thresholds at each single frequency is calculated, which is displayed in Figure 5.5. The correlation coefficients of Earcheck outcomes in stationary noise increase as frequency increases, reflecting that the greatest variation in this study population is in the higher frequencies. Except for the curves for SRTs in both high-pass filtered masking noise, which display weak correlation and are rather flat, all modified test conditions show strong correlations with audiometric thresholds, especially in the high frequency range. Results of the low-pass filtered condition show the highest correlation coefficients with hearing thresholds of 2 kHz and higher.

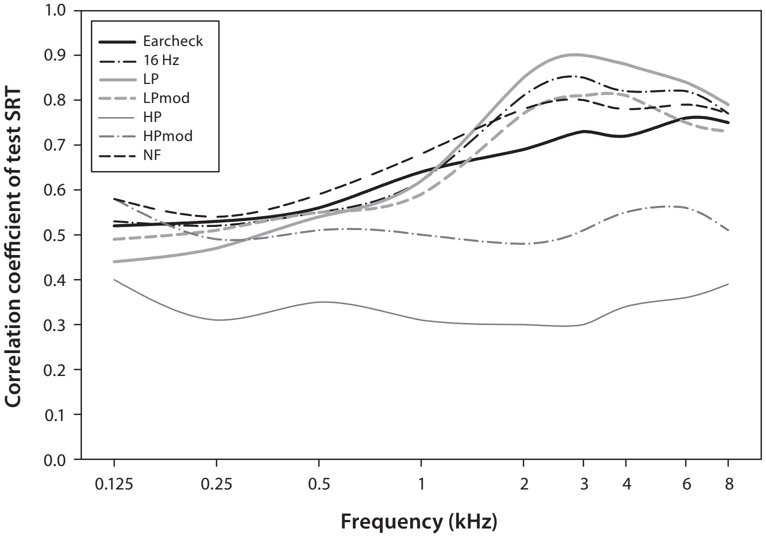


Figure 5.5. Correlation coefficients of SRTs (Pearson's r) with pure-tone thresholds of the total study population, displayed as functions of frequency, for each of the test configurations.

Sensitivity and specificity for NIHL

It is shown that the correlation coefficients between pure-tone thresholds and SRTs in the two low-pass filtered noise conditions are fairly high, and that also the largest group differences occur for these tests. This suggests that a speech-in-noise test in low-pass noise can be used to make better distinction between NH listeners and respondents with NIHL and thus to screen for noise-induced hearing loss.

When a test is deployed for screening purposes it should have both a high sensitivity and a high specificity. Test sensitivity refers to the percentage of hearing impaired participants classified correctly as having a hearing ability worse than 'good'; test specificity refers to the number of normal-hearing subjects that is correctly classified as normal-hearing listener. The sensitivity and specificity of a test depend critically on the cut-off criterion chosen to distinguish between normal hearing and noise-induced hearing impairment. Figure 5.6 shows receiver operating characteristics (ROC) curves that are calculated, based on the results of the first test session, to explore this relationship in more detail. A second curve is added to these ROC curves, representing the relationship between the sensitivity, specificity and the cut-off value for each test, when discriminating between hearing-impaired participants with either a narrow or a broad audiometric dip (Figure 5.6).

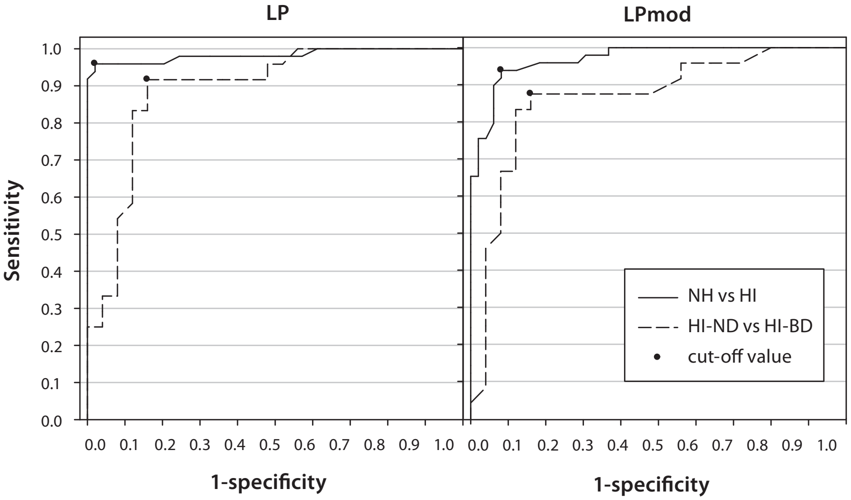


Figure 5.6. ROC-curve, showing sensitivity and specificity of Earcheck using LP noise (left) and using LPmod (right), depending on cut-off values discriminating between normal-hearing and hearing-impaired subjects (solid line) and values separating mild hearing loss (HI-ND) from severe hearing loss (HI-BD, dashed line). The symbols correspond to the chosen cut-off values, as represented in Table 5.5.

The ROC curves provide insight into the optimal combination of cut-off value and the resulting sensitivity and specificity. The SRT value that yielded a high sensitivity without a sufficient decrease in test specificity is chosen as the appropriate cut-off value. The chosen values of the two Earchecks with LP noise are shown in Table 5.5, and are compared to the values of the original test (see Chapter 4). After choosing the cut-off values in order to differentiate between normal-hearing and noise-induced hearing impairment, the Earcheck with low-pass filtered interfering noise showed the best combination of a high sensitivity and a high specificity of the two test alternatives. It turns out that the use of low-pass masking noise increased test sensitivity from 51% obtained for the stationary broadband test to 95%, while the specificity increased from 90% to 98%.

Table 5.5. Area under the curve (AUC), sensitivity and specificity of the original Earcheck (see Chapter 4) and Earcheck using either one of the low-pass filtered masking noises.

	AUC	Cut-off value 1	Cut-off value 2	Sensitivity	Specificity
Original EC	0.75	-10.0	-7.0	51%	90%
LP	0.98	-18.4	-12.7	95%	98%
LPmod	0.97	-14.3	-9.6	94%	92%

Cut-off value 1 discriminates between normal-hearing and hearing-impaired performance, and results in displayed sensitivity and specificity. Cut-off value 2 discriminates between mild NIHL and severe NIHL performance.

Discussion

Prevention of noise-induced hearing loss may be improved by an easily accessible self-test, as early detection of NIHL may lead to increased awareness about hearing loss and initiate proper audiological follow-up for affected individuals. The on-line speech-in-noise test Earcheck is considered to be useful for NIHL screening purposes, but an evaluation study showed that the sensitivity of this test to identify NIHL is too low (see Chapter 4). This study was conducted in order to improve the sensitivity of Earcheck for detecting noise-induced hearing loss, by investigating the effects of spectrally and temporally modified interfering noises on speech intelligibility, relative to the original test.

Results of the original test

The speech material of Earcheck was homogenized in order to increase test accuracy. After this process, the noise was adapted in order to maintain the match to the long-term average spectrum of the speech. Mean SRT for normal-hearing subjects on this test was -9.4 dB, which is somewhat poorer than the SRT found for the original version without homogenization (see Chapter 4). Both hearing-impaired groups showed slightly poorer performance as well. This resulted in greater SRT differences between the three subject groups when performing the homogenized EC. It is important to note that this improvement in discrimination was achieved while the overall test accuracy was unchanged. However, a larger learning effect was observed.

Modified test results

Earcheck using stationary speech-shaped noise generated the highest SRTs of all test conditions, since a spectrally matched broadband noise is the most effective masker.

The modified noises resulted in lower SRT values for all subjects. This improvement in SRT when going from stationary noise to modified noise can be defined as masking release. Interestingly, hearing-impaired participants showed less masking release than normal-hearing listeners for the majority of test conditions, resulting in greater differences relative to normal-hearing performance.

Interrupted masking noise

SRT values obtained in the 16-Hz interrupted noise improved due to listening in the dips of the masking noise. Although the greatest effect is seen in the normal-hearing group, some masking release is also present in hearing-impaired listeners. This decreases with increasing degree of hearing loss, resulting in only little masking release in the most severely affected HI listeners. The participants with NIHL showed reduced ability to take advantage of the dips in the masking noise relative to normal-hearing listeners. Consequently, the differences between normal-hearing and hearing-impaired listeners increase for interrupted masking noises, as was also reported in previous studies (Festen & Plomp 1990; Phillips et al, 1994; Stuart & Phillips, 1996; Smits & Houtgast, 2007; Peters et al, 1998; Bacon et al, 1998).

Low-pass filtered masking noise

Hearing-impaired listeners show the smallest masking release for the two low-pass filtered masking noises compared to all other noise conditions. Their masking release is considerably less than found for normal-hearing listeners; not only the subjects with broad audiometric dips - and poorer thresholds in the high frequencies - show very little benefit from the low-pass filtering of the masking noise, also participants with a narrow dip perform significantly poorer than normal. The high masking release found for the normal-hearing listeners shows that they are able to use the additional high-frequency speech information presented at a higher SNR due to the low-pass filtering. Hearing-impaired subjects have less advantage of this additional spectral information. Their masking release turns out to be strongly related to the degree of hearing loss, indicating that their lack of benefit probably originates from poorer audibility for speech due to higher hearing threshold levels in the high-frequency region.

The NH subjects and HI-ND listeners show higher mean SRTs in the low-pass filtered and interrupted noise condition (LPmod) than in the steady state low-pass filtered masking noise (LP). As relatively less silence occurs in the LPmod condition than in LP, this is consistent with the shape of the noise level distributions. On the contrary, the hearing-impaired participants with a severe loss do not show any difference between the SRTs obtained in the two low-pass filtered conditions. Apparently their ability to take advantage of the less masked high-frequency speech information is very limited since a small increase the degree of high-frequency masking does not affect their performance.

High-pass filtered masking noise

Finally two high-pass filtered background conditions were tested. We expected little or no effect of these conditions because the low-frequency hearing thresholds responsible for possible intelligibility differences in these noises, are comparable among all participants. These HP conditions were mainly chosen to complete our experimental design, in order to control for potentially unexpected results of the low-pass filtered conditions. However, these test conditions also provide useful additional information.

The most striking result of the HP conditions is that they do not discriminate between normal-hearing and hearing-impaired listeners at all. Unlike the higher SRTs in LP noise, hearing-impaired participants do not perform different from normal in HP noise. This indicates that their audibility for speech in noise at higher SNR in the *lower* frequency region is unaffected by their hearing loss. The addition of fluctuations to this noise type in HPmod results in significantly higher SRTs than in stationary HP noise for all subject groups. However, subjects with NIHL show slightly smaller masking release for the HPmod condition than normal-hearing listeners. Although their hearing thresholds at the lower frequencies are (nearly) normal, they already seem to show some signs of reduced temporal resolution. Because the high-pass filtered noise conditions did not discriminate well enough between the subject groups they are considered unsuitable to be used in a screening test for NIHL, as was expected.

Test reliability

Although all subjects benefit from both temporal dips in the noise and low-pass spectral filtering of the background, the benefits are consistently less for HI listeners. These alternative masking noises increase the contrasts between normal-hearing and hearing-impaired participants. However, these tests can only be applied as screening tests if the noise modifications do not adversely affect their reliability. The results of this study show that this is not the case.

Test precision is largely determined by the steepness of the slope of the psychometric function. All alternative test conditions have intelligibility functions with comparable slopes, which are similar to that of the original Earcheck (11.5%/dB). Shallower intelligibility functions for speech recognition in interrupted noise, as reported in earlier studies (Festen & Plomp 1990; Phillips et al, 1994; Stuart & Phillips, 1996) are not found in this study.

The standard error of measurement, expressing the consistency of a test, was 1.17 dB for the unmodified Earcheck. All alternative test conditions showed SEM values that are slightly higher but still comparable, except for the HP conditions that yielded much higher SEMs. The best value is found for the low-pass filtered noise condition, 1.25 dB. SEMs are much smaller than the inter-individual differences between the groups, so the precision of the tests should be high enough to differentiate between individual listeners. Another measure for test consistency, the intraclass correlation coefficient, also showed good reliability for the interrupted and LP noises, ranging from 0.83 to 0.93. Again, the best value is found for the low-pass filtered background. Finally, the test-retest differences are calculated for each condition. These are all comparable and much smaller than the average difference obtained for the original Earcheck. Although most of these test-retest differences are statistically significant, suggesting a small learning effect, the magnitude of this learning effect is smaller than the measurement error and can be considered as not clinically relevant.

Test validity

When a test is used for hearing screening, it is important that high correlations exist between its outcome and the presence of hearing loss. Generally, hearing ability is tested by pure-tone audiometry, and hearing loss for speech is determined by the Dutch sentences SRT test, which is considered the gold standard. SRT results of the unmodulated Earcheck show a moderate correlation with the Dutch sentence SRT test ($r = 0.71$). This level of association can be explained by the fact that speech material is quite different (sentences vs. words) as well as the response format (open set vs. closed set). The correlation coefficients found for most of the Earcheck variants are somewhat higher (0.79 to 0.80), except for the high-pass filtered noise conditions that both showed weak correlations with the Dutch sentence SRT test results (0.40 and 0.60).

Considering hearing tested by pure-tone audiometry, it is known that no perfect correlation exists between speech reception thresholds in stationary noise and pure-tone thresholds, because speech reception in noise requires more than just audibility. Previous studies found correlation coefficients ranging from 0.72 to 0.77 between sentence SRT in noise and $PTA_{2,4}$ for NH listeners and subjects with NIHL (Smoorenburg, 1992; Bosman & Smoorenburg, 1995). Current findings show a comparable correlation coefficient of Earcheck results in stationary noise and $PTA_{0.5,1,2,4}$ ($r = 0.75$). Word recognition in the modified test conditions showed a stronger association with PTA, with high correlation coefficients ranging from 0.82 for 16Hz to 0.91 for LP results. However, the HP noises again show much lower correlation coefficients with pure-tone thresholds. Figure 5.5 presents the results of an analysis of correlation coefficients of all SRT results compared with single-frequency pure-tone thresholds. The curves for the HP conditions are almost entirely flat with an r-value around 0.5 or less, indicating that SRTs measured in HP noise are not strongly associated with hearing ability. It is clear that the correlation coefficients between SRT and pure-tone thresholds attain a maximum value at the higher frequencies and obviously this variation is closely related to the SRT results for the modified noises. Again, SRT in low-pass filtered background yielded the highest correlation with hearing loss.

In summary, Earcheck employing low-pass filtered masking noise reveals the most pronounced differences between hearing-impaired and normal performance and shows the highest correlations with high-frequency pure-tone thresholds, without significantly reducing test reliability. Therefore this modification can be considered as the best alternative to screen for NIHL. Indeed analysis revealed a high sensitivity of 95%, as only two mildly hearing-impaired listeners had recognition performances within the normal range. The test specificity of 98% was also very high.

The approach used here may be assumed to be applicable to other types of speech-in-noise tests as well. If this hypothesis is valid, manipulation of the masking noise can increase the sensitivity and specificity with respect to NIHL for different types of speech-in-noise screening tests. In addition, the results of this study may be generalized to any individual with a high-frequency sensorineural hearing loss regardless of the etiology, which is important for the broad application of screening tests such as Earcheck in the Netherlands.

Study limitations

The results presented here were obtained in lab situations, under well-controlled test conditions and with standardized parameters. A potential limitation of a broad application of this screening test over the internet might be the lack of control over environmental variables at the remote test site, such as ambient noise levels, and over testing conditions such as PC settings, presentation level and transducers used. These parameters have no or little influence on test outcome when presenting speech in a broadband stationary noise with a spectrum matched to the long-term average speech spectrum (Plomp, 1986; Smits et al, 2004; Culling et al, 2005; Ozimek et al, 2009). However, their influence on test results using these modified masking noises is unknown.

Online the presentation level will be individually adjusted before starting the screening test. In this study noise level was fixed at 65 dBA, representing general conversation level. It is reasonable to expect that this could be below the preferred volume of hearing-impaired participants, and that actual testing will be done at higher presentation levels. There is no level effect when measuring SRTs in stationary noise (Plomp, 1986; Wagener & Brand, 2005), but SRT results in interrupted noise become better with increasing presentation level (Rhebergen et al, 2010b). As audibility plays an important role in the intelligibility differences in the low-pass filtered noise conditions, higher presentation levels could result in better SRTs in this noise condition as well. Nevertheless, this effect will be limited by the noise floor, resulting in a more or less equal audibility for all subjects. Also, the presence of background noise at the remote test site may affect test results. However, the added noise floor may limit this effect as well; it is set at -15 dB to ensure that the noise floor was above uncontrollable ambient noise levels.

When performing the test at home, participants are strongly advised to perform the test in a quiet room with headphones, although the use of speakers is allowed as well. Culling et al. (2005) showed that variations in the type of headphone had negligible effects on SRT in stationary noise. In addition, SRTs found for speech-in-noise tests in a living room environment were similar to those obtained under headphones in laboratory conditions (Ozimek et al, 2009), even when loudspeakers were used (Culling et al, 2005). However, other studies indicate that a different set of reference values is needed when speech-in-noise tests are performed using loudspeakers instead of headphones (Smits et al, 2006a; Jongmans et al, 2008).

So there are some remaining uncertainties that need to be investigated before the test can be implemented online for use at home. However, this study shows that the lab results of Earcheck with low-pass filtered masking noise are promising for the purpose of screening for NIHL.

Conclusion

The aim of this study was to examine whether the online speech-in-noise test Earcheck would perform better as a screening test for noise-induced hearing loss when using a modified masking noise, as compared to the original version of the test. Earcheck with a low-pass filtered masking noise showed the best discriminative power between subjects, and was strongly correlated with results of the Dutch sentence SRT test and with pure-tone thresholds, especially in the high-frequencies. This speech-in-noise test can be considered a very useful test for (early) NIHL, due to the small measurement error and the large spread in SRT values in this population of listeners with different degrees of NIHL. The test had a sufficiently high sensitivity of 95% and specificity of 98%, and is thus considered applicable to use as a valid screening test for NIHL. This manipulation of the masking noise may also be applicable to increase the sensitivity and specificity with respect to NIHL or any type of high-frequency hearing loss for other types of speech-in-noise screening tests.

3

Application of the alternative approach

6

Speech-in-noise screening tests by internet, **Part 3**: Test sensitivity for uncontrolled parameters in domestic usage

M.C.J. Leensen
W.A. Dreschler

Revised manuscript submitted to International Journal of Audiology

Abstract

Objective: The online speech-in-noise test ‘Earcheck’ is sensitive for noise-induced hearing loss (NIHL). This study investigates effects of uncontrollable parameters in domestic self-screening, such as presentation level and transducer type, on speech reception thresholds (SRTs) obtained with Earcheck.

Design: Subjects performed 26 Earchecks that differed regarding presentation level (65, 71, and 77 dBA), presentation mode (monotic or diotic), and masking noise (two different low-pass filtered noises) in the lab. To investigate effects of test environment, participants conducted 8 additional Earchecks at home using different transducer types (headphones or loudspeakers).

Study sample: Thirty noise-exposed workers, either normal-hearing (n=10), or with different degrees of NIHL (n=20), participated.

Results: There was a minor effect of presentation levels exceeding 65 dBA in severely impaired listeners. Diotic presentation mode yielded lower SRTs compared to monotic presentation mode. Normal-hearing test results at home were poorer than in the laboratory, whereas hearing-impaired subjects performed better in domestic testing. Using loudspeakers deteriorated SRTs significantly in comparison to headphones, but only in hearing-impaired subjects.

Conclusions: A monotic presentation mode using headphones is recommended for domestic screening. Since domestic testing affect SRT results, a follow up study using a large study population should assess Earcheck’s validity when performed at home.

Introduction

Noise-induced hearing loss (NIHL) is a permanent loss of hearing caused by sustained exposure to intense noise, present in occupational settings and/or during recreational activities. Because NIHL occurs gradually over several years it comes on insidiously, causing the risk of hearing loss to be easily underestimated (Vogel et al, 2008). Therefore, consciousness about hearing problems may be increased by an accessible and reliable hearing screening test. Internet-based screening tests have attracted widespread interest and their rapid growth and increasing usage illustrate that the internet is a suitable medium to contact the general public (Smits et al, 2006a; Koopman et al, 2008; Swanepoel & Hall, 2010). An easily accessible online self-test offers a new approach to facilitate awareness about the risks of hearing loss due to noise or music exposure. Through early detection of hearing loss, it can help to prevent further development of hearing damage in individuals already affected. Furthermore, internet-based self-testing may offer new methods to monitor hearing health in certain noise-exposed populations, such as in occupational settings.

Since subjects with NIHL often encounter difficulties understanding speech in noisy situations (Chung & Mack, 1979; Smoorenburg et al, 1982; Smoorenburg, 1992; Bosman & Smoorenburg, 1995), a speech-in-noise test can be considered a suitable measure for screening. More importantly, speech presented in stationary noise is a very suitable method for online application, because it has less strict acoustical requirements than traditional pure-tone audiometry due to suprathreshold presentation, requires less strict calibration and enables automated test administration by the use of a simple adaptive procedure. In the past years, several simple and automatic online speech-in-noise screening tests became available, such as digit triplet tests in several countries (Smits et al, 2004; Jansen et al, 2010; Zokoll et al, 2012; Watson et al, 2012), and 'Earcheck' in the Netherlands using monosyllables (Albrecht et al, 2005). These tests have proven to be well accepted by the users and offer a reliable self-screening test for hearing loss in general.

Unfortunately, most of these tests lack sensitivity to specifically detect the (mild) high-frequency hearing loss that is typical for beginning NIHL (see Chapter 4), because the words, usually presented in a closed set, can be recognized from low-frequency cues only. However, a recent study showed that when the online speech-in-noise test Earcheck used low-pass filtered masking noise instead of broadband noise, the discriminative power of the test increased substantially (see Chapter 5). Employing a low-pass filtered masking noise facilitates the use of high-frequency speech information, which is advantageous for normal-hearing listeners. Reduced audibility of this high-frequency speech information will limit the potentially positive effect of low-pass filtered noise in subjects with NIHL, since their hearing is affected in this frequency region. Consequently, the test's sensitivity to detect NIHL improved from

51% to 95%, and this was shown to be possible without a reduction in test reliability (see Chapter 5)[3].

Domestic testing has no or limited influence on test outcome when speech is presented in a broadband stationary noise with a spectrum matched to the long-term average speech spectrum (Plomp, 1986; Smits et al, 2004; Culling et al, 2005; Ozimek et al, 2009). However, this may not be the case when the masking noise is spectrally filtered. The increased test sensitivity was shown in a well-controlled experiment performed in a lab environment. Self-testing over the internet may be affected by the lack of control over testing conditions, such as the presentation level, the testing environment (such as ambient noise present), and the expected variety of equipment used by the respondents (such as PC settings). In order to investigate the effect of these parameters on Earcheck with either one of the two types of low-pass filtered noise, test results obtained under different test conditions are compared in this study. These test conditions differed with regard to:

- presentation level: Presentation level is set individually before starting the test. Level is no critical factor in measuring SRTs in stationary speech-shaped noise since it depends upon the SNR rather than upon the absolute level, as long as speech level clearly exceeds the individual's threshold (Plomp & Mimpen 1979b; Wagener & Brand 2005; Theunissen et al, 2009) and is within the range of moderate conversation levels where effects of uncomfortable listening levels do not deteriorate speech intelligibility (Studebaker et al, 1999; Dubno et al, 2006; Summers & Cord, 2007). However, in low-pass filtered noises, level-dependency of the SRTs measured may be introduced, as the audibility of unmasked high-frequency speech information might increase at higher levels.

- presentation mode: Currently, online speech-in-noise tests for domestic screening presents speech and noise diotically. A monotic presentation mode that allows the testing of both ears separately by headphones may be highly beneficial in cases of mild to moderate asymmetric hearing loss. On the other hand, it is known that speech discrimination under diotic conditions is superior to monotic listening (Plomp & Mimpen 1979a; Kaplan & Pickett 1981; Davis et al, 1990; McArdle et al, 2012). In order to choose the appropriate presentation mode, differences between Earcheck results in monotic and diotic presentation mode will be considered in this study.

---

[3] Similar effects were found for a stationary low-pass filtered noise combined with a high-pass interrupted noise replacing the removed high-frequency part of the noise.

- test environment: Earlier studies on speech-in-noise testing in a living room environment, using either headphones (Ozimek et al, 2009) or loudspeakers (Culling et al, 2005), yielded results that were highly comparable to those obtained in laboratory conditions. However, these living room environments were simulated and hence were similar to all participants. In addition, the variability in computers to be used for domestic testing, with different sound cards and various possible PC settings, is largely unknown. These specifications of domestic equipment might become relevant for test outcomes when part of the masking noise is removed by low-pass filtering, as is the case in the Earcheck test investigated in this study.

- transducer type: Several experiments revealed that variations in the type of headphone had negligible effects on SRT in stationary noise (Culling et al, 2005). However, when SRTs are obtained in a low-pass filtered noise, effects of (differences in) the frequency response of the transducer used may influence speech recognition, especially when there is a peak in high frequency region with less masking energy. Instead of using headphones, testing can also be administered using loudspeakers. The advantage of loudspeakers over headphones is their availability as they are standard PC equipment. In addition, the use of loudspeakers offers the ability to assess hearing capacity with hearing aids. However, when a speech-in-noise test is presented over loudspeakers, rather than headphones, the acoustics of the test environment, such as room reverberation and ambient noise levels, could degrade speech intelligibility (Culling et al, 2005; Soli & Wong 2008; Theunissen et al, 2009).

This study consists of two experiments investigating the performance of Earcheck in different testing conditions. In experiment A participants perform Earcheck in different testing conditions under well-controlled lab conditions, in order to investigate the effect of presentation level and presentation mode on SRT results. In addition, test-retest reliability is assessed. In experiment B the same participants perform Earcheck at home, to examine the influence of test environment and transducer type on the SRTs obtained by Earcheck.

Methods

Participants

A selection of male construction employees aged 18 years or older who recently had a periodic occupational health examination was invited to participate in this study. In total 30 participants were included, all employed in different occupations in construction industry. Twenty-four of them reported job related noise exposure. All subjects were native speakers of the Dutch language. The study population was divided into three subgroups. The first subgroup consisted of 10 listeners considered as normal hearing (NH), with pure-tone thresholds below or equal to 20 dB HL for the octave frequencies between 0.25 and 6 kHz, including 3 kHz. Their ages ranged from 32 to 60 years (mean age 46 yrs, SD = 8.7 years), and mean job tenure in construction was 25.6 years (SD = 9.8 years). The remaining 20 subjects were mildly-to-moderately hearing-impaired (HI) participants. They had high-frequency hearing loss defined as one or more pure-tone thresholds greater than 25 dB HL at 2 to 6 kHz. All had normal hearing (≤ 20 dB HL) at frequencies below 2 kHz and none of them used hearing aids or had known middle ear problems. They were employed in construction for 32.2 years (SD = 6.7 years) on average, and all reported some kind of noise exposure, either occupational and/or recreational, in the past.

Analogous to the earlier studies concerning Earcheck described in Chapters 4 and 5, the hearing-impaired participants were divided into two subgroups having either a narrow audiometric dip (HI-ND, n=14, mean age 50 years, SD = 7.0 years), corresponding with early NIHL, or a broad dip (HI-BD, n=6, mean age 56 years, SD = 5.6 years), corresponding with more severe hearing loss. Distinction was made based on whether or not their hearing threshold at 2 kHz was affected; when hearing threshold at 2 kHz exceeded the pure-tone average of 0.5 and 1 kHz by more than 15 dB, the patient was classified as having a broad dip. For each of the three groups, mean audiometric hearing thresholds are displayed in Figure 6.1.

Test stimuli

The online speech-in-noise test Earcheck (in Dutch: www.oorcheck.nl), was used to assess speech recognition performance in noise (Albrecht et al, 2005). Earcheck consisted of a closed set of nine different monosyllabic words (thumb /dœym/, goat / γεit/, chicken /klp/, rat /rɑt/, fire /vyr/, lion /lew/, cat /pus/, saw /zaχ/ wheel /w¡l/). These words were randomly presented in background noise. On screen, nine response buttons were shown, containing a written representation of the words and a corresponding picture. A tenth button, saying ‘not recognized’, was added to prevent guessing. The test was automated using an up-down procedure with fixed noise level and speech level varied adaptively: after an incorrect response, or ‘not recognized’, the next stimuli was presented at a 2 dB higher level, after a correct response the stimulus level was lowered by 2 dB. The test consisted of 27 trials. The signal-to-noise ratios (SNR) of the last 20 presentations were averaged to calculate the speech reception threshold (SRT), defined as the SNR at which 50% of the speech stimuli was recognized correctly. The SNRs of Earcheck ranged from -30 to -6 dB, and the starting level was fixed at -10 dB.

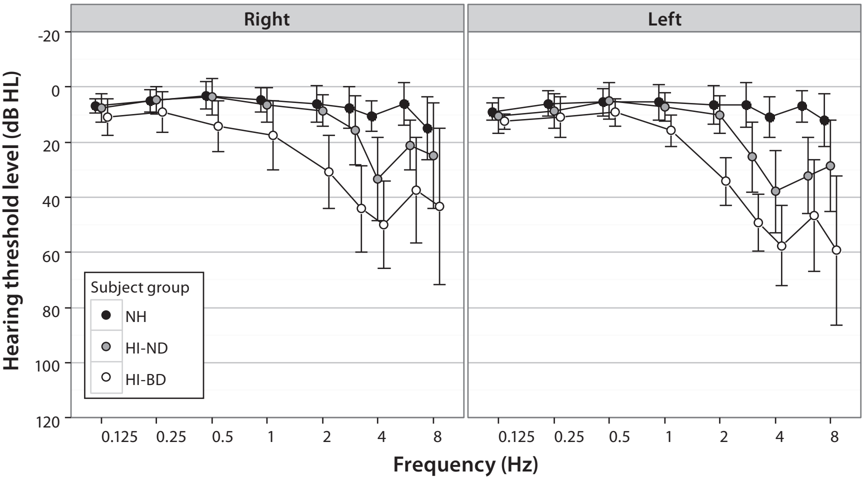


Figure 6.1. Mean audiometric thresholds for each group for left and right ear. Error bars represent one SD.

The masking noise used in this test was a low-pass filtered noise, either without (LP) or with temporal modulations in the high-frequency part (LPmod). Both masking noises were derived by digitally filtering a stationary broadband noise with a long term average spectrum similar to that of the speech stimuli, using a low-pass filter with a cut-off frequency of 1.4 kHz and with a steep roll-off slope (100 dB/octave). To generate the LP noise, a noise floor was added after filtering, that consisted of the speech-shaped noise attenuated by 15 dB (see Figure 6.2 for schematic representation). In order to create the LPmod noise, the stationary speech-shaped noise was modulated by a 0.016-kHz square wave with 50% duty cycle and modulation depth of 15 dB, to generate fluctuating noise. This fluctuating noise was digitally filtered by a high-pass filter with a cut-of frequency of 1.4 kHz and steep roll-off slopes (100 dB/octave) and was added to the low-pass filtered stationary noise (Figure 6.2). Seefor more details Chapter 5.

Procedure and set-up

This study consisted of two experiments. Experiment A was completely conducted at the laboratory of the department of Clinical and Experimental Audiology at the AMC Amsterdam. In experiment B, participants performed the internet-based Earcheck at home in 8 different conditions. The experimental protocol and all procedures in this study were approved by the ethics committee of the University of Amsterdam (approval number: 2001_187).

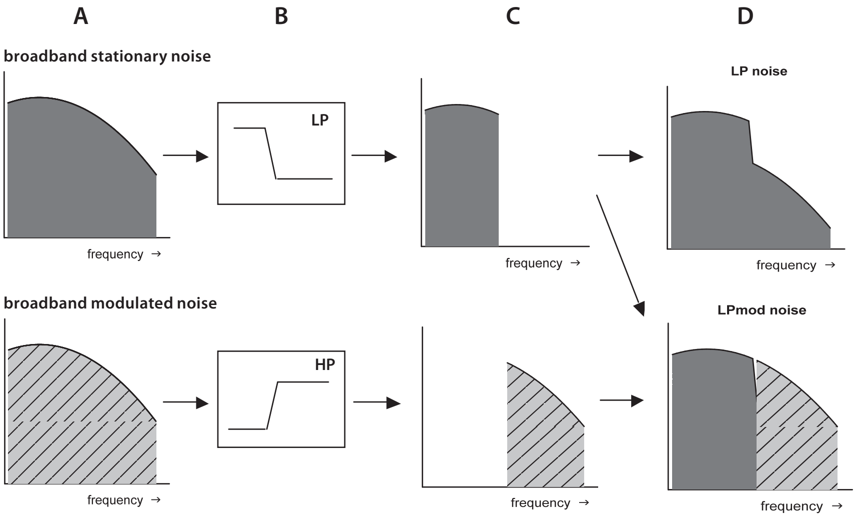


Figure 6.2. Schematic presentation of the creation of the two different masking noise conditions. A: spectral representation of the broadband masking noises indicated as the stationary noise (dark grey in upper section) and 16 Hz modulated noise (light grey shadowed in lower section). B: schematic presentation of the filters; LP shows the low-pass filter, HP shows the high-pass filter. C: representation of the filtered spectra of the stationary and modulated noise. D: schematic representation of the modified masking noises. The upper section shows the stationary LP conditions after combing the filtered results of C with a -15 dB noisefloor. The lower section shows the LPmod condition after combining the filtered results of the stationary filtered noise with a complementary modulated filtered noise as represented in C.

Experiment A

During their visit to the AMC Amsterdam, the subjects were tested individually in a double-walled sound-proof booth (size: lxwxh = 300x200x220 cm). At the start of each test session, pure-tone audiometry was conducted, using a Decos Audionigma PRO audiometer connected to Telephonic TDH-39 headphones. Air conduction thresholds of both ears were obtained at octave frequencies from 0.125 Hz to 8 kHz, including 3 and 6 kHz. The audiometer was calibrated according to ISO-389.1 (1998). Following audiometric testing, participants performed Earcheck in several different testing conditions. Speech and noise signals were played via an Echo soundcard (Gina 24/96) on a PC at a sample frequency of 44.1 kHz. Then they were routed via a TDT Programmable Attenuator (PA4), ensuring separate attenuation as required for a given presentation level, and through a TDT Headphone Buffer (HB7). Subjects received the signals via HDA 200 Sennheiser headphones. The noise levels of the test were calibrated with a B&K type 2260 sound level meter and a B&K type 4153 artificial ear with flat-plate adaptor.

Standard Earcheck administration procedure was utilized in this study (Albrecht et al, 2005). Respondents were instructed to listen carefully and enter their responses through the corresponding button on the computer screen. To allow subjects to become acquainted with test stimuli and procedure, a preliminary sequence of the test words was presented preceding testing. This sequence was presented at an individually chosen comfortable listening level that was set prior to the administration of the actual speech-in-noise tests. Mean individually set preferred listening level was 66.2 dB (range 56.2 – 73.0 dBA, SD = 5.6 dBA).

In order to assess the effect of presentation level, Earcheck was performed at three fixed intensities. Noise levels were set at either normal conversation level or levels that were 6 dB or 12 dB higher; 65, 71 and 77 dBA. Also, testing was done in either monotic (testing each ear consecutively) or diotic signal presentation. The diotic presentation mode was not tested at 71 dB to reduce testing time in order to limit experiment A to one single visit to the AMC. This led to eight different conditions that were all tested in test and retest (referred to as 'repetition'). Monotic retest measurements were limited to a single ear to reduce testing time. The ear to be retested was randomly chosen and remained the same throughout the experiment in each participant. See Table 6.1 for an overview of the different testing conditions.

The various test conditions were conducted in two blocks. Each block contained the 13 different tests using either one of the two masking noises; in 8 tests and 5 retests (Table 6.1). To avoid possible order effects, the tests in each experimental block were counterbalanced across subjects; the eight tests varying in presentation level and presentation mode were counterbalanced using an 8x8 Latin square design, and the five retests were presented in opposite order. For each subject, masking noise condition for the first block of trials was randomly selected. In total, the laboratory experiments comprised 26 Earcheck tests in different conditions.

Experiment B

In order to evaluate differences in SRT results when Earcheck was performed at home as opposed to in a well-controlled lab situation, all participants repeated a selection of eight Earchecks at home (Table 6.1), using their own personal computer. Participants were instructed to complete the test in a quiet environment, at a volume level that is comfortable. Six of the test conditions (3 tests differing in presentation mode (monotic, testing both ears, and diotic) in 2 masking noises (LP/LPmod)) had to be performed with headphones (Table 6.1). To reduce variability in testing equipment in this study, all participants received relatively simple headphones (HQ, type HP 113 LW) for domestic testing[4]. Finally, Earcheck with both types of masking noise was done using participants' own PC loudspeakers.

Table 6.1. Summary of different conditions tested in experiment A (26 tests) under laboratory conditions and experiment B (8 tests) obtained at home.

Masking noise	*Presentation mode*	**Experiment A: lab**						**Experiment B: home**	
		test			**retest**				
		Presentation level (dBA)			*Presentation level (dBA)*			*Transducer type*	
		65	*71*	*77*	*65*	*71*	*77*	*headphones*	*speakers*
LP	*monotic*	2	2	2	1	1	1	2	-
	diotic	1	-	1	1	-	1	1	1
LPmod	*monotic*	2	2	2	1	1	1	2	-
	diotic	1	-	1	1	-	1	1	1

Experiment A was performed in two experimental blocks, each containing the test and retest measurements with either one of the two masking noises. When monotic testing was performed twice in a certain condition, both left and right ear are measured.

Statistical analyses

In total, Earcheck was performed in 34 different conditions by each of the participants. To account for effects of repeated measurements within each individual, linear mixed effect models were used to estimate the difference in SRT over the various testing conditions. This method could also handle the missing variables that were incorporated in the acquired data, due to the experimental design. To perform correct analyses including all variables of interest, the total dataset of SRT results was split up into three subsets, before linear mixed effects models were estimated:

- Dataset A included the 26 test results of experiment A, and was used to investigate the effect of presentation level, presentation mode and repetition.
- Dataset B was used to investigate the differences between test environments and contains all 32 results of tests conducted with headphones in experiments A and B.
- Dataset C was used to study the influence of transducer type and consisted of the results of the four diotic tests either through headphones or through loudspeakers, performed at home in experiment B.

---

[4] Here we deviate from "real-life" domestic testing, but the application of low-cost headphones is a feasible option in a screening program. This approach also solves the problem that headphones are not always available for each PC.

For all analyses of SRTs over different conditions, the variation between 'subjects' and 'ears' within subject were treated as random effects. In addition, Earcheck was performed in two experimental blocks, each with a different interfering noise. Because of this experimental design, variation between 'masking noise' within subjects was also treated as a random effect. For the same reason 'repetition' was considered as a random effect, but this was not shown to be a significant random factor. Fixed factors of primary interest were 'presentation level', 'presentation mode', 'repetition' (in dataset A), 'test environment' (in dataset B), and 'transducer type' (in dataset C). In each model, 'masking noise' and 'subject group' were included as fixed factors, to account for their known systematic differences in SRT results (see Chapter 5). Two-way interaction terms between each of the fixed variables and 'masking noise' and 'subject group' were incorporated in the model as well, since the primary interest of this study is whether shown effects of the tested parameters are similar in both masking conditions and in the three subject groups. When there was a reasonable a priori expectation of interaction between two other factors, these were also included. Only the factors and interaction terms that showed a significant contribution to the fitted model, tested with conditional F-tests at the 0.10 level, were investigated for significant coefficients of each level. When coefficients proved to be significant, the term was retained in the model. Results of the models are displayed as the estimated effects for the fixed effect levels and interaction terms retained in the model and their 99% confidence intervals, relative to the reference condition of monotic test results of the normal-hearing group, presented at 65 dBA in LP noise. In case of an interaction between two variables, the difference between a certain condition and this reference is obtained by summing up the coefficients obtained for each separate factor contributing to that interaction.

The repeated measurements of SRT in both test and retest, and over the different testing conditions, were used to assess the test-retest reliability of Earcheck. The intraclass correlation coefficient (ICC) is a relative index of test precision. It represents the ratio of the between-subjects variability to the total variability in the data, and is used to determine the consistency of the position of individual scores relative to others (Weir, 2005). In addition, the standard error of measurement (SEM), a measure of absolute consistency, was calculated. This measure quantifies the precision of individual outcomes on a test, and assesses the reliability within individual subjects (Weir, 2005). It combines an overall standard deviation (SD) of all measurements and the ICC as follows: $SEM = SD*\sqrt{(1-ICC)}$. When there are two levels of trials, which is the case in test and retest, SEM calculation can be simplified by dividing the SD of the differences by $\sqrt{2}$. The data were analysed using SPSS (version 19.0) and R software (R Foundation 2008, from http://www.R-project.org).

Results

Results of experiment A

The mean SRT results of the three subject groups in the different testing conditions are shown in Figure 6.3. Linear mixed effect models were run as previously described and all fixed variables showed significant effects, explained in more detail below. The coefficients of the full model are presented in Table 6.2.

Effect of masking noise

In this model, 'masking noise' showed a significant main effect (F[1,119] = 69.13, $p < 0.001$), as was expected based on previous results (see Chapter 5). On average, SRTs obtained in low-pass noise were 4.0 dB better ($p < 0.001$) than SRTs obtained under low pass fluctuating noise. There were no significant interactions found between 'masking noise' and the other fixed factors, except for 'subject group' (described below), indicating that the effects of the various testing conditions were similar in both interfering noises.

Effect of subject group

The main effect of 'subject group' (F[2,27] = 25.97, $p < 0.001$) showed that SRTs were different between subject groups. However, only SRTs of the hearing-impaired participants with a broad dip were significantly higher (thus poorer) than found for the participants in both other subject groups. HI-BD subjects showed SRT results that were on average 7.4 dB worse than the NH results ($p < 0.001$). Although SRTs of the hearing-impaired group with a narrow dip were 1.4 dB poorer than those of the normal-hearing group, these differences turned out to be not statistically significant ($p = 0.116$) (Table 6.2). 'Subject group' showed significant interactions with both 'masking noise' and 'presentation level'. The significant interaction between 'subject group' and 'masking noise' (F[2,27] = 5.29, $p = 0.012$) indicated that the mean SRT difference of 4.0 dB between both noise conditions differed between the three subject groups. Indeed, the difference between the two noise conditions was significantly smaller (2.3 dB) in the HI-BD group than in the two other subject groups ($p = 0.039$).

Effect of presentation level

Increasing presentation levels over the range from 65 to 77 dBA showed only minor effects on SRTs measured. The main effect of 'presentation level' was not significant (F[2,593] = 0.58, $p = 0.558$), but there was a significant interaction between 'subject group' and 'presentation level' (F[4,593] = 2.22, $p = 0.065$): the model indicated an improvement in SRT at higher presentation levels only in the severe hearing-impaired participants. In the HI-BD group, SRTs were on average 1.0 dB better at both higher

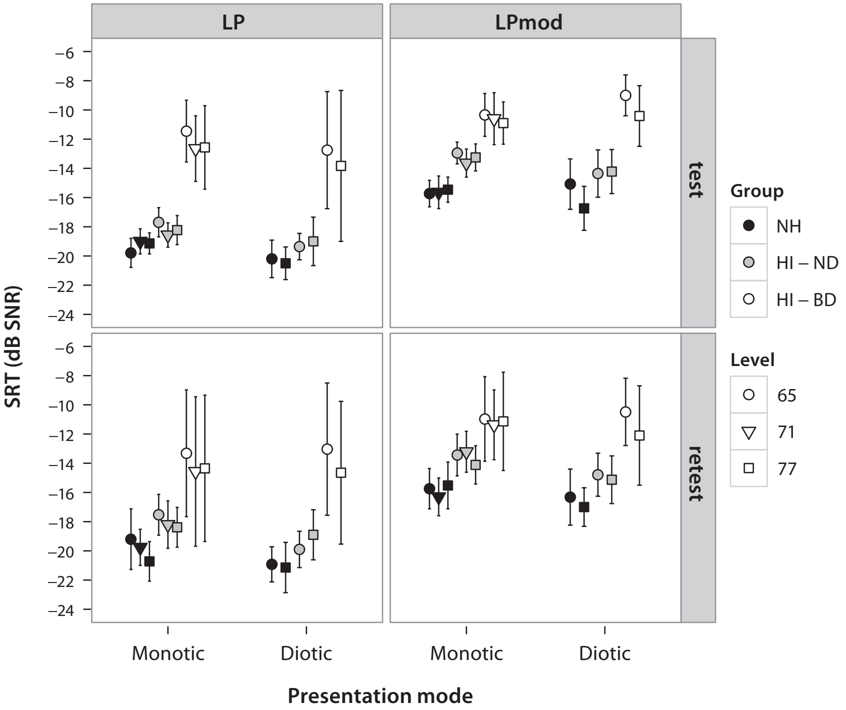


Figure 6.3. Mean SRT results for the three subject groups over the 26 different experimental conditions of experiment A. Error bars represent 95% CI. Results in the stationary LP noise are displayed in the left panel, results in the fluctuating LP noise in the right. The upper panel shows test results, retest results are presented in the lower panel.

levels compared to testing at 65 dBA ($p = 0.024$ for 71 dBA and $p = 0.026$ for 77 dBA). In the normal-hearing and the mildly hearing-impaired listeners groups SRTs at different presentation levels were similar (Table 6.2 and Figure 6.3).

Effect of presentation mode

SRTs were obtained in two different presentation modes: monotic and diotic. In monotic presentation mode both left ear and right ear were tested separately. Although we were not primarily interested in inter-ear differences, our analysis showed that average SRTs obtained from testing the left ear were 0.55 dB higher, thus worse, than SRTs resulting from right ear testing. However, since the focus of this analysis was on presentation mode rather than on differences between ears, the full model considered results obtained in either monotic or diotic presentation mode, regardless of the ear tested.

Table 6.2. Model coefficients of dataset A: all test results of the 26 laboratory conditions.

Model terms	Coefficient	99% CI
Intercept	-19.29	-20.58 – -18.01
Masking noise: LPmod	4.02	3.03 – 5.05
Subject group: HI – ND	1.38	-0.36 – 3.13
Subject group: HI – BD	7.41	5.22 – 9.59
LPmod * HI – ND	0.71	-0.58 – 2.01
LPmod * HI – BD	-1.71	-3.33 – -0.09
Level: 71	-0.07	-0.58 – 0.44
Level: 77	-0.23	-0.66 – 0.19
Level 71 * HI – ND	-0.29	-0.95 – 0.37
Level 71 * HI – BD	-0.96	-1.78 – -0.13
Level 77 * HI – ND	0.02	-0.54 – 0.59
Level 77 * HI – BD	-0.80	-1.50 – -0.10
Mode: diotic	-0.85	-1.19 – -0.52
Repetition: retest	-0.54	-0.77 – -0.31

All coefficients are expressed in dB. In case of an interaction between two variables, the particular difference between a certain condition and the reference situation is obtained by summing up the different individual coefficients contributing to that condition.

'Presentation mode' was a significant main effect in the model (F[1,119] = 26.18, $p < 0.001$). Results from diotic speech-in-noise testing were slightly better than when testing was done monotically. Averaged over all testing conditions, listeners showed a benefit of 0.85 dB performing the test with both ears over monotic testing. There was no statistical interaction between 'presentation mode' and 'subject group', demonstrating all groups benefitted equally from diotic presentation.

Effect of repetition

Most of the test conditions were performed in test and retest, in order to assess a possible test-retest difference. A change in SRT between similar test conditions over time might indicated a learning effect (if the SRT improved) or signs of fatigue (if the SRT deteriorated). Indeed, the full model showed a main effect of 'repetition' (F[1,119] = 20.71, $p < 0.0001$); retest results were slightly better than test results, indicating an average learning effect of 0.54 dB. This effect was similar in all subject groups and in the two masking noise conditions.

An important feature of a measurement procedure is the test-retest reliability. Reliability could be described as the consistency of a test's results across series of observations. Intraclass correlation coefficients and standard error of the measurements were calculated for each of the five conditions tested in test and retest and over the total of 13 obtained SRTs in both masking noises, and are displayed in Table 6.3. For the stationary low-pass filtered noise ICC values between 0.78 and 0.86 were found, resulting in an overall ICC value of 0.81. For Earcheck using low pass filtered modulated noise, overall ICC was somewhat lower, 0.68. The coefficients for the five different conditions ranged from 0.65 to 0.81. The obtained SEMs ranged from 1.33 to 1.76 (Table 6.3) and were similar in both masking noise conditions.

Table 6.3. The intraclass correlation coefficienti (ICC) and standard error of measurements (SEM) calculated for the different test conditions (based on test and retest results) and over the total set of 13 repeated measurements.

	LP		**LPmod**	
	ICC	*SEM*	*ICC*	*SEM*
Diotic 65 dBA	0.84	1.51	0.76	1.63
Diotic 77 dBA	0.80	1.70	0.79	1.43
Monotic 65 dBA	0.83	1.53	0.81	1.33
Monotic 71 dBA	0.78	1.60	0.73	1.53
Monotic 77 dBA	0.85	1.39	0.65	1.76
Total	**0.81**	**1.62**	**0.68**	**1.70**

ICC are two-way random model, absolute agreement, single measure.

Results of experiment B

In experiment B, participants repeated Earcheck in 8 conditions at home, using their own personal computer. Only 26 participants completed these tests; domestic test results were unavailable for three normal-hearing listeners and one hearing-impaired listener with a broad dip. The laboratory results of these participants were excluded before comparing test results acquired in thelab environment to the outcomes obtained at home, using database B. Because the volume setting in domestic testing is unknown the linear mixed model did not account for the factor 'presentation level'. In addition, the factor 'repetition' could not be taken into account, since measurements at home were only performed once. As a result, the linear mixed modeling averaged SRTs obtained in the laboratory over the three presentation levels and over test and retest. Mean SRT results for the different subject groups and test conditions are presented in Figure 6.4.

All fixed test parameters showed significant main effects. Since the main effects of 'masking noise' (F[1,75] = 233.23, p < 0.001) and 'subject group' (F[2,23] = 21.91, p < 0.001) concerned the reference condition of lab testing, these effects were similar to those found in experiment A, even though the study population was slightly smaller. There were significant interactions between 'subject group' and both 'masking noise' and 'test environment', and between 'test environment' and 'presentation mode'. The coefficients of the full model are presented in Table 6.4.

Effect of test environment

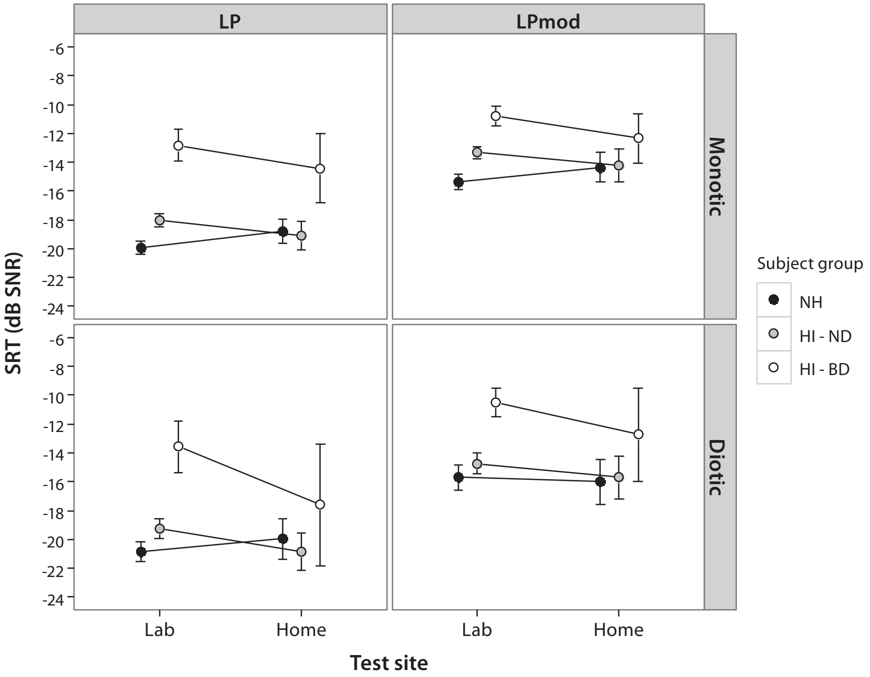


Figure 6.4. Mean SRT results for the three subject groups over the 32 different experimental conditions of experiment B, displayed for lab and domestic testing. The upper panel shows test results obtained in monotic presentation mode, results obtained in diotic presentation mode are displayed in the lower panel.
Results in the stationary LP noise are displayed in the left panel, results in the fluctuating LP noise in the right. Lab results are averaged over the three presentation levels used. Error bars represent 95% CI.

The significant main effect of 'test environment' (F[1,672] = 12.41, p < 0.001) indicated that results of at home testing were 1.0 dB poorer than SRTs obtained in the laboratory environment (p < 0.001). However, this only held for the normal-hearing listeners. There was a significant interaction between 'test site' and 'subject group' (F[2,672] = 27.12, p < 0.001) showing that the domestic SRT results for the hearing-impaired subjects were significantly better than the outcomes in lab conditions (Table 6.4). These differences, of 0.78 dB (p < 0.001) and 1.90 dB (p < 0.001) for HI-ND and HI-BD respectively, seemed to increase with hearing loss (Figure 6.4). In addition, the interaction term between 'test environment' and 'presentation mode' was shown to be significant (F[1,672] = 5.79, p = 0.016). The positive main effect of diotic listening (F[1,51] = 18.75, p < 0.001) in the lab environment of 0.86 dB (p = 0.001) was even larger when Earcheck was performed at home; the diotic SRT results at home were on average 1.60 dB better than monotic test outcomes (p = 0.016).

Table 6.4. Model coefficients of dataset B: all test results of the 32 tests with headphones.

Model terms	Coefficient	99% CI
Intercept	-19.88	-21.39 – -18.37
Masking noise: LPmod	4.68	4.07 – 5.30
Subject group: HI – ND	1.63	-0.34 – 3.60
Subject group: HI – BD	7.11	4.77 – 9.44
LPmod * HI – ND	0.04	-0.72 – 0.80
LPmod * HI – BD	-2.24	-3.14 – -1.34
Test environment: at home	1.04	0.46 – 1.62
At home * HI – ND	-1.82	-2.49 – -1.14
At home * HI – BD	-2.94	-3.74 – -2.14
Mode: diotic	-0.86	-1.26 – -0.46
At home * diotic	-0.74	-1.34 – -0.14

All coefficients are expressed in dB. In case of an interaction between two variables, the particular difference between a certain condition and the reference situation is obtained by summing up the different individual coefficients contributing to that condition.

Effect of transducer type

The 26 participants that conducted Earcheck at home performed the diotic tests with both headphones provided for this study and loudspeakers of their own personal computer, using each of the two interfering noises (dataset C). Mean SRTs of the three subject groups are displayed in Figure 6.5. Since all tests were performed in diotic presentation mode, 'tested ear' was not included as a random factor in the linear mixed model. The coefficients of the full model are presented in Table 6.5.

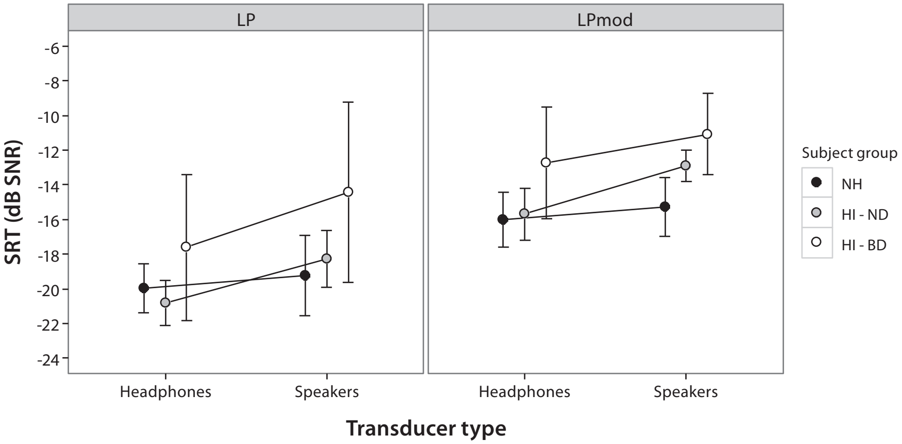


Figure 6.5. Mean SRT results for the three subject groups displayed for the two types of transducer used in experiment C.Results in the stationary LP noise are displayed in the left panel, results in the fluctuating LP noise in the right. Error bars represent 95% CI.

This model also showed significant main effects of 'masking noise' (F[1,25] = 158.41, p < 0.001) and 'subject group' (F[2,23] = 4.14, p = 0.029). Whereas the effect of masking noise was similar as found in previous analyses, the differences between subject groups were smaller than obtained in the previous models. Because the hearing impaired groups performed better at home than under lab conditions, especially when signals were presented diotically, the differences between subject groups in this dataset were reduced. HI-ND still showed SRTs that did not differ significantly from normal performance (p = 0.099), whereas HI-BD subjects showed SRT results that were on average only 2.8 dB higher than the NH results (p < 0.001). This probably also explained why the interaction term between 'subject group' and 'masking noise' did not contribute significantly to this model.

'Transducer type' did not show a main effect on test outcomes (F[1,49] = 1.37,

Table 6.5. Model coefficients dataset C: all test results of the four binaural domestic tests.

Model terms	Coefficient	99% CI
Intercept	-20.30	-22.05 – -18.56
Masking noise: LPmod	4.63	3.87 – 5.39
Subject group: HI – ND	-0.28	-2.46 – 1.90
Subject group: HI – BD	2.81	0.23 – 5.41
Transducer type: speakers	0.73	-0.52 – 1.98
Speakers * HI – ND	1.95	0.40 – 3.50
Speakers * HI – BD	1.70	-0.13 – 3.54

All coefficients are expressed in dB. In case of an interaction between two variables, the particular difference between a certain condition and the reference situation is obtained by summing up the different individual coefficients contributing to that condition.

p = 0.247), but there was a significant interaction between ‘transducer type’ and ‘subject group’ (F[2,49] = 3.36, p = 0.043). Detailed analyses showed that SRTs with headphones did not differ significantly from SRTs obtained via speakers in the normal-hearing group (p = 0.247). However, there was a significant effect of transducer type in both hearing-impaired groups (see Figure 6.5); when the speech and noise signals were played through speakers HI-ND and HI-BD listeners have SRTs that were on average 2.68 (p = 0.015) and 2.43 dB (p = 0.068) higher, respectively.

Discussion

This study aims to investigate the effects of several testing parameters that may vary when the online speech-in-noise test Earcheck is used for domestic screening. Subjects with and without NIHL performed Earcheck in several conditions, both in a well-controlled lab environment and at home using their own personal computer.

Test-retest reliability

The lab results of experiment A prove that Earcheck is a reliable test, yielding relatively high ICCs and good within-subject variability. Yet, the ICCs found in this study are slightly lower than the values found in the evaluation study described in Chapter 5. Also, the SEMs are slightly higher than those reported in previous studies, which range from 0.9 - 1.4 dB (Plomp & Mimpen, 1979a; Smits et al, 2004; Vaillancourt et al, 2005; Van Wieringen & Wouters, 2008). The reliability of Earcheck using stationary LP noise is slightly better than that with modulated LP noise, in terms of ICC and SEM. Test-retest reliability for testing at home could not be assessed, but the higher variability in domestic results suggests that this is somewhat poorer than in the lab (see variability in Figure 6.4).

The test-retest measurements in low-pass filtered masking noise show a small but significant learning effect of 0.5 dB that is comparable to the test-retest differences obtained in the evaluation study described in Chapter 5, and in other studies developing a speech-in-noise screening test (Smits et al, 2004; Vailancourt et al, 2005, Wagener & Brand, 2005; Jansen et al, 2010). Although small, this effect should be accounted for when performing consecutive tests for domestic hearing screening.

Effect of masking noise condition

For all subject groups, the SRTs in LP noise are somewhat lower (e.g. better) than in LPmod noise, consistent with the shape of the noise level distributions of both noises. But the average difference is significantly higher in the NH group than in HI-BD group (4.0 versus 2.3 dB, respectively). These findings are similar to the differences between LP and LPmod noises shown in the previous evaluation study (see Chapter 5). Apparently, the ability to benefit from the unmasking of the high frequencies is smaller for listeners with severe hearing impairment, since additional fluctuations in the noise have a less disturbing effect in this group. As a result, SRTs obtained in LP noise differentiate better between subjects groups than the SRTs in LPmod. There are no differences regarding the effects of any of the tested measurement parameters between both masking noise conditions.

Presentation mode

Earcheck was performed in diotic as well as monotic presentation mode, testing both ears consecutively. When testing monotically, a small difference of 0.6 dB between left and right ears is observed. This right ear benefit in speech recognition is often seen in speech-in-noise testing, and can be explained by left hemispheric dominance for speech and language processing (Kimura, 1961). Our results agree with differences of 0.6-0.7 dB reported in literature (Plomp & Mimpen, 1979a, McArdle et al, 2012). Also, it is known that speech discrimination under diotic conditions is superior to monotic listening (Kaplan & Pickett, 1981; Davis et al, 1990; Van Hoesel & Litovsky, 2011; McArdle et al, 2012). The diotic benefit in this study is 0.9 dB in the lab and 1.6 dB at home (see Table 6.4). Previous studies showed an improvement in SNR due to diotic listening ranging from 1.3 dB to 1.5 dB (Davis et al, 1990), for either sentences in noise (Plomp & Mimpen, 1979a; Bronkhorst & Plomp, 1998) or digit triplets (Smits et al, 2006a). Hence, the diotic advantage found in the lab experiment is smaller than expected.

Effect of test environment

The most important finding in this study is that the SRT results in domestic testing differ from the SRT results obtained under lab conditions. In agreement with the findings of Culling et al. (2005), a slight overall deterioration in domestic test results could be expected because of poorer testing conditions at home, attributable to possible room reverberation or ambient noise, and quality of used sound cards and headphones that is likely to be lower than in the lab. Accordingly, the SRTs of the normal-hearing group are on average 1 dB worse when the test was performed at home rather than in the lab. Conversely, a positive effect of domestic testing is seen in both hearing-impaired groups. Although this overall effect is small in HI-ND subjects (0.8 dB), HI-BD subjects show an average benefit of almost 2 dB.

In order to gain more insight into the origin of these dissimilarities, we calculated the individual differences between the SRTs obtained in the lab (results were calculated separately for monotic and diotic conditions tested at 65 dB and 77dB) and at home. These differences were plotted against the pure-tone average of the hearing thresholds at 3, 4 and 6 kHz ($PTA_{3,4,6}$) for the corresponding ear of the individual participants in Figure 6.6. This shows that the difference in SRT, in favor of domestic testing, becomes larger with increasing hearing loss. Apparently, the more severely hearing-impaired listeners show larger benefit from domestic testing, suggesting an increase in audibility of the high-frequency speech information when testing is conducted at home.

Though these effects need further analyses, we will discuss the possible role of differences between lab testing and domestic testing regarding spectral differences between the stimuli, the level of presentation, and the presence of background noise.

Spectral differences

The first possible factor of influence might be differences in the spectra of the test signal, possibly induced by different sound cards, PC settings, or the transfer characteristics of the transducers used. With well-matched speech and noise passing through the same playback device, the variation in devices will leave the signal-to-noise ratio in each frequency band unaffected (Culling et al, 2005), and intelligibility is mainly determined by SNR (Plomp, 1986; Wagener & Brand, 2005). However, part of the masking noise used in the Earcheck test investigated in the present study is removed by low-pass filtering, hence the specifications of domestic equipment may become relevant for the test outcomes and its discriminative power. Specific amplification of the less masked high-frequency region can improve audibility of the speech stimuli and hence test performance.

Headphone presentation, instead of loudspeaker use, offers the advantage of monotic testing, eliminates reverberation, and reduces ambient noise. The use of the

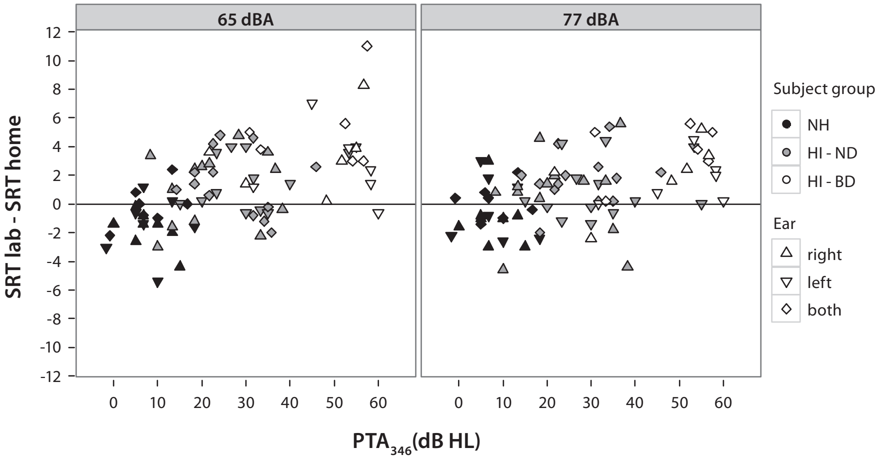


Figure 6.6. Individual differences in SRT obtained in the lab and at home plotted against their $\text{PTA}_{3,4,6}$ for results obtained in LP noise. A positive difference indicates a superior SRT home relative to SRT lab. Left panel shows the differences relative to lab results at 65 dBA, right panel shows differences relative to lab results at 77 dBA. Black symbols represent the results of the NH subjects, grey symbols denote the HI-ND group and white symbols represent the NI-BD group, for results of the left ear (▽), the right ear (△) and the diotic condition (◇).

same headphones in this study reduces one of many sources of spectral variability in speech reception tests at home. But the headphones used at home and in the lab differ both in quality and in frequency response. As a consequence, a systematic difference in SRT results cannot be excluded. Although the type of standard headphones was selected for its relatively flat frequency response, a slight difference in frequency response could have been more beneficial for the hearing-impaired groups if it amplifies the high-frequency region. By using the same headphones in the current study, the possible effect of differences in frequency responses is not established, but should be kept in mind when interpreting domestic Earcheck results.

Level effects

Another factor may be that the stimuli at home were presented at a higher, self-chosen, presentation level than used in the lab. It is likely that particularly hearing-impaired respondents increase the presentation level to maximize their performance. This is confirmed when comparing the mean individually set presentation levels of the different groups recorded preceding the lab tests: 64.6 dB for NH versus 70.6 dB for

HI-BD. This difference in presentation level may be even more pronounced when the test is performed in the presence of ambient noise in a non-isolated environment.

Previous studies reported that speech intelligibility in interrupted noise improves with increasing presentation level (Stuart & Phillips, 1997), due to increased audibility and steeper forward masking slopes at higher noise levels (Rhebergen et al, 2010b). Whereas the effect of the latter factor only applies to intelligibility in LPmod noise, audibility is an important factor in speech reception in both noises. Because of the high masking release in the low-pass filtered noises, relatively low SNRs are reached, and as a result speech at threshold is presented at low levels. Consequently high-frequency speech information may fall below the elevated hearing thresholds of the HI subjects, who may be expected to experience more benefit as presentation level increases (Summers & Molis, 2004; Rhebergen et al, 2010b). Indeed, the lab results of experiment A show that presentation level has no effect on SRTs in NH and HI-ND subjects, but HI-BD subjects have a 1 dB benefit for presentation levels exceeding 65 dBA. This finding demonstrates that audibility is improved at higher presentation levels for subjects with more severe high-frequency losses.

Presence of background noise

In order to assess the effect of transducer type used on the SRT outcome of Earcheck, domestic diotic test results obtained with loudspeakers and headphones were compared. Previous findings of studies investigating domestic screening methods indicated an advantage of using headphones that ranged from 1.1 dB to 1.4 dB (Culling et al, 2005; Smits et al, 2006a; Van Son & Jellema, 2011), due to detrimental effects of poor listening conditions when loudspeakers are used. In this study, the expected poorer outcomes for loudspeaker testing are only shown in both hearing-impaired subject groups. Loudspeaker presentation produces a significant elevation of SRT of 2.6 dB averaged over all hearing-impaired participants. Normal-hearing listeners, however, do not show any significant effects of the transducer type used.

Differences between subject groups

Current study aimed to investigate the effects of uncontrollable parameters in domestic speech-in-noise testing. To assess whether effects were different for NH listeners and subjects with NIHL, three subject groups were examined. All the fitted models (presented in Tables 6.2, 6.4 and 6.5) show that SRTs increase with increasing hearing loss, but only the mean SRT of the HI listeners with severe hearing loss is significantly poorer than the mean SRTs of the other groups. Yet, SRTs of the mildly hearing impaired group do not differ significantly from normal-hearing performance, indicating a limited ability to differentiate between NIHL and normal performance. Deviant from the first evaluation study (see Chapter 5), current subject groups were not selected on their hearing, but composed by classifying the partaking construction workers, with a continuum of hearing threshold levels, after their inclusion in the study. Subjects were considered normal-hearing when all pure-tone thresholds were below or equal to 20 dB HL. As a result, several near normal-hearing subjects, with only slightly elevated audiometric thresholds at 1 or 2 frequencies, are included in the HI-ND group. Although these are the subjects particularly of interest in case of screening purposes, this might explain the insignificant differences found in SRTs between these groups. Our current results show that Earcheck is not sensitive enough to detect these very mildly hearing impairments.

General applicability

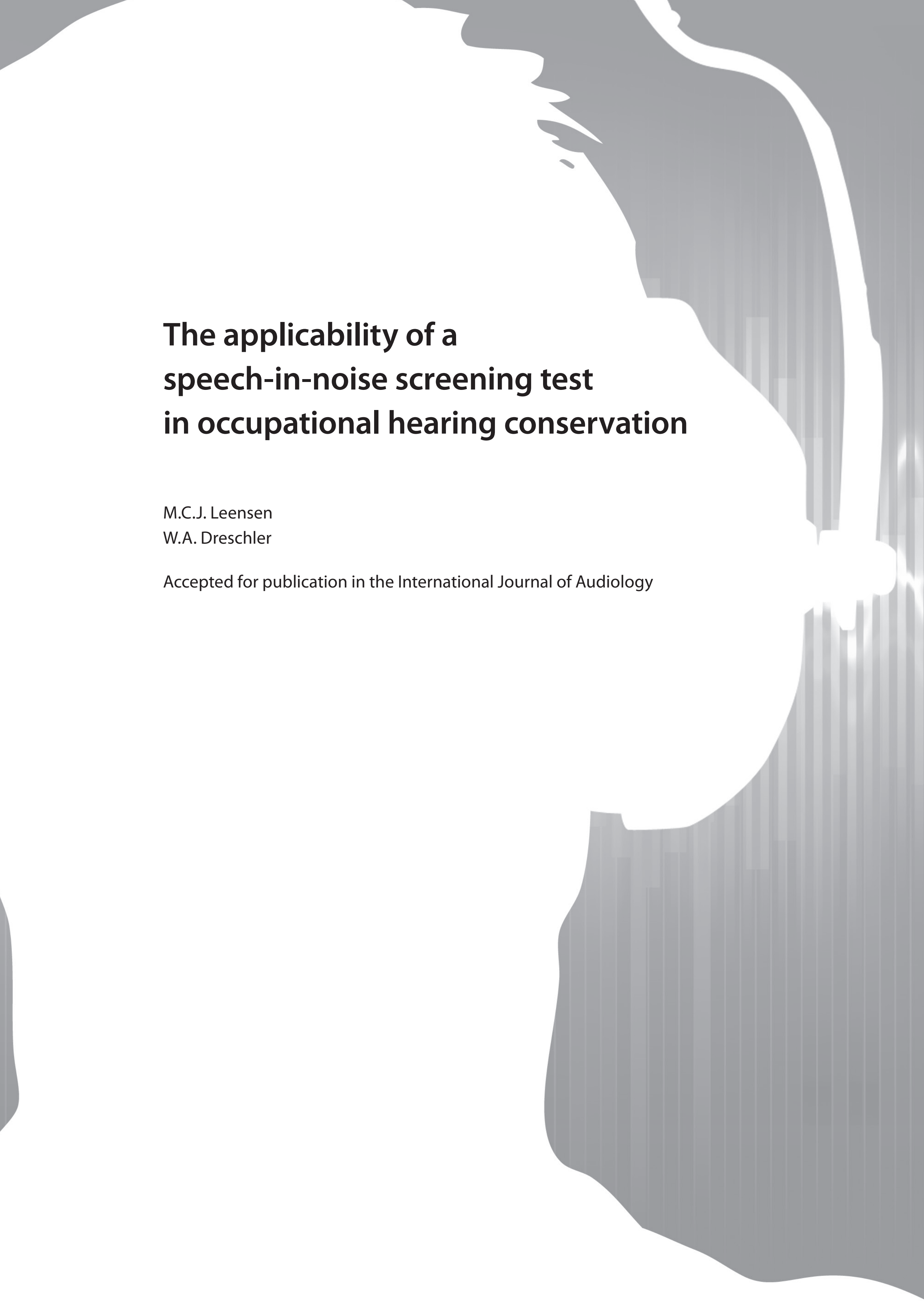
Overall, LP and LPmod noises used in this study are equally affected by the different test parameters investigated in the experiments, meaning that both masking conditions can be employed in Earcheck for domestic screening purposes. Considering the greater inter-individual spread in SRT results in LP noise, and the slightly better reliability results for this noise condition, the stationary LP noise is considered the best alternative for online NIHL screening.

Although better SRTs are obtained in diotic testing, monotic testing using headphones has the great advantage of testing each ear separately. This produces a more comprehensive assessment of the respondent's hearing ability, particularly in the case of a mildly or moderately asymmetric hearing losses. Also, headphones are preferred to eliminate the acoustic effects of test environment. If one does use Earcheck for diotic screening in domestic testing, the diotic benefit of 1.6 dB should be taken into account.

Our results do not show any differences between the NH and the HI-ND groups, and significant differences between the NH and HI-BD groups found in the lab experiment decrease when testing was done at home; the normal-hearing listeners show reduced SRTs for domestic testing, whereas the both HI groups improve their performance. Now that domestic testing turns out to affect SRT results, actual results of Earcheck performed at home by a larger study population than currently tested should be compared to pure-tone audiometry, in order to assess Earcheck's validity and reliability when performed at home.

Nevertheless, an online self-test such as Earcheck provides a fast and easy way to reach many people, which makes the test highly applicable for adult hearing screening. When the discrimination between normal and impaired hearing could be improved by taking into account the variability arising from domestic test administration, Earcheck can be considered as a valuable test for NIHL screening at home.

7



The applicability of a speech-in-noise screening test in occupational hearing conservation

M.C.J. Leensen
W.A. Dreschler

Accepted for publication in the International Journal of Audiology

Abstract

Objective: Noise-induced hearing loss (NIHL) is the most reported occupational health disease in the Netherlands. The internet-based speech-in-noise test Earcheck is designed to detect beginning NIHL and can be a valuable tool in occupational hearing health surveillance. The aim of this study is to investigate the validity of Earcheck compared to regular screening audiometry.

Design: Subjects performed online Earcheck tests at home. The results were compared to a pure-tone screening audiogram obtained during regular occupational health examination. A subgroup performed the measurements twice to assess test-retest reliability.

Study sample: 249 male construction employees who recently had a periodic occupational health examination participated.

Results: An average learning effect of -1.6 dB was found, that reduced with increasing test number. The test-retest variability was 1.6 dB. Sensitivity to detect beginning NIHL was 68%, with a specificity of 71%.

Conclusions: Although sensitivity and specificity values are only moderate, the broad internet application still promises a valuable addition to current practice. The relatively high learning effect indicates that more reliable results can be obtained after a longer test session. When this is put into practice some improvement in sensitivity and specificity may be expected as well.

Introduction

Noise represents one of the most common environmental health hazards. Long term exposure to high daily noise levels may cause noise-induced hearing loss (NIHL), a permanent sensorineural hearing impairment (ISO-1999, 1990). In modern society, large groups of people are frequently exposed to noise, either during recreational activities or in occupational settings. In the past years NIHL was the most reported occupational disease in the Netherlands (Van der Molen, 2010; 2011), especially in sectors with a high concentration of noisy occupations such as the construction industry. In 2012, 12.9% of the workers employed in this sector reported NIHL (Arbouw, 2012). Internationally, a large US analysis showed that the construction industry had the highest number of workers with self-reported hearing impairment attributable to their employment (Tak & Calvert, 2008).

NIHL can mostly be prevented by reducing or eliminating the exposure to noise. Hearing conservation programs in construction mainly focus on employee's use of hearing protection devices (HPDs) rather than on controlling the noise exposure at its source (Neitzel & Seixas, 2005). Yet few workers use hearing protection consistently enough to prevent hearing damage (Lusk et al, 1998; Hong et al, 2005). NIHL affects the high frequency region (3-6 kHz) and develops gradually (Rösler, 1994). Consequently, it is often unnoticed, especially when hearing loss is still in an early stage (Vogel et al, 2009; Shah et al, 2009). When the damage is substantial and severe enough to be measured, hearing impairment is irreversible. Hence it is of major importance to assess the possibility of hearing damage as early as possible, so precautionary measures can be taken to prevent further development of NIHL (Meyer-Bisch, 1996).

The early detection of hearing loss is thus a crucial aspect of hearing conservation. Periodic testing enables monitoring of the hearing ability of employees at risk, and may help to prevent further development of NIHL. Workers that have been demonstrated to have reduced hearing ability can be motivated better to use HPDs properly (Zohar et al, 1980). The European Directive 2003/10/EC states that employees whose daily noise exposure exceeds the lower exposure action value of 80 dBA have the right to preventive audiometric testing (EPC, 2003). In the Netherlands, pure-tone screening audiometry is incorporated in the periodic occupational health examinations (POHE), to which all construction employees are invited once in at least every 4 years. Participation in this POHE occurs on a voluntary basis.

However, audiometric testing is not always a viable option for construction workers because it is logistically difficult and time-consuming. Despite the attention given to occupational health in the Dutch construction sector, 40-50% of the employees invited for a POHE do not respond to their call. This relatively high non-response reflects the difficulty in attaining the entire population at risk, because of the widespread and very transient workforce in this sector that is characterized by many small companies, subcontracting, and mobility between workplaces.

Additionally, there are some requirements to assure reliable and valid outcomes in pure-tone audiometry that are not always met in occupational testing, which for practical reasons is often performed at the company or worksite. First, an accurate audiogram requires adequate acoustical isolation, but a sound proof booth is not available in most cases. As a result, possible ambient noise may negatively affect threshold determination. Second, test administration is mostly done during a working day, without a predefined noise-free period prior to testing. Although respondents are advised to wear HPDs prior to examination, hearing threshold assessment might be biased by a temporary threshold shift due to recent noise exposure. Third, this method requires a well-calibrated audiometer and qualified test administrators, making it a costly and time-consuming test method. Therefore alternatives for hearing screening in occupational health need to be found.

Since the effects of NIHL are typically experienced in challenging listening situations such as in background noise, speech-in-noise testing may be a valuable alternative for hearing screening in occupational health. This has the advantage that all stimuli are presented suprathreshold, which is less demanding with regards to the test environment. This makes a quiet room adequate for accurate testing, and there is no need for an isolated testing booth. Speech-in-noise testing can be automated using an adaptive procedure, so it can be used as a self-administered test (Smits et al, 2004; Jansen et al, 2010). Furthermore, calibration can be controlled programmatically and it is relatively insensitive to absolute presentation level or small variations in equipment (Plomp, 1986;Smoorenburg, 1992; Smits et al, 2004; 2006a).

Due to these characteristics, a speech-in-noise test is very suitable for internet-based application. Using the internet to distribute a screening test entails the possibility of evaluating hearing status remotely at home and provides a fast way to reach many employees at risk (Stenfelt et al, 2011). As a result, online hearing screening may lead to higher participation rates in the transient workforce of the construction industry.Moreover, it requires no specialized equipment, can be performed more easily after a period free of occupational noise (e.g. during weekends), and hearing status can be tested more frequently than once every four years, e.g. when complaints arise. So, an easily accessible speech-in-noise test can be a valuable addition to current practice of occupational hearing screening in the construction industry.

In the Netherlands, such a test was developed by the National Hearing Foundation in association with the Leids University Medical Centre; Earcheck (Albrecht et al, 2005). Earcheck is an online speech-in-noise test that measures the ability to understand words in noise by determining the signal-to-noise ratio (SNR) that corresponds to 50% intelligibility. The original version of this test used a stationary broadband noise and was not sensitive enough to detect mild-to-moderate NIHL (see Chapter 4). By modifying the broadband masking noise this sensitivity increased strongly; by using a low-pass filtered noise that forces the listener to use high-frequency speech information, the test discriminates well between normal and impaired performance due to NIHL. Lab results show that the test has a high sensitivity and specificity of 95% and 98% respectively (see Chapter 5). However, these promising findings were based on results obtained under well-controlled conditions in the laboratory. The question is how this holds when the online test is broadly applied for remote testing in occupational health, when there is lack of control over environmental variables and testing conditions such as settings, presentation level or transducers used.

A recent evaluation study showed that the influence of these parameters on outcomes of domestic Earcheck testing were small, but significant (Chapter 6). Normal- hearing performance at home was on average 1 dB poorer than in the laboratory, whereas hearing-impaired subjects showed a significant beneficial effect of 1.5 dB from domestic testing. These findings indicate that Earcheck's laboratory-based cut-off points need to be refined for adequate discrimination in domestic screening application. The evaluation study also showed that Earcheck results were significantly affected by the chosen presentation mode and transducer type. Although diotic presentation yielded 1.5 dB better results than monotic listening, monotic presentation is recommended since it has the great advantage of testing each ear separately (see Chapter 6). Intrinsic to monotic presentation is that testing should be done via headphones. This choice is also favorable because test outcomes are more reliable when measured with headphones than using loudspeakers (Smits et al, 2006a; Van Son & Jellema, 2011).

The results of Chapter 5 and 6 indicated that two types of low-pass filtered masking noise showed good applicability; a stationary low-pass filtered noise and a stationary low-pass filtered noise combined with a high-pass interrupted noise that replaces the removed high-frequency part of the noise. These additional fluctuations place demands on temporal resolution of the listener, and might increase the discriminative power between normal and impaired speech recognition performance even more. The current study compares both masking noises when applying the online Earcheck in a noise-exposed population, to come to the best alternative for domestic hearing screening.

The aim of this study is to investigate the value of Earcheck in identifying (early) noise-induced hearing loss in addition to regular pure-tone screening audiometry. To do so, Earcheck results of a large population of construction workers are compared to their screening audiograms. The obtained test reliability, sensitivity, and specificity in this occupationally exposed population will determine the value of Earcheck as a screening tool for occupational NIHL.

Methods

This investigation consisted of a cross-sectional comparison of data from construction workers. Survey data collection was completed through internet-based speech-in-noise testing. Pure-tone audiometric data were collected as part of regular periodic occupational health examinations and were provided by occupational health services. The experimental protocol and all procedures in this study were approved by the ethics committee of the University of Amsterdam (approval number: 2001_187).

Participants

The participants in this study were employees of the construction industry, who were possibly exposed to occupational noise and therefore were at risk of developing noise-induced hearing loss. The eligible study population included only male employees aged 18 years or older, including office workers. Female workers comprised only a small percentage of the total workforce in construction (3.5%), and they were discarded because of their relatively high concentration in non-noise-exposed jobs. Since outcomes of the online Earcheck were to be compared to the regular occupational screening audiogram as performed during a periodic occupational health examination (POHE), only subjects who recently had such an examination were selected for participation.

The selection procedure was carried out by Arbouw, the Dutch national institute on occupational health and safety in the construction industry. This organization collected data from medical records of POHEs of all employees in the construction sector throughout The Netherlands.

First, all medical records from POHEs conducted within 3 months prior to the selection date were selected. Records with invalid or missing audiometric data were excluded from the selection, as were subjects reporting known hearing problems due to other etiologies than noise exposure. Then, three random samples of approximately 1.000 subjects were drawn from all selected records, according to a weighing procedure based on age. Since the mean age in the total population of construction employees is 40 years, the selected population was stratified into three age groups, 18-34 yr, 35-44 yr and 45-64 yr, to get a proportional representation of age groups in the study population. The 2.937 selected employees were invited to participate by sending them an invitation letter.

Pure-tone audiometry

Screening audiograms were assessed during a periodic occupational health examination. POHEs are provided for all employees in the construction industry, irrespective of occupational noise exposure. Pure-tone audiometry was conducted at the workplaces, if possible in a mobile unit equipped with a soundproof booth. Manual audiometers coupled with TDH-39 headphones were used. These audiometers were annually calibrated according to the ISO-389.1 (1998) standard. Pure-tone air-conduction thresholds were determined for both ears at frequencies 0.5, 1, 2, 3, 4, 6 and 8 kHz, in 5-dB steps. Testing was done during the work shift, but subjects were advised to wear HPDs on the day of testing if they had to work in a noisy environment.

Earcheck

In the Netherlands Earcheck was developed as an automatic online speech-in-noise test (Albrecht et al, 2005) that was adapted to be used for NIHL screening (see Chapter 5). In brief, it consists of nine different monosyllables, randomly presented in a low-pass filtered interfering noise. On screen, nine response buttons are shown. A tenth button, saying ‘not recognized’, is added to prevent respondents from guessing.

After the presentation of a word, the subject’s task was to identify the word he had been presented by clicking on the corresponding button on the computer screen. The level of the noise was fixed and the level of presented words varied according to an up-down procedure, with a step size of 2 dB. This was based on the testing procedure according Plomp & Mimpen (1979a), except for the first stimulus of Earcheck that was presented only once at a fixed SNR of -10 dB. A list of 27 presented words was used to estimate the signal-to-noise ratio at which 50% of the speech material was reproduced correctly. This was defined as the speech reception threshold (SRT), and was calculated by taking the arithmetic average of the SNRs of the last 20 presentations. The masking noise used in this test was a low pass filtered noise, either without (LP) or with temporal modulations in the high-frequency part (LPmod). Both masking noises have been derived by digitally filtering a stationary broadband noise with a long term average spectrum similar to that of the speech stimuli, using a low-pass filter with a cut-of frequency of 1.4 kHz and with a steep roll-off slope (100 dB/oct). To generate the LP noise, a noise floor of -15 dB was added after filtering. In order to create the LPmod noise, the low-pass filtered noise was combined with a high-pass filtered noise that was modulated by a 16-Hz square wave with 50% duty cycle and modulation depth of 15 dB. Further details of the development of noise stimuli can be found in Chapter 5.

Procedure

Subjects were recruited by sending them an information letter about the study and inviting them to participate. Subjects willing to participate were asked to log in to a secured website of Arbouw, to sign an informed consent and give permission to retrieve audiometric data from their medical record for the purpose of this study. After this, they were automatically led through a short online questionnaire that serves as a first screening for including adequate participants. Respondents were asked to complete questions regarding noise exposure at work and during leisure time, hearing ability and native language. Reported job title was used to estimate daily noise exposure levels of individual construction workers; time weighted average (TWA) noise exposure levels for standardized job titles were extracted from a database of Arbouw (Arbouw, 1998; for more details see Chapter 2). Registered subjects were assigned a random identification number and were sent an instruction letter and e-mail. The letter was accompanied by a standard low-cost pair of headphones (HQ, type HP 113 LW) to assure that every participant had one in order to conduct the domestic screening test properly. Subjects were asked to perform Earcheck at home, using their own personal computer with corresponding settings. They used their identification number to log in to a special experimental website to perform the online speech-in-noise tests.

Standard Earcheck administration procedure was utilized in this study. At the start of the test session, a word without noise was presented repeatedly, and participants used their PC volume control or a slider on screen to adjust the volume to a level at which the presented word was clearly understandable. This presentation level was used for the presentation of all consecutive test stimuli. To allow subjects to become acquainted with the stimuli, a preliminary sequence of the nine different test words was presented in noise preceding the actual tests.

All testing was done monaurally. Both ears were tested consecutively, and the ear to be tested first was chosen randomly. In addition, both ears were tested using LP and LPmod masking noise, resulting in four different test conditions per test session. The masking noise condition was counterbalanced across subjects. To assure this balanced design of the tests, the four test conditions were programmed to appear in the correct order and participants were automatically led through the complete testing procedure[5]. Participants were instructed to perform the first Earcheck during the weekend (preferably on Sunday) to prevent occupational noise exposure from biasing the SRT results. Reminders were sent to participants that did not perform their test on the given date, until the test was completed. To assess Earcheck's reliability when used for domestic screening, a subpopulation of 32 employees performed the test in test and retest. The repeated test session was performed on the same day as the first test session.

Statistical analysis

The data were analysed using SPSS (version 19.0) and R software (R Foundation 2008, from http://www.R-project.org).

To account for effects of repeated measurements within both ears of each individual, linear mixed effect models were used to estimate the effects of different

---

[5] Subjects that logged in with an odd identification number performed the test with the right ear in LP noise first and even numbered subjects started with their left ear in LPmod noise.

masking noise conditions, tested ear, and test repetition on SRTs. This method accounted for the nested data structure of repeatedly testing two ears within one participant, and also handled missing data that were introduced in the data due to some incomplete measurements. The variations between both 'subjects' and 'ear' were treated as random effects. Fixed factors of primary interest were 'masking noise' and 'ear'. Since the systematic difference between both masking noise conditions was known, the primary interest of this study was whether shown effects are similar in both masking conditions. In order to investigate this, two-way interaction terms between each of the predicting variables and 'masking noise' were incorporated in the model as well. Only the factors and interaction terms that showed a significance contribution to the fitted model, tested with conditional F-tests at the 0,05 level, were investigated for significant coefficients of each level. When coefficients proved to be significant, the term was retained in the model.

To assess test reliability, a similar linear mixed effects model was fitted to the data of the subpopulation, also incorporating the fixed factor 'repetition' in the model, to assess systematic differences in SRT outcomes between test and retest sessions. Test-retest reliability can be expressed by two different parameters; the intraclass correlation coefficient (ICC), a relative index of reliability representing the ratio of the between-subject variability to the total variability in the data, and the standard error of measurement (SEM), a measure of absolute consistency reflecting the precision of individual outcomes (Weir, 2005). Two-way random, absolute agreement, single measures ICCs were calculated over the test and retest results within each masking noise condition. This type of ICC takes the systematic difference between test sessions into account and yields results that can be extrapolated to other situations (Weir, 2005). SEM can be derived from ICC, according to the following equation:

$$\text{SEM} = \text{SD}\ \sqrt{(1\text{-ICC})} \qquad \textit{(Equation 7.1)}$$

where SD is the standard deviation of the test and retest scores from all subjects.

However, when there are two levels of trials, there is an alternative way of calculating the SEM, by dividing the SD of the differences (SDdiff) resulting from paired t-testing by $\sqrt{2}$.

Finally, the obtained SEM was used to define the minimum detectable difference (MDD), which can be considered as a real change in a subject's score, above measurement error. MDD was calculated from SEM, according to

$$\text{MDD} = \text{SEM} * 1.96 * \sqrt{2} \qquad \textit{(Equation 7.2)}$$

creating a so-called 95% confidence interval of change.

Pearsons product correlation coefficients were calculated to determine if there was an association between Earcheck results and hearing thresholds of the screening audiogram.

To further explore the relationship between Earcheck and pure-tone thresholds, linear regression was performed. Since this analysis compared two test methods, Deming's regression was used. Whereas the ordinary linear regression method assumes that only the independent measurements are associated with measurement errors, the Deming method (Deming, 1964) takes measurement errors for both methods into account. To use this technique it was necessary that the ratio of the variances of measurement error of the two test methods was known. The measurement errors for the Earcheck using either LP or LPmod noise were derived from the test-retest analysis. For the PTA-value obtained by a screening audiogram, a value of 5 dB could be taken (Helleman & Dreschler, 2012).

Receiver operating characteristics (ROC) curves were calculated to establish the correct cut-off values of Earcheck and assess corresponding test sensitivity and specificity in detecting NIHL. Sensitivity is the ability of Earcheck to detect participants that actually show audiometric hearing loss, and specificity is the ability of Earcheck to detect absence of hearing loss in those showing no elevated hearing threshold levels.

Results

Participants

In total, 256 male construction employees signed up to participate in this study. After application, seven participants did not complete the online hearing tests. Of the 249 participants that did complete the Earcheck session, nine were excluded from analysis based on their answers in the online questionnaire; three of them reported otological problems other than NIHL or presbyacusis, two used hearing aids and four were not native speakers of the Dutch language.

In addition, 30 participants performed their first hearing test at the end of a working day, instead of during the weekend. Since possible temporary effects of noise exposure on their speech intelligibility cannot be ruled out entirely, these subjects were omitted as well. This leaves 210 participants included in the analyses. The distribution of audiometric hearing threshold levels of this population of 210 participants is displayed in Figure 7.1.

Participants were predominantly middle-aged (mean age 45.7 years, SD = 10.0 years). The majority of them, 83.3%, were exposed to job related daily noise levels exceeding 80 dBA. Their mean job tenure in construction was 25.8 years (range 1 – 46 years,

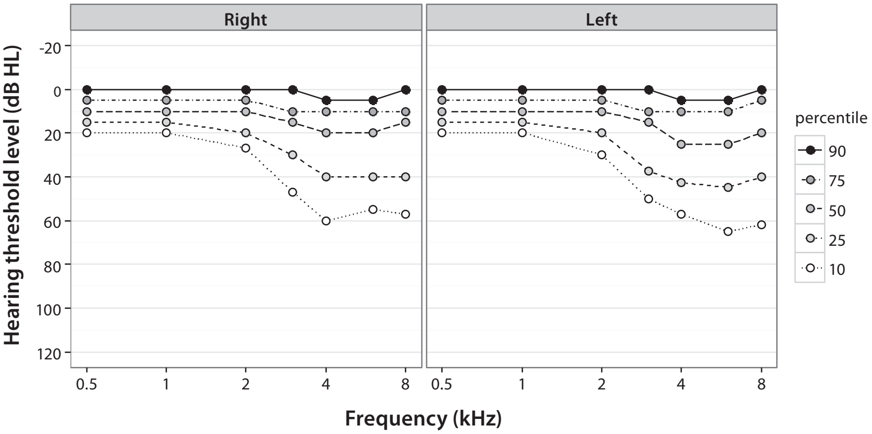


Figure 7.1. Audiometric thresholds of the study population.

SD = 11.4 years). Two-thirds of the study population reported to often wear their HPDs and only 10.5% never used HPDs. About 74.2% of the study participants indicated that their hearing ability wasn't good, and 19.1% reported to have tinnitus. In addition, 70.3% of the participants experienced some problems understanding speech in noise, and 27.7% had also problems in quiet.

Of the total population, 32 subjects conducted a test and retest session on the same day to assess test-retest reliability: the test-retest subgroup. T-tests and chi-square analyses indicated that there were no significant differences in characteristics between the subgroup and the total population.

SRT outcomes

To assess possible systematic differences in SRT results between masking noise conditions and ears linear mixed effect models were run with random effects for 'subject' and 'ear', and fixed effects for 'masking noise' (LP, Lpmod) and 'ear' (left, right). A two-way interaction between these fixed effects was also included in the model, but this did not significantly contribute to it (F[1,403] = 3.31, p = 0.0696). 'Masking noise' significantly affected SRT outcomes (F[1,403] = 1109.64, p < 0.001), as was expected beforehand. On average, SRTs obtained in low-pass noise were 4.6 dB better than SRTs obtained under low pass filtered modulated noise, reflecting the differences in SRT between masking noise found in earlier studies described in Chapters 5 and 6. 'Ear' did not show any significant contribution to the model (F[1,207] = 1.69, p = 0.194). Since significant right-left differences were not found, data were pooled into a group of 420 ears for further analyses.

Test-retest reliability

A subpopulation of 32 participants, and thus 64 ears, performed test and retest measurements on the same day. The mean test and retest results of this population in both noise conditions are shown in Table 7.1.

Learning effect

The difference in SRT over the two test sessions was calculated. This difference estimated a systematic change in performance due to effects of either learning or fatigue. The results in Table 7.1 show that the performance on the second test was better than on the first, in both masking noises. A linear mixed effect model was fitted to the test-retest data with random effects for 'individual ear', and fixed effects for 'masking noise', 'ear' and 'repetition' (test, retest). Two-way interactions between the fixed effects were also included in the model, but these showed no significant contribution to the model. The factors 'masking noise' (F[1,187] = 360.17, $p < 0.001$) and 'ear' (F[1,62] = 0.07, $p = 0.786$) showed effects similar to those found in the total study group.

Table 7.1. Mean SRTs and test-retest characteristics of Earcheck in both masking noise conditions (n=64 ears).

	SRT 1 *mean (SD)*	SRT 2 *mean (SD)*	SRT diff *mean (SD)*	ICC	SEM	MDD
LP	-17.3 (3.4)	-18.7 (2.7)	-1.4 (2.3)	0.65	1.63	4.49
LPmod	-13.0 (2.7)	-14.8 (2.9)	-1.8 (2.4)	0.54	1.68	4.61

ICCs are two-way random model, absolute agreement, single measure.

The model showed that 'repetition' (F[1,187] = 51.62, $p < 0.001$) significantly affected SRT results; test and retest outcomes differed by 1.6 dB averaged over all testing conditions. The insignificant interaction terms of 'repetition' with either 'masking noise' or 'ear' indicated that the learning effect was similar for both ears and in both masking noises.

The results of the test and retest sessions were compared against each other in Figure 7.2; the number of datum points laying below the diagonal line representing absolute agreement indicated the better SRT results of the second test. The intra-session Pearson correlation coefficients of 0.74 for LP noise and 0.65 for LPmod noise showed reasonably good agreement between test and retest results.

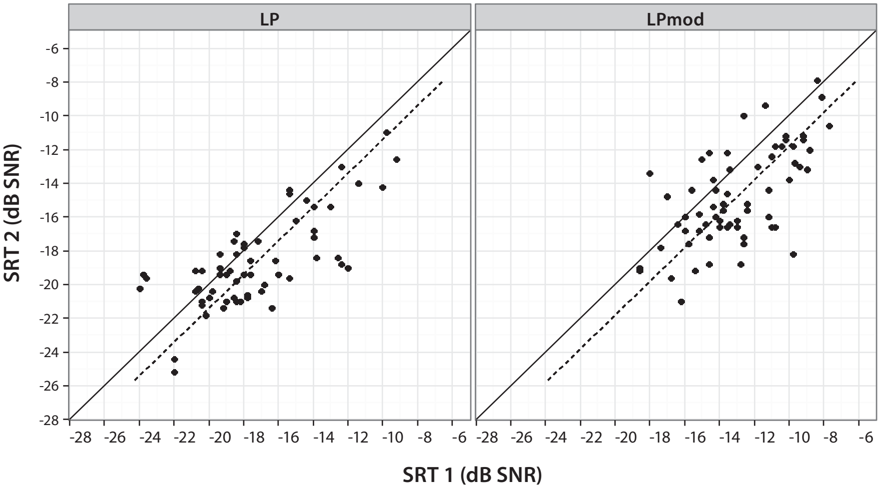


Figure 7.2. Domestic Earcheck results of the retest are plotted against results from the first test, separated for masking noise. Solid straight lines represent perfect agreement, dashed lines indicate the learning effect in each masking noise condition.

To explore the observed learning effect of -1.6 dB in more detail, the effect of practice over the subsequent trials was investigated. To do so, the improvements in listeners' performance over the eight tests during their test and retest session were assessed. Figure 7.3 shows the mean SRT at the particular position during the test, as a function of the trial number (1-8) during the test sessions, determined for both noise conditions separately.

In order to characterize this learning effect, an exponential curve, described by Rhebergen et al. (2008), was fitted to the data according to the following equation:

$$SRT(n) = SRT_{final} + a*2^{-(n-1)/N} \qquad \textit{(Equation 7.3)}$$

where *n* denotes the trial number (n = 1,...,8), SRT_{final} denotes the SRT value that is reached after the learning has leveled off, *a* is the size of the total learning effect (expressed in dB), and *N* is the average amount of tests that is required to halve the size of the learning effect. SRT_{final}, a, and N are free parameters that have been estimated by a least squares fit to the individual data.

The exponential curve was fitted for each of the two masking noise conditions, and the estimated parameters are displayed in Table 7.2. The amount of learning (*a*) was estimated to be 2.6 dB in LP noise and 3.5 dB in LPmod noise.

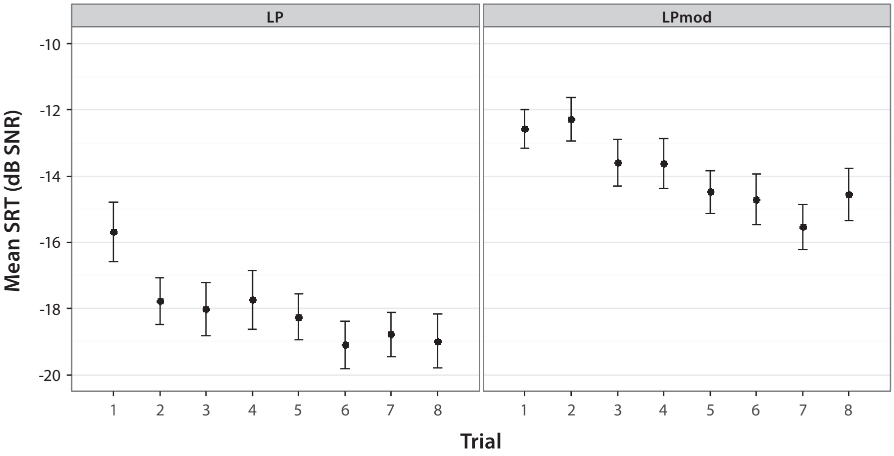


Figure 7.3. Mean SRT as function of trial number in both test sessions for the test-retest subgroup. Mean SRT is calculated across combinations of listeners at that trial, and for both noise conditions separately.

For LP noise, the parameter *N* was 1.3, which means that with an observed SEM of 1.6 dB, about 1.9 trials are required to bring the SRT within one SEM of its final value. In addition, about 2.8 trials are needed to bring the SRT within one decibel of the final SRT. Since the learning effect *a* was larger in LPmod noise, more trials would be required to reach final SRT; the amount of trials required to bring the SRT either within one SEM or one dB of the final SRT value were 2.6 and 3.5 respectively.

Table 7.2. Parameter estimates obtained by the least squares fit to the individual data, for both noise conditions.

	LP	LPmod
SRT_{final} (dB)	-19.47	-15.86
a (dB)	2.57	3.48
N	1.32	1.40

Test-retest reliability

An important measure to evaluate a specific measurement procedure is the test-retest reliability; the consistency of a test's results across series of observations. Two-way random, absolute agreement, single measures ICCs were calculated over the test and retest results within each masking noise condition (Table 7.1). For the stationary low pass filtered noise an ICC of 0.65 was found. For Earcheck using low pass filtered modulated noise, the ICC was somewhat lower, 0.54.

The obtained standard errors of measurement for both masking noise conditions were comparable (Table 7.1). The minimum difference to be real, derived from this SEM, indicated that a difference between two Earcheck measures should be at least 4.5 dB to be considered as a real difference in speech intelligibility (Table 7.1).

Validity

Test validity relates to the correlation between the test's results and other, accurate, measures of the same behavior. To assess Earcheck's validity, the SRT results obtained during the first test session were compared to single frequency hearing threshold levels of the corresponding ear, as obtained by routine screening audiometry. In addition, the pure-tone average (PTA) over the noise-sensitive frequencies 3, 4 and 6 kHz was calculated ($PTA_{3,4,6}$) and compared to SRT results. In Table 7.3 the Pearson's correlation coefficients for the SRTs obtained with Earcheck and the different audiometric parameters are given. All coefficients showed significance at 0.001 level. Significant positive linear associations between speech-in-noise intelligibility measured with Earcheck and audiometric hearing thresholds were observed. The association became stronger with increasing frequency, since the higher frequencies are more prone to noise-induced damage. The correlation coefficients indicated only moderate relations, which was best for Earcheck with LP noise compared to $PTA_{3,4,6}$ (r = 0.58).

Table 7.3. Pearson correlation coefficients of Earcheck outcomes and hearing thresholds (HTLs) at different frequencies from 1 to 6 kHz and $PTA_{3,4,6}$. All correlation coefficients are significant at p-value <0.001.

	HTL 1kHz	HTL 2kHz	HTL 3kHz	HTL 4kHz	HTL 6kHz	$PTA_{3,4,6}$
LP	0.177	0.397	0.532	0.532	0.482	0.575
LPmod	0.192	0.321	0.438	0.471	0.422	0.496

The sensitivity and specificity of Earcheck to detect NIHL relative to the screening audiogram depend on the cut-off values that were used to distinguish between normal and impaired hearing. First, $PTA_{3,4,6}$, was used as the audiometric reference measure, and hearing loss was defined as having $PTA_{3,4,6}$ > 40 dB HL. Then the relationships between the sensitivity and specificity of Earcheck were explored in more detail by calculating the receiver operating characteristic (ROC) curve. The ROC curve exhibits the sensitivity versus 1-specificity) of the model for each possible threshold in the range of estimated probabilities. The area under the curve (AUC) reflects the overall discriminative value of the model (Table 7.4). This was slightly higher for the Earcheck in LP noise (0.82) than when LPmod noise was used (0.79). Then the SRT values of Earcheck above which a subject is classified as having impaired hearing were calculated using the Deming's regression equation and corresponding cut-off values were -15.1 dB for Earcheck with LP noise and -11.2 dB for Earcheck with LPmod noise. Details of the regression are presented in Table 7.4.

Table 7.4. Results from Deming's regression analysis, fitting the relationship between SRT and $PTA_{3,4,6}$, and from fitted ROC curves.

	Intercept (dB) (95% CI)	Slope (95% CI)	Cut-off $PTA_{3,4,6}$	Cut-off SRT	AUC	Se	Sp
LP	-20.71 (-21.22 to -20.20)	0.14 (0.12 to 0.16)	40	-15,1	0.82	62.4	86.7
			35	-15.8	0.79	59.3	83.9
			30	-16.5	0.80	67.7	82.0
			25	-17.2	0.78	67.9	71.0
LPmod	-14.83 (-15.33 to -14.34)	0.09 (0.08 to 0.11)	40	-11.2	0.79	67.1	71.2
			35	-11.7	0.76	67.3	68.3
			30	-12.1	0.76	70.7	65.5
			25	-12.6	0.74	74.4	58.8

Different criteria are used to define NIHL, and corresponding cut-off values of the SRT outcomes are calculated. Based on these cut-off values, and analysis of the ROC curve, the area under the curve (AUC), sensitivity (Se) and specificity (Sp) are obtained.

Applying these calculated cut-off values showed that Earcheck performed at home had a moderate sensitivity and specificity in detecting NIHL compared to the screening audiogram; respectively 62% and 87% in LP noise, and 67% and 71% in LPmod noise. The observed sensitivity and specificity also depend on the value of $PTA_{3,4,6}$ defined to classify NIHL, and the procedure described above was repeated for different cut-off values. The corresponding parameters are presented in Table 7.4, showing an increase in sensitivity to distinguish between normal and impaired hearing with decreasing PTA-value, although the area under the curve slightly decreased.

Since this study aimed to assess the applicability of Earcheck in NIHL screening purposes, the sensitivity to detect mild NIHL was of main interest. Beginning NIHL was defined as having $PTA_{3,4,6}$ exceeding 25 dB HL. Using the regression equation, cut-off values of -17.2 dB in LP noise and -12.5 dB in LPmod noise were established (see Table 7.4). For this criterion, Earcheck with LP noise showed a sensitivity of 68% and specificity of 71% in detecting beginning NIHL. For Earcheck with LPmod noise the observed sensitivity was higher, 74%, but the corresponding specificity was lower (59%).

The scatterplots of the outcomes of both Earcheck conditions versus the $PTA_{3,4,6}$ reference measure and the Deming's regression line are shown in Figure 7.4.

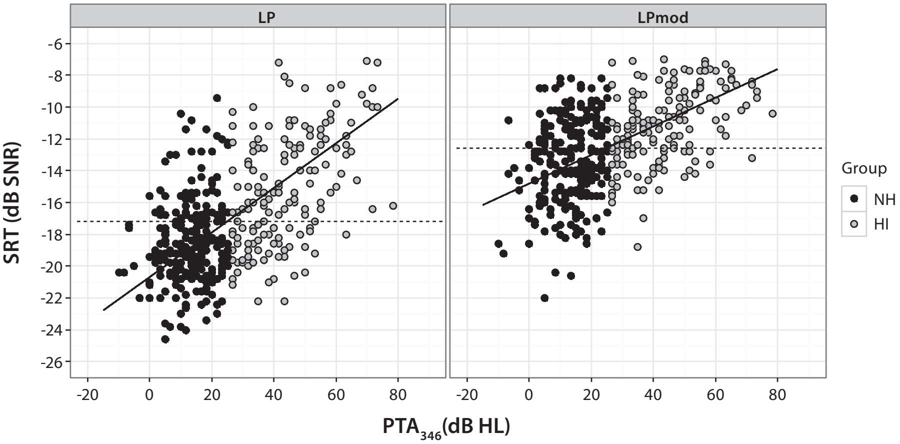


Figure 7.4. Scatterplots of SRT values against $PTA_{3,4,6,}$ for both masking noise conditions. The black symbols represent normal-hearing ears, grey symbols represent hearing-impaired ears ($PTA_{3,4,6}$ > 25 dB HL). Dashed lines represent the cut-off value for SRT that corresponds to hearing impairment. Solid lines represent the Deming's regression line.

Discussion

The internet-based Earcheck has been validated in this study against the current practice of pure-tone screening audiometry. The results of 210 construction workers at risk for noise-induced hearing loss demonstrate that domestic hearing screening by Earcheck has a moderate sensitivity to detect beginning NIHL; the observed sensitivity and specificity are 68% and 71% when stationary low-pass filtered masking noise is used, and 74% and 59% in low-pass filtered noise with modulations.

These values are lower than the test characteristics obtained in a previous evaluation study in which all measurements were conducted under well-controlled laboratory conditions (Chapter 5); those findings showed very high sensitivity and specificity of 94% and 92% in LPmod noise, and even higher values in stationary LP noise (95% and 98% respectively).

It is well known that speech intelligibility in noise is not perfectly related to pure-tone thresholds (Smoorenburg, 1992), because both methods assess different functionalities of hearing. Since 70.3% of the study population indicated to experience difficulties with understanding speech in noisy situations, testing speech intelligibility in noise has a higher relevance for hearing in daily functioning. Although there are large inter-individual differences, there is a weak relationship between the measured SRT values and self-reported speech intelligibility in noise. In the subgroup reporting no difficulties, the average SRT is -18.7 dB, in the participants who sometimes experience difficulties the average SRT is -17.2 dB and in the subgroup often having difficulties this is -15.0 dB ($F[2,403] = 33.20$, $p < 0.001$) in the LP noise condition. Nevertheless, pure-tone audiometry is generally considered the gold standard to which new hearing tests should be compared.

Several issues may be considered to have played a role in the moderate association between speech reception in noise and audiometric results observed in this study. First, field testing may incorporate effects of uncontrollable parameters that increase variability in both screening audiometry and domestic speech-in-noise testing. Second, some methodological considerations should be taken into account.

When Earcheck is administered at home, some uncontrollable parameters could affect test outcomes, such as presence of ambient noise, individually set presentation levels, and variety in the equipment used regarding its quality and frequency response. A previous study showed that the differences between normal and impaired performance were smaller at home than under well-controlled lab conditions (Chapter 6). Although an effect of presentation level could not be convincingly established, performing the test at a higher presentation level, might have been responsible for the average improvement of 1.2 dB in hearing-impaired domestic SRTs. On the other hand, unfavorable effects of remote testing were shown in the normal-hearing subjects, who had 1 dB poorer test results at home than in the lab (Chapter 6). In addition, the test-retest reliability in the current study shows that the variability of Earcheck testing at home of 1.63 dB is slightly greater than the SEM of 1.25 dB observed in LP noise in a clinical setting (Chapter 5). Also, industrial screening audiometry is less reliable than diagnostic audiometry obtained in a clinical setting, due to less controlled test conditions. Dobie (1983) observed that workers referred for otologic evaluation have hearing levels that were, on average, 5 dB better than indicated by screening audiometry. Additionally, test-retest variability in industry, expressed in SD_{diff}, was approximately 3 dB higher (6.7-8.3 dB up to 4 kHz) than reported for clinical audiometry (4-5 dB up to 4 kHz) (Dobie, 1983). In conclusion, both measurement methods compared here may show greater variability in the field than opposed to lab testing. Nevertheless, the quality of the test results in this validation study does reflect the quality of these measurements in daily practice (Grobbee & Hoes, 2009).

Another complicating factor is that for screening on NIHL the insensitivity of Earcheck for conductive hearing losses is to be preferred above the sensitivity of the gold standard, pure-tone audiometry. Listeners with conductive hearing loss (sometimes due to a recent flue or cold) will demonstrate elevated air conduction thresholds. Bone conduction thresholds, allowing for identification of conductive hearing losses, cannot be measured reliably in screening audiometry. As a consequence these patients will be referred for further diagnostic testing in order to assess potential NIHL. Earcheck results, however, will most likely indicate that this is not necessary. So, the discrepancy between Earcheck and screening audiometry for conductive losses is one of the reasons for poorer sensitivity and specificity values, while for this group the Earcheck outcomes may be expected to be more reliable than the gold standard of screening pure-tone audiometry.

Furthermore, some methodological issues that might have affected the observed test parameters should be considered. First of all, the observed sensitivity and specificity depend upon the study population used. Although sensitivity and specificity are considered to be more or less constant and not directly influenced by the prevalence of a disease, it has been shown that they vary according to differences in disease severity, and thus an indirect effect of the prevalence of the disease cannot be ruled out (Grobbee & Hoes, 2009).

The current study population consists of construction workers selected randomly without any criterion regarding their hearing status, as opposed to the study population participating in our previous evaluation study described in Chapter 5. Those participants were selected to have either normal hearing or NIHL, which resulted in a NIHL prevalence of 50%. As a result, listeners with beginning NIHL that are of particular interest for Earcheck validation were not involved in the evaluation study. Current study participants have a continuum of hearing threshold levels, of which 40% is considered hearing-impaired. Now that subjects with small losses are also included, they may blur the previously observed clear distinction between NH and NIHL.

Second, Earcheck outcomes could be influenced by the fact that starting level is fixed at -10 dB for each participant, hence the intelligibility of the first stimulus depends on amount of hearing loss. Figure 7.5 shows the mean signal-to-noise ratio for the different positions in the adaptive procedure, stratified for 1-dB SRT groups and for both noise conditions. Only data points representing means from at least 30 SNRs are shown. This shows that for the majority of participants the SNR for the different positions in the adaptive procedure decreases as function of presentation number. Although the first seven presentations are omitted in calculating the SRT, Figure 7.5 shows that the mean signal-to-noise ratios still improve after these 7 presentations, at least for the better performing listeners with highly negative SRTs. The final SRT obtained for these listeners might thus be an underestimation of the true speech reception threshold. Consequently, test reliability could be improved by choosing a fixed starting point at a lower SNR than -10 dB to exclude effects of the fixed starting level in NH listeners.

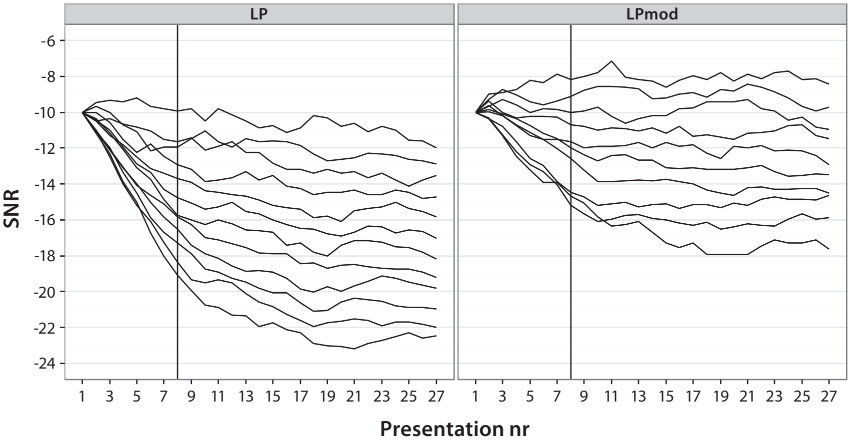


Figure 7.5. Mean signal-to-noise ratio for the different presentation numbers in the adaptive procedure, separated for both masking noise conditions. Results are shown for different SRT groups, representing results of at least 30 participants. The vertical line divides the procedure in presentations used to approach threshold, and presentations used to calculate the SRT. For the LP noise, the lines represent SRTs of -9 dB until -22 dB, for the LPmod noise, the lines represent SRTs of -7 dB until -17 dB.

To assure intelligibility of the first stimulus for each respondent, another possibility is to set the starting point individually, for instance by starting the actual adaptive test procedure of 27 presentations after the first two or three incorrect responses are given. Besides changing the starting point, a reduction of the number of stimuli used to calculate the SRT could also lead to more accurate SRT derivations, hence improve test reliability.

Third, the mean learning effect of -1.6 dB observed in the test results is rather high, compared to previous findings that report learning effects for Earcheck ranging from -0.5 to -0.7 dB (see Chapter 5 and 6). However, these results are derived from well-controlled lab testing as opposed to the domestic test administration in current study. Besides the uncontrollable parameters that may influence domestic test precision, an important difference between both test situations is the amount of instruction offered prior to testing. In the present study, participants received only written instructions; they were displayed on screen, and sent by e-mail and in the letter accompanying the provided headphones. After logging in into the experimental website, the test words are presented on screen and are played once, before test administration begins. Although similar steps were taken in the lab study, the test administration was preceded by an oral explanation of Earcheck when showing the response screen. Before starting the test, participants were able to consult the test leader if there were any questions. The small amount of instruction at home might have led to an underestimation of SRT results of the tests conducted first. This is confirmed by assessing SRTs as a function of trial number. The results of this analysis show that test outcomes improve over the eight trails conducted (Figure 7.3). An exponential curve could be fitted to this data, which estimated that after two trials the result is within one SEM of the final SRT in LP noise, and after three trials in LPmod noise.

The systematic error in repeated measures, which is expressed as the learning effect, is taken into account in the reliability calculations and causes lower ICC values and accordingly higher SEMs. To adequately assess reliability parameters, the significant systematic error should be diminished. In the case of a learning effect, additional trials should be conducted until a plateau in performance occurs, and the ICC should be calculated on the trials in the plateau region only (Weir, 2005). When the 2-3 first measurements from the test-retest subgroup are excluded, the reliability parameters could be calculated over the last 5-6 tests that roughly form a plateau. Indeed, the ICC improved, to 0.69 for LP noise and 0.67 for LPmod noise, and the SEM decreased slightly, to 1.5 dB for both masking noise conditions, indicating less variability in the test. Of course, this is only an optimized estimation of Earcheck's reliability when performed at home. In order to truly eliminate the systematic error, several consecutive Earchecks should be performed under identical test conditions to reach a plateau in performance, and accurately assess test reliability over the trials in the plateau region.

Practically this implies that Earcheck, when used for NIHL screening purposes, should also incorporate at least two practice trials prior to actual testing. The exact number of trials needed, and whether or not this should be repeated every time the same individual performs the test, should be addressed by future research.

Another alternative method to reduce the learning effect due to insufficient amount of instruction might be to apply Earcheck for testing small groups of construction employees at one central location, for instance at the company site. Although the great advantage of easy accessible domestic testing is lost this way, the test is still easy to implement and quick to perform. The test set-up could be adequately installed in a quiet but not necessarily sound-isolated room, and accurate instruction and education could be given.

Moreover, group testing might motivate more workers to participate in hearing screening, and after a controlled administration these workers could continue testing their hearing at home. Earcheck might even be considered as a supplementary low-cost test performed during POHE, to assess additional information about the functional hearing status of the workforce in the construction industry.

The big advantage of an internet-based screening test is the easy accessibility and low requirements of the test. This study showed that it is easy to develop and administer an online experiment to remotely test hearing ability. One possible application of a reliable internet test would be to monitor subtle changes in speech recognition in noise over time. To illustrate this, a pilot experiment was conducted simultaneously with the experiment described here, to investigate a possible effect on hearing of occupational noise exposure during a representative working week. For details, see the appendix.

In short, 140 participants repeated a second Earcheck after ending their last working day of the week following their first test. Their difference in SRT outcomes is 0.5 dB in LP noise and 1.0 dB in LPmod noise. Although this was significantly lower than the obtained learning effect participants still perform better when conducting Earcheck for the second time. Since the difference in SRT was not negatively associated with the intensity of noise exposure, an effect of exposure to noise during the intermediate period could not be proven. Apparently, the effects of noise exposure during a working week on SRT are, if any, too small to be detected by Earcheck in its current form. If an adapted procedure will allow a better precision, it seems worthwhile to repeat such a study. With the current procedure, monitoring of speech recognition only seems to be applicable over longer time intervals, as larger differences are expected to occur over longer time periods,.

Test-retest results showed that the minimum detectable difference of current domestic Earcheck testing equals 4.5 dB. This is quite large. Reduction of the learning effect and SEM, by performing more tests, should result in lower MDD values. A considerable reduction is necessary to arrive at SEM and MDD values that are comparable to SEM and MDD values reported for screening audiometry (Helleman & Dreschler, 2012).

The present study was conducted to validate Earcheck using two different masking noises. Again, results of this study show only small differences between test performance using stationary low-pass filtered noise or modulated low-pass filtered noise. Better discrimination between normal and impaired performance due to additional fluctuations is not proven. Although the test sensitivity to detect beginning hearing loss was slightly higher in LPmod noise, the specificity is lower. Both parameters of Earcheck with LP noise are more balanced, and accordingly the area under this ROC curve was slightly higher than for LPmod noise. In addition, results obtained in LP noise showed stronger correlation with pure-tone audiometry than SRTs in LPmod, and also reliability parameters were slightly better for this condition, as was the case in our previous studies described in Chapter 5 and 6. Furthermore, in LP noise fewer practice trials are needed to reduce the learning effect than in LPmod noise. So, although the differences are small, stationary LP noise is the recommended masking noise condition in Earcheck for domestic screening purposes.

Conclusion

Present study has validated Earcheck against a pure-tone screening audiogram. Earcheck provides the benefit of an easily accessible, self-administered alternative method for hearing screening. The Earcheck version with low-pass masking noise has a sensitivity of 68% and a specificity of 71% to detect beginning NIHL. Test-retest reliability was relatively high, 1.6 dB, as was the mean learning effect of -1.6 dB. Since this systematic difference was mainly observed in the first tests, improved instruction and the use of practice trials is recommended in order to reduce this effect and increase test reliability.

The broad internet application still promises an attractive and valuable tool for hearing screening in large populations. The easy accessibility of this test facilitates addressing large segments of the workforce, which is a major advantage over pure-tone screening audiometry in a widespread and very transient population as in the construction industry. Moreover, Earcheck allows for more frequent and on demand testing, and enables testing during leisure time, which facilitates the presence of an adequate noise-free period before testing.

Earcheck cannot replace a pure-tone screening audiogram, but can be a valuable addition to current occupational health practice by better reaching the target population and raising their awareness about noise exposure and the risk of hearing loss. Future modifications in the procedure, including a better instruction and the use of practice trials may be expected to increase the test-retest reliability, and thereby the applicability of Earcheck as a screening test in occupational health care.

Appendix

Applicability of Earcheck in monitoring small changes in speech intelligibility over time.

The major advantage of Earcheck is the easy accessibility due to its internet application. This makes the test not only suitable for hearing screening, but also for monitoring hearing status over time. The potential of Earcheck to detect small longitudinal changes in speech intelligibility is investigated in a small pilot study investigating the effects of noise exposure duringone working week on hearing status; the so-called week-effect.

Methods

140 Of the participants in current study performed the Earcheck test session twice; before and after their working week. This was done in order to investigate the influence of occupational noise exposure on SRTs measured using Earcheck. If a 'week-effect' could be established, Earcheck might be used to detect small changes in SRT outcomes, at least averaged over group results. The first Earcheck was always conducted during the weekend and the second test was performed on Friday after finishing work. After completing the second test, participants were linked to a short closing questionnaire, that consisted of 6 questions regarding noise exposure prior to the test and during the intermediate working days. In order to remind participants to conduct the second test on Friday, all received an e-mail after completing the first test.

The noise exposure during the intermediate workweek was estimated using the time-weighted average (TWA) set for the reported job title (Arbouw, 1998). Based on the equal energy principle this TWA was adjusted according to the answers in the closing questionnaire. When subjects stated they were exposed to noise for less than 50% of the time, the estimated TWA was reduced with 3 dB. When subjects reported to 'often' have worn HPDs the TWA was also decreased by 3 dB. When a subject was involved in one or more noisy recreational activities during the intermediate period, his noise exposure level was increased by 3 dB. The noise exposure levels derived from these calculations were divided into three categories; no noise exposure (TWA < 80 dBA), moderate noise exposure (TWA 80-89 dBA) and severe noise exposure (TWA ≥ 90 dBA).

Results

To assess the week-effect, a linear mixed effects model was fitted to the SRTs of the 140 participants that performed two test sessions. The variations between 'individual ear' was treated as a random effect. Fixed factors of primary interest were 'masking noise' (LP, LPmod), 'ear' (left, right) and 'repetition' (test session 1, test session 2). The main effects of 'masking noise' and 'ear' were the same as those observed in the total population (Table 7.A1). 'Repetition' showed a significant effect on SRTs measured (F[1,789] = 8.06, p = 0.005), and also significantly interacted with 'masking noise' (F[1,789] = 4.05, p = 0.045). SRTs obtained during the second test session were on average better than results of the first test session; this difference was 0.48 dB in LP noise and 0.96 dB in LPmod noise.

Since these differences in test session were smaller than those obtained in the test-retest subgroup, a second mixed effect model was fitted to the data of all participants that conducted two Earcheck sessions. This model incorporated a fixed factor for the subgroup that performed the tests ('test-retest', 'week-effect') and an interaction term of 'repetition' and 'subgroup' (Table 7.1A). The significant interaction term showed that the differences between SRTs obtained during the first and second test session were indeed 0.98 dB smaller in the week-effect subgroup than in the test-retest subgroup (F[1,948] = 12.81, p = 0.004).

Table 7.1A. Coefficients and p-values of the fixed factors in the three different mixed models.

Model	Fixed factors	Coefficient	p-value
I	Masking noise: LPmod	4.92	<0.001
	Ear: right	-0.49	0.1059
	Repetition: retest	-0.49	0.0046
	Retest * LPmod	-0.48	0.0455
II	Masking noise: LPmod	4.81	<0.001
	Repetition: retest	-1.46	<0.001
	Subgroup: week-effect	0.23	0.6571
	Retest * LPmod	-0.45	0.0330
	Retest * week-effect	0.98	0.0004
III	Masking noise: LPmod	0.49	0.0158
	Noise exposure: none	0.08	0.8302
	Noise exposure: moderate	0.39	0.0402
	Noise exposure: severe	1.07	0.0002

All coefficients are expressed in dB. In case of an interaction between two variables, the particular difference between a certain condition and the reference situation is obtained by summing up the different individual coefficients contributing to that condition.

This smaller difference between test sessions in the week-effect subgroup may have been caused by a reduced speech recognition performance due to occupational noise exposure during the intermediate test period. In order to investigate this hypothesis, the relationship between the week-effect, calculated as the difference in SRT outcomes between the first and the second test session, and noise exposure categories was examined.

A third linear mixed effect model, with week-effect as the dependent variable, 'individual ear' as random effect and fixed effects of 'masking noise' and 'noise exposure' (none, moderate, severe) was fitted. Both 'masking noise' (F[1,251] = 5.91, p = 0.016) and 'noise exposure' (F[1,251] = 3.32, p = 0.038) significantly affected the observed week-effect. There was no significant interaction between these terms. The model showed that the week-effect for LP noise was 1.07 dB for the severely exposed workers, and this was closer to the learning effect of the test-retest group than for the moderately exposed and non-exposed workers who showed almost no weekeffect (Table 7.1A). In LPmod noise, the week-effects were on average 0.49 dB larger.

Discussion

The results of this pilot experiment show a smaller difference in SRT outcomes between the first and second Earcheck session measured over a working week than in the test-retest group. Apparently, these smaller difference in the week-effect subgroup cannot be explained by negative effects of intermediate noise exposure. Either, the exposure to noise did not significantly affect SRT results, or the time between the end of workday and test performance at retest was long enough to establish recovery of temporary hearing damage. A third possibility is that Earcheck is not sensitive enough to detect these small differences, even over group results. The relatively high MDD of 4.5 dB observed in this study (Table 7.1) indicates that indeed Earcheck can only detect differences between consecutive SRT results when they are quite large.

Although monitoring small changes in SRT over time is a potentially valuable application of Earcheck, the results of this study cannot prove that Earcheck in its current form can be used for this purpose.

8

General discussion

Noise-induced hearing loss (NIHL) remains a significant public health problem, despite the widespread attention to hearing conservation. NIHL is the most reported occupational disease in the Netherlands over the past few years (Van der Molen & Lenderink, 2012) and is highly prevalent in construction. NIHL can be prevented by taking sufficient precautions, therefore hearing conservation programs are established. In practice, the use of hearing protection devices (HPDs) is the preventative measure most frequently applied in hearing conservation. However, proper usage of HPDs among construction workers is less than required (Lusk et al, 1998; Neitzel & Seixas, 2005; Seixas et al, 2011), and the effectiveness of HPDs during exposure to high noise levels is jeopardized by irregular use; selective wearing of HPDs suggests good protection while the effective attenuation is significantly reduced by noise exposure during the short periods of non-use (Else, 1973). Moreover, the cumulative contribution of non-occupational exposure to loud music and other recreational noise, may pose an additional risk to hearing (Sorgdrager & Dreschler, 2010).

Yet, the importance of good hearing in our highly communicative society is great, especially in the ageing workforce due to the trend that retirement age is increasing, and in the ageing population in general. Hence, prevention of NIHL is essential.

The aim of this thesis was to investigate screening methods for noise-induced hearing loss in hearing conservation. NIHL develops gradually and it is often unnoticed until the impairment is substantial and irreversible damage has been established. Early detection of hearing loss is therefore a crucial element in NIHL prevention programs, that helps to increase awareness and facilitates effective prevention strategies to stop further development of hearing damage.

Noise-induced hearing loss in construction

National and international studies have shown that NIHL is a highly prevalent problem in the construction industry, where noise levels frequently exceed the safety limits of occupational standards (Passchier-Vermeer et al, 1991; Suter, 2002; Tak & Calvert, 2008; Hong et al, 2011; Arbouw, 2012). Indeed, the audiometric assessments of a large population of Dutch construction workers described in Chapter 2 and 3 confirm that NIHL is still a problem in the construction industry. Noise-exposed employees have poorer hearing ability than expected based on their age. The cross-sectional analyses of Chapter 2 show that the hearing threshold levels (HTLs) of construction workers, either noise-exposed or non-exposed, are higher, and thus worse, than median HTLs predicted for their age based on annex A of ISO-1999 (1990). The employees exposed to daily noise levels exceeding 80 dBA show poorer hearing than their non-exposed colleagues, demonstrating the detrimental effects of noise exposure on hearing ability.

In addition,a longitudinal analysis described in Chapter 3 concerning a selection of the baseline cohort that also had a follow-up assessment shows that for the majority of the participants in the study, this hearing impairment develops into more hearing loss over time. Hearing ability relates more strongly to the duration of noise exposure than to estimated noise exposure level, while on the other hand the development of hearing loss is significantly associated with noise intensity only. These findings highlight the importance of adequate strategies for hearing loss prevention in this sector.

The value of the traditional approach: pure-tone audiometry

Generally, occupational NIHL is detected using behavioural audiometric techniques. The pure-tone audiogram is considered the gold standard for describing hearing sensitivity (Sataloff & Sataloff, 1993), and the diagnosis of NIHL is based on audiometric evaluation in combination with a history of noise exposure. In the Netherlands, screening audiometry is part of the periodic occupational health examination (POHE). Participation in this POHE is completely voluntary, and about 50-60% of the construction workforce participates each year.

The goal of this hearing screening is to identify those workers suffering from hearing loss, most likely attributable to occupational noise exposure. An additional application is audiometric database analysis; the analysis of periodically obtained HTLs of a group of employees, that can be used for the evaluation of hearing conservation effectiveness (Hétu, 1979; Royster & Royster, 1986). The analyses in Chapter 2 and 3 show that the obtained audiometric survey data can be adequately used for describing the hearing status and trends in hearing loss development on group level.

Limitations of pure-tone audiometry

In order to detect beginning NIHL, measurement should be precise and accurate. However, as described in Chapter 3, the quality of audiometry will have a direct impact on the accuracy of the diagnosis of hearing loss. The combination of the baseline data collection (Chapter 2) with follow-up measurements provides insight in the quality of this real world audiometric survey data. This data quality appears to leave room for improvement and is considered to be not optimal.

Not only a substantial proportion (15%) of the data had to be excluded because of inconsistencies in the consecutive datasets, probably caused by typographical errors and recall bias, also the ability to determine low HTLs was limited. Literature already showed that survey data yield poorer thresholds than those obtained under laboratory, or clinical, conditions (Dobie, 1983; Schlauch & Carney, 2012). In our analyses, the mean hearing threshold levels reflecting normal hearing, such as HTLs at the low frequencies 0.5 and 1 kHz as well as those obtained in the youngest and non-exposed workers, actually fall at 10 dB HL. This indicates that the lowest hearing threshold level that could be reliably detected was 10 dB HL instead of 0 dB HL. Since hearing loss definitions are determined relative to 0 dB HL, these higher normal HTLs have consequences for defining and detecting mild NIHL.

Although audiometric survey may yield poorer normal-hearing thresholds, useful conclusions about trends in hearing loss over time can still be drawn from the data. Nevertheless, the longitudinal dataset in Chapter 3 show large variability in HTLs obtained at two measurement occasions, resulting in a remarkable improvement in hearing ability for part of the study population. This is most likely related to measurement variability in screening audiometry rather than the reflection of an actual improvement in hearing ability. Causes of this variability may include tester and participant experience and motivation, test equipment, test procedure and test conditions. Many of these error sources can be minimized by careful control of the testing environment and cautious implementation of the standard procedure (Hétu, 1979). Reliable audiometric measurements therefore require a really quiet environment with low-level background noise levels, adequately calibration of the equipment, as well as a noise-free period of 14-16 hours prior to testing in order to eliminate TTS effects (Franks, 2001). However, in occupational screening settings these requirements are not easily met, and the observed improvement of HTLs caused by measurement variability in Chapter 3 stresses the importance of these testing requirements.

Nevertheless, when the sources of error are kept at a minimum, pure-tone audiometry still does not have perfect precision (Hétu, 1979; Schlauch & Carney, 2012). For clinical audiometry, standard errors of measurement range from 2.1 to 4.4 dB depending on frequency, meaning that only threshold shifts greater than or equal to 10 to 15 dB can be considered as significant deteriorations in hearing (Hétu, 1979). Since the variability in industrial screening may be even larger, small shifts for an individual employee due to beginning NIHL cannot easily be distinguished from measurement variation, indicating that the ability of pure-tone screening audiometry to detect early signs of potential NIHL is limited.

The screening and monitoring of hearing ability over time by pure-tone audiometry could be improved by incorporating more information on confounding variables, such as otologic history, otoscopic examination outcomes, non-occupational noise exposure, HPD usage, and on testing parameters, such as the type of audiometer, background levels, calibration date, and the possibility of TTS, in the data collection. Furthermore, alternative, or additional, methods are sought that can help to improve the early detection of NIHL in occupational health surveillance.

The value of an alternative approach; online speech-in-noise testing

As described in Chapter 1, measuring otoacoustic emissions (OAEs) can be considered a useful alternative for NIHL screening in occupational health, as this technique is suggested to be more sensitive to early signs of NIHL than the audiogram (Lapsley Miller & Marshall, 2007). However, evidence regarding this topic remains ambiguous, and the value of OAEs in hearing surveillance is currently investigated by Helleman et al. using different field studies (2010; 2012). Screening on NIHL trough OAEs is beyond the scope of this thesis.

Modern technology offers possibilities for tele-audiology (Swanepoel & Hall, 2010); performing hearing tests by internet-based applications. An online hearing test could measure hearing ability in the participant's home environment using only headphones and a home computer, making hearing screening easily accessible for a broad population. Since one of the first and most common complaints of people with NIHL is difficulty understanding speech in noise, one of the automatic online Dutch speech-in-noise self-tests can be considered a valuable method for NIHL screening. Unfortunately, the evaluation study described in Chapter 4 reveals that – although reliable – the original Earcheck that presents CVC words in a stationary, broadband noise is not sensitive enough to adequately distinguish mild NIHL from normal performance (see Figure 4.4). Because of their preserved hearing ability for the low and mid-frequencies, listeners with a beginning noise notch perform similarly to normal-hearing (NH) listeners.

In Chapter 5 possible adaptations to improve Earcheck performance in NIHL screening are investigated. First, the speech stimuli are adjusted in level to achieve equal perceptual difficulty. A homogeneous intelligibility of stimuli in noise is important for an accurate assessment of speech recognition when using an adaptive procedure (Theunissen et al, 2009). Although this adaptation does not yield the expected steeper slope of the performance intensity function that reflects test precision, it does lead to slightly better test-retest reliability expressed as SEM and ICC, stronger correlations with high-frequency pure-tone thresholds and slightly greater differences between NH and hearing-impaired (HI) listeners (see Table 8.1).

A higher impact regarding Earcheck's discriminative power was obtained when the masking noise was modified, and some of the proposed masker types increased NIHL sensitivity extensively. It was known from literature that NH subjects perform much better than HI listeners in fluctuating noise, as they benefit from the short periods of relatively low noise levels in this type of noise (Festen & Plomp, 1990). The results of Chapter 5 confirm this; Earcheck with fluctuating instead of stationary noise leads to better discrimination between NH and NIHL. However, the highest discriminative power is observed when a low-pass filtered masking noise is used. Employing this type of masking noise facilitates the use of high-frequency speech information, where limitations imposed by reduced audibility in this frequency region will impair speech intelligibility in subjects with NIHL. Consequently, the results of Chapter 5 show that Earcheck's sensitivity to detect NIHL improved from 51% to 95%, with a high specificity of 98%. This is shown to be possible without a reduction in test reliability. See Table 8.1 for an overview of the specifications of the different types of Earcheck.

Table 8.1. Test-retest reliability and validity specifications of the different types of Earcheck investigated in this thesis.

Type of Earcheck	Study type (Chapter)	Learning effect (dB)	SEM (dB)	ICC#	Correlation $PTA_{3,4,6}$	Se for NIHL (%)	Sp for NIHL (%)
Original	Lab (4)	0.82	1.24	0.75	0.62	51	90
Homogenized	Lab (5)	1.62	1.17	0.84	0.76	-	-
LP	Lab (5)	0.71	1.25	0.93	0.92	95	98
	Lab (6)*	0.54	1.62	0.81	0.83†	-	-
	Field (7)	1.40	1.63	0.65	0.56	68	71
LPmod	Lab (5)	0.48	1.39	0.87	0.88	94	92
	Lab (6)*	0.54	1.70	0.68	0.73†	-	-
	Field (7)	1.80	1.68	0.54	0.50	74	59

* *Reliability results were obtained over the 13 test conditions performed in the lab, instead of comparing only test and retest results.*

\# *Different forms of ICCs are presented, those obtained in Chapter 4 and 5 are two-way mixed model, type consistency and single measures ICCs, those from Chapter 6 and 7 are ICC are two-way random model, absolute agreement, single measure ICCs. This makes that values cannot be compared directly.*

† *These correlation coefficients were not calculated in Chapter 6, but are presented here for reasons of comparison. Correlations are calculated for results of monotic lab testing at 65 dBA.*

Limitations of online speech-in-noise testing

The lab results in this thesis demonstrate a successful improvement of Earcheck, which is highly promising for NIHL screening purposes. Yet, these results are obtained in the laboratory under well-controlled test conditions and the question is how these hold when the online test is used for remote testing, e.g. in the field of occupational health. Using Earcheck for remote hearing screening introduces uncontrollable testing paramaters, such as background noise, individually set presentation levels, and variety in the equipment used. Earcheck is rather insensitive for influences of these parameters when broadband stationary noise with the same long term average spectrum as the speech stimuli is used. The results of the study described in Chapter 6 show however, that the noise-filtering and resulting spectral differences in SNR do pose some limitations to the domestic implementation of Earcheck. Whereas normal Earcheck performance was negatively affected by testing in a domestic setting, HI performance improved, reducing the test's ability to differentiate NH from HI listeners in home-based screening.

Although testing at different presentation levels in the lab induces only small differences in test outcomes for the severely impaired listeners, hearing impaired participants are likely to have benefitted from higher presentation levels when conducting Earcheck at home. In addition, variations in frequency responses of sound cards and/or transducers could have played a role. Since all participants use the same type of headphones, provided by the study, variability in domestic results might increase when different types of transducers are used for testing.

Since the results of Chapter 6 indicate that domestic testing affect SRT results, validity of home-based use of Earcheck for NIHL detection should be determined in a field study, that also assesses its reliability for home testing and revises the cut-off values defined in the lab study of Chapter 5. In Chapter 7 results of domestic Earcheck testingare compared to screening pure-tone audiometry for approximately 250 noise-exposed construction workers, to assess the test's applicability in the field of occupational health.

A subgroup of 32 participants performed Earcheck twice on the same day, in test and retest. These results display a rather large learning effect for remote testing of 1.6 dB and slightly poorer test-retest reliability than that observed in lab studies (see Table 8.1). Since the systematic learning effect is mainly observed in the first tests of the experimental sequence, improved instruction and use of practice trials may reduce this effect and increase test reliability.

Moreover, the increased variability observed in both Earcheck and pure-tone audiometry when performed in the context of NIHL screening yields a less strong correlation between these two methods of 0.56 (Table 8.1). As a result, the sensitivity of the domestic Earcheck to detect beginning NIHL is 68%, with a specificity of 71%. Although, it must be realized that these parameters only hold for this specifically tested study population, we believe however, that these participants are adequate representatives for the population to be using the test.

The obtained validity of Earcheck to detect beginning NIHL when used for remote screening purposes, such as in the field of occupational health, is not considered optimal, especially when detecting and monitoring subtle effects.

All studies described in this thesis concerning the ability of Earcheck to detect NIHL used two types of low-pass filtered noise; a stationary low-pass filtered noise (LP) and a stationary low-pass filtered noise combined with a high-pass interrupted noise replacing the removed high-frequency part of the noise (LPmod). Although the investigation in Chapter 5 already showed that LP noise had the best discrimination and more reliable results, LPmod showed the second best results, and is also considered a good alternative for NIHL screening. However, the results for LPmod in the field, obtained in Chapter 6 and 7, were either similar to or slightly poorer than the results obtained in LP noise. So LP noise is considered the best alternative for NIHL screening.

Future research

Although promising results of Earcheck with low-pass filtered noise for NIHL screening application were obtained under well-controlled laboratory circumstances, the field study showed that the online Earcheck performance was not optimal in terms of reliability and validity. Future research could investigate ways to improve online low-pass filtered Earcheck performance for NIHL screening purposes.

Analyses in Chapter 7 already show that the first tests in a session of multiple Earchecks are mainly responsible for the large learning effect and part of the resulting higher test-retest variability in domestic outcomes (Figure 7.3). This suggests that performing practice tests might reduce the learning effect and increase test reliability. Although the fit of an exponential curve to the data shows that a plateau in SRT is reached after the first 2 tests using LP noise (see Table 7.2), field research should investigate the actual number of practice tests needed and should describe how this affects the observed learning effect and test reliability. In doing so, also the degree of instruction should be taken into account.

In addition, the individual starting level (i.e. SNR) of the test affects the estimation of SRT and thus test reliability. Analyses in Chapter 7 show that by fixing the starting point at -10 dB SNR, the actual SRT of particularly good performers is underestimated (Figure 7.5). Generally, the SNR for the different stimuli in the adaptive procedure decreases as function of presentation number. The SNRs of the first seven presentations, considered as run-up to the SRT, are omitted in SRT calculation. However in some participants this amount of presentations is insufficient to reach the level of approximately 50% intelligibility. In those cases the difference between the starting SNR and the actual SRT is too large. Smits & Houtgast (2006) showed that this difference should be less than 5 dB to minimize its effect on the SRT. This problem could thus be solved by choosing a starting SNR closer to the estimated SRT, which could be achieved by setting this starting SNR individually, for instance by starting the actual adaptive test procedure of 27 presentations after the first one or two incorrect responses are given. This results in a similar number of presentations around the actual threshold and comparable audibility of the presentations used to calculate the SRT for all participants.

An alternative solution would be to consider a longer run-up, at the cost of the number of averaged responses. However, a reduced number of presentations used for SRT estimation may lead to reduced accuracy of the test outcome. Increasing the number of stimuli in the test would counteract this, but increases testing duration.

A possibility to increase the ability to differentiate between NH listeners and subjects with NIHL, might be choosing a different target point in the up-down procedure. Generally, SRT is determined for 50% of correct intelligibility, since the slope of the performance intensity (PI) curve is steepest around this point. Because the PI curve is usually shallower for HI than NH listeners, these groups deviate more at higher target points. Hence a higher target point, reached by a different up-down procedure, may result in better discrimination between normal and impaired performance (Smits et al, 2011).

Finally, speech material is an important parameter of a speech-in-noise test. Earcheck uses nine different CVC words, each with a unique vowel. In order to increase test precision, the individual words were homogenized. This was done for over 100,000 results obtained by online testing in the original stationary broadband masking noise. As Earcheck is improved by using a low-pass filtered masking noise, homogenization should be repeated for this masking noise condition. Although analysis of Smits et al. (2007) showed that correction factors determined for one type of noise can also be used for homogenization of the stimuli in other types of noise, this might not be a reliable assumption when low-pass filtered noise is used. Because of the partial elimination of the high-frequency region of the masking noise, spectral differences between speech and noise occurred, and perceptual audibility of the words might have changed. It would be best if this was done for both NH and HI listeners.

Based on those results, the chosen set of stimuli may be reconsidered. The Occupational Earcheck described in Chapter 4 uses a speech set specifically chosen to contain a higher proportion of high-frequency consonants and more similar vowels (Ellis et al, 2006). Although this did not lead to higher sensitivity for NIHL when presented in stationary noise, this might be the case when low-pass filtered noise that stresses the available high frequency information is used, which might increase the discriminative power of Earcheck.

Application of the online speech-in-noise screening test

Possibilities in occupational hearing conservation

Although the test performance of Earcheck in NIHL screening is not optimal yet, the internet-based speech-in-noise application has great advantages over pure-tone audiometry in current occupational practice.

The most important benefit of the online screening tests is its accessibility. Using Earcheck for remote testing, for instance at home, has the potential of reaching almost all participants at risk, and can result in higher participation rates. This would be especially the case in the construction sector, an industry characterized by a highly mobile and widespread workforce with a high rate of self-employment that challenges the participation in regular POHEs. Another advantage is that testing can be done more frequently than the 2-4 year cycle of POHEs, and also testing on demand is possible, for example when complaints arise. This may also positively affect the participation rate.

With regards to occupational hearing screening, Earcheck testing has low requirements; only a PC with internet connection and a pair of headphones are needed. A quiet room will be sufficient for adequate testing, the calibration is software-set and since it is a self-test no trained technician is needed to perform the test. This would highly reduce the costs of screening compared to pure-tone audiometry. Finally, as testing can be performed at home, screening can be done at any given time. Testing during a day-off or before the start of the workday, may be very efficient in preventing biasing TTS effects of occupational noise exposure preceding the test.

In order to make Earcheck completely feasible for practical occupational hearing screening, appropriate instructions and more design work is required to ensure proper test administration (Culling et al, 2005). In addition, attention should be given to the educational part of a screening assessment, and to the handling of abnormalities by forwarding the findings to an audiological or occupational specialist.

Although not an alternative method for audiometric evaluation, these advantages make Earcheck a valuable addition to the current practice in occupational health. This can be implemented in various ways:

- By regular performance of the online screening test hearing ability can be monitored, and audiometric evaluation should follow when abnormalities arise. In order not the miss workers with NIHL, the proportion of false negatives can be reduced by increasing the test sensitivity. Although this would inevitably lead to a lower specificity, this procedure would result in a smaller number of audiometric evaluations than is the case in current practice.
- Due to additional implementation of online speech-in-noise testing audiometry can be performed less frequent, for instance every other POHE. Savings from the resulting reduction in costs may be invested in improving audiometric testing conditions, such as the availability of audiometric booths.
- Online Earcheck turned out to be highly valuable for NIHL screening when performed in a well-controlled test environment. This offers the possibility of measuring employees at a central location with a standardized and calibrated test set-up, for instance at the company or POHE site, either or not after a pre-selection of workers showing abnormalities by means of domestic testing.

This way, Earcheck can be used as a supplementary low-cost test performed during POHE, to assess additional information about the functional hearing status of the construction.

Other applications of tele-audiology

The data collection conducted in Chapter 7 was completely internet-based, and proved to be a feasible method for data collection. This makes tele-audiology, in the form of hearing screening by an online self-test highly applicable in other situations, such as:

- A measurement instrument in large epidemiological studies

Pure-tone audiometry is not very suitable for large-scale population-based epidemiological studies because of the expensive and complicated procedures required for accurate measurements of hearing ability. Internet-based testing offers the possibility of cheap and easy hearing measurement, that can be very useful in longitudinal studies of hearing ability (as was done the National Longitudinal Study on Hearing by Nachtegaal et al 2009), especially in monitoring intra-individual changes over time (Honeth et al, 2010).

- Assessing hearing status at remote sites in underserved areas

Earcheck is a reasonably useful and valid alternative when audiometric testing is not available. The web-based hearing screening test has the potential to provide widespread access to services that are affordable and do not require specialist personnel on site. This offers important healthcare coverage for rural areas and developing countries where specialized audiological services are limited (Seren, 2009; Honeth et al, 2010; Swanepoel & Hall, 2010)

- Offer self-administered hearing screening tests to the general public

Besides for occupational screening, self-administered hearing tests for individual health checkups can also be used for screening in the general public. This is already proven successful by the implementation of the National Hearing Test in the Netherlands (Smits et al, 2004; Smits & Houtgast, 2005), as well as similar tests in other European countries and the US (Zokoll et al, 2012; Watson et al, 2012). It provides a more accessible means for people to have their hearing checked and to help them decide whether they should seek professional help. In addition, Earcheck can be used for hearing screening of specific groups at risk of NIHL, such as youngsters that are exposed to high recreational noise level.

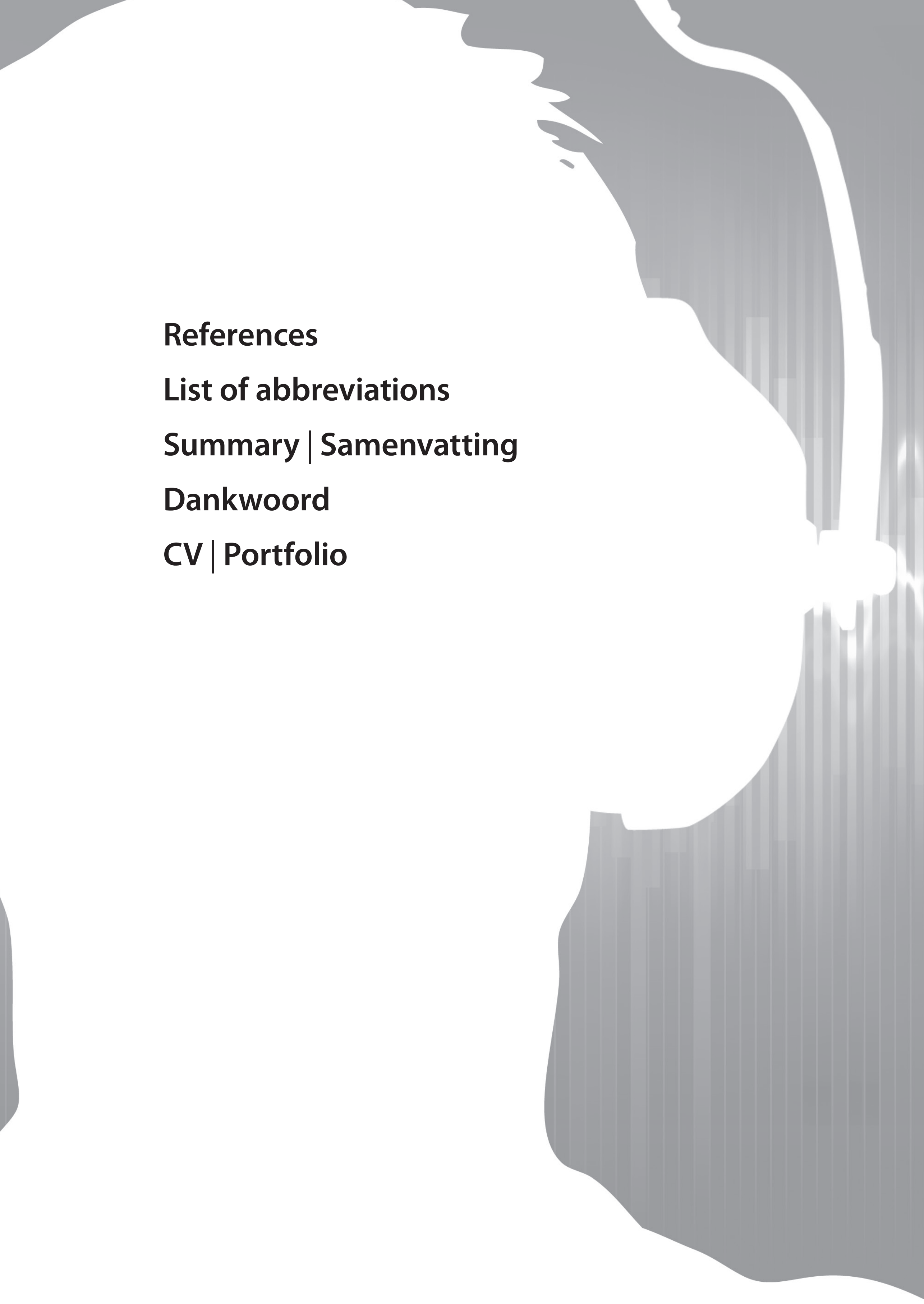
- Usage as smartphone application

A recent development is the use of smartphones, which are communication devices with the ability to program and control their audio output. Using this ability, so called applets can be installed that can test hearing ability (Stenfelt et al, 2011, Kam et al, 2012). This provides an easily accessible and convenient means for people to take a valid hearing test, that could be readily offered to the general population to raise the public awareness on hearing health (Kam et al, 2012).

General conclusion

Despite the essence of early detection of NIHL in hearing conservation programs, current methods for NIHL identification show large variability. In the light of the search for an alternative or additional method to improve NIHL screening by audiometry in occupational health surveillance, Earcheck can be considered as a valuable addition to pure-tone audiometry, in terms of accessibility, simplicity, and low requirements. When performed under well-controlled testing conditions Earcheck has a high sensitivity and specificity for NIHL detection.

However, in its current stage, the online Earcheck application shows too much variability to be considered a more sensitive measure of early NIHL than pure-tone audiometry. Nonetheless its internet-based application provides widespread access the hearing tests, raising public awareness on hearing health and offering other useful possibilities for utilization.



References

List of abbreviations

Summary | Samenvatting

Dankwoord

CV | Portfolio

References

Agrawal, Y., Niparko, J.K. & Dobie, R.A. 2010. Estimating the effect of occupational noise exposure on hearing thresholds: the importance of adjusting for confounding variables. *Ear Hear*, 31, 234-237.

Albera, R., Lacilla, M., Piumetto, E. & Canale, A. 2010. Noise-induced hearing loss evolution: influence of age and exposure to noise. *Eur Arch Otorhinolaryngol*, 267, 665-671.

Albrecht, J., van Elewout, L., Verhage, L., Verweij, C. 2005. *Oorcheck: de validering van een interactief screenings-instrument*. Leiden, Internal report, LUMC.

ANSI S3.44. 1996. *Determination of Occupational Noise Exposure and Estimation of Noise-Induced Hearing Impairment*. New York, American National Standards Institute.

ANSI S3.5. 1997. *American National Standard Methods for Calculation of the Speech Intelligibility Index*. New York, American National Standards Institute.

ANSI S1.42-2001. 2011. *Design Response of Weighting Networks for Acoustical Measurements*. New York, American National Standards Institute.

Arbouw. 1998. *Lawaai in de bouw. Arbouw-advies voor de bouwnijverheid*. Harderwijk, Stichting Arbouw.

Arbouw. 2006. *Richtlijn gehoorstoornissen*. Harderwijk, Stichting Arbouw.

Arbouw. 2009. *Bedrijfstakatlas 2009*. Harderwijk, Stichting Arbouw.

Arbouw. 2012. *Bedrijfstakatlas 2012*. Harderwijk, Stichting Arbouw.

Arezes, P.M. & Miguel, A.S. 2005. Individual Perception of Noise Exposure and Hearing Protection in Industry. *Hum Factors*, 47, 683-692.

Arlinger, S. 2003. Negative consequences of uncorrected hearing loss – a review. *Int J Audiol,* 42 Suppl 2, 2S17-20.

Arndt. V., Rothenbacher, D., Brenner, H., Fraisse, E., Zschenderlein, B., Daniel, U., Schuberth, S. & Fliedner, T.M. 1996. Older workers in the construction industry: results of a routine health examination and a five year follow up. *Occup Environ Med*, 53, 686-691.

ASA. 1951. *American standard specification for audiometers for general diagnostic purposes* (Z24.5-1951). New York, American Standards Association.

Attias, J., Bresloff, I., Reshef, I., Horowitz, G. & Furman, V. 1998. Evaluating noise induced hearing loss with distortion product otoacoustic emissions. *Br J Audiol*, 32, 39-46.

Attias, J., Horovitz, G., El-Hatib, N. & Nageris, B. 2001. Detection and clinical diagnosis of noise-induced hearing loss by otoacoustic emissions. *Noise Health*, 3, 19-31.

Bacon, S.P., Opie, J.M. & Montoya, D.Y. 1998. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J Speech Lang Hear Res*, 41, 549-563.

Badri, R., Siegel, J.H. & Wright, B.A. 2011. Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiogram. *J Acoust Soc Am,* 129, 852–863.

Bauer, P., Körpert, K., Neuberger, M., Raber, A. & Schwertz, F. 1991. Risk factors for hearing loss at different frequencies in a population of 47.388 noise-exposed workers. *J Acoust Soc Am,* 90, 3086-3098.

Biassoni, E.C., Serra, M.R., Richtert, U., Joekes, S., Yacci, M.R., Carignani, J.A., Abraham, S., Minoldo, G. & Franco, G. Recreational noise exposure and its effects on the hearing of adolescents. Part II: development of hearing disorders. *Int J Audiol*, 44, 74-85.

Bielefeld, E.C., Hu, B.H., Harris, K.C. & Henderson, D. 2005. Damage and threshold shift resulting from cochlear exposure to paraquat-generated superoxide. *Hear Res*, 207, 35–42.

Bosman, A.J. 1989. *Speech perception by the hearing impaired*, Thesis, University of Utrecht.

Bosman, A.J. & Smoorenburg, G.F. 1995. Intelligibility of Dutch CVC syllables and sentences for listeners with normal-hearing and with three types of hearing impairment. *Audiology*, 34, 260-284.

Bronkhorst, A.W. & Plomp, R. 1988. The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J Acoust Soc Am*, 83, 1508-1516.

Brownell, W.E. 1990. Outer hair cell electromotility and otoacoustic emissions. *Ear Hear*, 11, 82-92.

Cacciatore, F., Napoli, C., Abete, P., Marciano, E., Triassi, M. & Rengo, F. 1999. Quality of life determinants and hearing function in an elderly population: Ossevatorio Geriatrico Campano Study Group. *Gerontology*, 45, 323-328.

Carabellese, C., Appollonio, I., Rozzini, R., Bianchetti, A., Frisoni, G.B. & Trabucchi, M. 1993. Sensory impairment and quality of life in a community elderly population. *J Am Geriatr Soc*, 41, 401-407.

Chan, V.S., Wong, E.C. & McPherson, B. 2004. Occupational hearing loss: screening with distortion-product otoacoustic emissions. *Int J Audiol*, 43, 323-329.

Chia, E.M., Wang, J.J., Rochtchina, E., Cumming, R.R., Newall, P. & Mitchell, P. 2007. Hearing impairment and health-related quality of life: The Blue Mountains hearing study. *Ear Hear*, 28, 187-195.

Chung, D. Y. & Mack, B. 1979. The effect of masking by noise on word discrimination scores in listeners with normal-hearing and with noise-induced hearing loss. *Scand Audiol*, 8, 139-143.

Chung, J.H., Des Roches, C.M., Meunier, J. & Eavey, R.D. 2005. Evaluation of noise-induced hearing loss in young people using a web-based survey technique. *Pediatrics*, 115, 861-867.

Clark, W.W. & Bohne, B.A. 1999. Effects of noise on hearing. *JAMA*, 281, 1658–1659.

Clark, W.W. & Bohl, C.D. 2005. Hearing levels of firefighters: risk of occupational noise-induced hearing loss assessed by cross-sectional and longitudinal data. *Ear Hear*, 26, 327-340.

Culling, J.F., Zhao, F. & Stephens, D. 2005. The viability of speech-in-noise audiometric screening using domestic audio equipment. *Int J Audiol*, 44, 691-700.

Daniel, E. 2007. Noise and hearing loss: a review. *J Sch Health*, 77, 225-231.

Davies, H., Marion, S. & Teschke, K. 2008. The impact of hearing conservation programs on incidence of noise-induced hearing loss in Canadian workers. *Am J In Med,* 51, 923-931.

Davis, A., Haggard, M. & Bell, I. 1990. Magnitude of diotic summation in speech-in-noise tasks: performance region and appropriate baseline. *Br J Audiol*, 24, 11-16.

Dawson-Saunders, B. & Trapp, R.G. 1994. *Basic and Clinical Biostatistics*. 2nd ed. Connecticut, Appleton Lange.

Deming, W.E. 1964. *Statistical adjustment of data*. 2nd edn. New York, Dover Publications, Inc.

De Moraes Marchiori, L.L., De Almeida Rego, F.E. & Matsuo, T. 2006. Hypertension as a factor associated with hearing loss. *Bras J Otorhinolaringol*, 72, 533-540.

Desai, A., Reed, D., Cheyne, A., Richards, S. & Prasher, D. 1999. Absence of otoacoustic emissions in subjects with normal audiometric thresholds implies exposure to noise. *Noise Health*, 1, 58-65.

Dobie, R.A. 1983. Reliability and validity of industrial audiometry: implications for hearing conservation program design. *Laryngoscope*, 93, 906-927.

Dobie, R.A. 2001. *Medical-legal evaluation of hearing loss*. 2nd ed, San Diego, Singular publishing group, Thomson learning.

Dobie, R.A. 2006. Methodological issues when comparing hearing thresholds of a group with population standards: the case of the ferry engineers. *Ear Hear,* 27, 526-537.

Dobie, R.A. 2007. Noise-induced permanent threshold shifts in the occupational noise and hearing survey: an explanation for elevated risk estimates. *Ear Hear,* 28, 580-591.

Dobie, R.A. 2008. The burdens of age-related and occupational noise-induced hearing loss in the United States. *Ear Hear*, 29, 565-577.

Dubno, J.R., Horwitz, A.R. & Ahlstrom, J.B. 2006. Spectral and threshold effects on recognition of speech at higher-than-normal levels. *J Acoust Soc Am*, 120, 310-320.

Edelson, J., Neitzel, R., Meischke, H., Daniell, W., Sheppard, L., Stover, B. & Seixas, N. 2009. Predictors of hearing protection use in construction workers. *Ann Occup Hyg*, 53, 605-615.

Ellis, N., Kuijpers, M., van der Pijl, S. & Verbiest, E. 2006. *Ontwikkeling van een gehoorscreening voor het bedrijfsleven*. Leiden, Internal report, LUMC.

Else, D. 1973. A note on the protection afforded by hearing protectors--implications of the energy principle. *Ann Occup Hyg*, 16, 81-83.

Engdahl, B. 1996. Effects of noise and exercise on distortion product otoacoustic emisions. *Hear Res*, 93, 72-82.

European Parliament and Council (EPC). 2003. Directive 2003-10-EC on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical agents. *Official J Europ Union*, 15.2.2003.

Festen, J.M. & Plomp, R. 1990. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*, 88, 1725-1736.

Festen, J.M. & Smits, C. 2007. Is speech reception in stationary noise suitable for screening noise induced hearing damage? *Ned Tijdschr KNO-Heelk*, 13.

Franks, J.R. 2001. Hearing measurement. *In:* Goelzer, B., Hansen, C.H. & Sehrndt, G.A. (eds). *Occupational exposure to noise: evaluation, prevention and control*. Dortmund, Germany : World Health Organization, pp. 183-232.

Gao, W.Y., Ding, D.L., Zheng, X.Y., Ruan, F.M. & Liu, Y.J. 1992. A comparison of changes in the stereocilia between temporary and permanent hearing losses in acoustic trauma. *Hear Res*, 62, 24-41.

Gerges, S.N.Y., Vedsmand, L. & Lester, H. 2001. Personal measures and hearing conservation. *In:* Goelzer, B., Hansen, C.H. & Sehrndt, G.A. (eds). *Occupational exposure to noise: evaluation, prevention and control*. Dortmund, Germany : World Health Organization, pp. 297-316.

Glorig, A., Quiggle, R., Wheeler, D.E. & Grings, W. 1956. Determination of the normal hearing reference zero. *J Acous Soc Am*, 28, 1110–1113.

Gorga, M.P., Neely, S., Johnson, T.A., Dierking, D.M. & Garner C.A. 2007. Distortion-Product Otoacoustic Emissions in relation to hearing loss. *In*: Robinette, M.S. & Glattke, T.J. (eds) *Otoacoustic Emissions: Clinical Applications*. 3rd ed. New York. Thieme Medical Publishers, pp. 197-275.

Griffin, S.C., Neitzel, R., Daniell, W.E. & Seixas, N.S. 2009. Indicators of hearing protection use: self-report and researcher observation. *J Occup Environ Hyg*, 6, 639-647.

Grobbee, D.E. & Hoes, A.W. 2009. *Diagnostic research. Clinical Epidemiology: Principles, Methods, and Applications for Clinical Research*. Sudbury. Jones & Bartlett Learning, pp. 58 - 102.

Hall, A.J. & Lutman, M.E. 1999. Methods for early identification of noise-induced hearing loss. *Audiology*, 38, 277-280.

Hamernik, R.P., Patterson, J.H., Turrentine, G.A. & Ahroon, W.A. 1989. The quantitative relation between sensory cell loss and hearing thresholds. *Hear Res*, 38, 199-211.

Health and Safety Authority (HSA). 2007. *Guidelines on hearing checks and audiometry under the safety, health and welfare at work regulations 2007; Control of noise at work*. Dublin, HSA.

Helleman, H.W., Jansen, E.J.M. & Dreschler, W.A. 2010. Otoacoustic emissions in a hearing conservation program: General applicability in longitudinal monitoring and the relation to changes in pure-tone thresholds. *Int J Audiol*, 49, 410–419.

Helleman, H.W. & Dreschler, W.A. 2012. Overall versus individual changes for otoacoustic emissions and audiometry in a noise-exposed cohort. *Int J Audiol*, 51, 362-372.

Henderson, D., Subramaniam, M. & Boettcher, F.A. 1993. Individual susceptibility to noise-induced hearing loss: an old topic revisited. *Ear Hear*, 14, 152-168.

Henderson, D. & Saunders, S.S. 1998. Acquisition of noise-induced hearing loss by railway workers. *Ear Hear*, 19, 120-130.

Henderson, D., Bielefeld, E.C., Harris, K.C. & Hu, B.H. 2006. The role of oxidative stress in noise-induced hearing loss. *Ear Hear*, 27, 1-19.

Hessel, P.A. 2000. Hearing loss among construction workers in Edmonton, Alberta, Canada. *J Occup Environ Med*, 42, 57-63.

Hétu, R. 1979. Critical analysis of the effectiveness of secondary prevention of occupational hearing loss. *J Occup Med*, 21, 251-254.

Hetu, R. & Fortin, M. 1995. Potential risk of hearing damage associated with exposure to highly amplified music. *J Am Acad Audiol*, 6, 378-386.

Honeth, L., Bexelius, C., Eriksson, M., Sandin, S., Litton, J.E., Rosenhall, U., Nyrén, O. & Bagger-Sjöbäck, D. 2010. An internet-based hearing test for simple audiometry in nonclinical settings: preliminary validation and proof of principle. *Otol Neurotol*, 31, 708-714.

Hong, O. & Csaszar, P. 2005. Audiometric testing and hearing protection training through multimedia technology. *Int J Audiol*, 44, 522-530.

Hong, O.S. 2005. Hearing loss among operating engineers in American construction industry. *Int Arch Occup Environ Health*, 78, 565-574.

Hong O.S., Ronis, D.L. & Antonakos, C.L. 2011. Validity of self-rated hearing compared with audiometric measurement among construction workers. *Nurs Res*, 60, 326-332.

Hotz, M.A., Probst, R., Harris, F.P. & Hauser R. 1993. Monitoring the effects of noise exposure usins transiently evoked otoacoustic emissions. *Acta Otolaryngol*, 262, 217-224.

Hoeymans, N., Eysink, P.E.D. & de Hollander, A.E.M. 2005. *Arbeidsomstandigheden en ziektelast. Een haalbaarheidsstudie*. RIVM rapport 270052001/2005. Bilthoven, RIVM.

Hu, B.H., Henderson D. & Nicotera T.M. 2002. Involvement of apoptosis in progression of cochlear lesion following exposure to intense noise. *Hear Res*, 166, 62-71.

ISO-6189. 1983. *Acoustics – Pure tone air conduction threshold audiometry for hearing conservation purposes*. Geneva, International Organisation for Standardization.

ISO-8253.1. 2010. *Acoustics – Audiometric test methods – Part 1: pure-tone air and bone conduction audiometry*. Geneva, International Organisation for Standardization.

ISO-1999. 1990. *Acoustics - Determination of occupational noise exposure and estimation of noise-induced hearing impairment*. Geneva, International Organisation for Standardization.

ISO-398.1. 1998. *Acoustics – Reference zero for the calibration of audiometric equipment - Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones*. Geneva, International Organisation for Standardization.

Jansen, E.J.M., Helleman, H.W., Dreschler, W.A. & de Laat, J.A.P.M. 2009. Noise induced hearing loss and other hearing complaints among musicians of symphony orchestras. *Int Arch Occup Environ Health*, 82, 153–164.

Jansen, S., Luts, H., Wagener, K.C., Frachet, B., Wouters, J. 2010. The French digit triplet test: a hearing screening tool for speech intelligibility in noise. *Int J Audiol*, 49, 378-387.

Johnson, D.L. 1991. Field studies: industrial exposures. *J Acoust Soc Am*, 90, 170-174.

Jongmans, P., Martens, M. & Dreschler, W.A. 2008. *Analyse van de hoortesten van de Nationale Hoorstichting*. Leiden, Nationale Hoorstichting.

Kam, A.C., Sung, J.K., Lee, T., Wong, T.K. & van Hasselt, A. 2012. Clinical evaluation of a computerized self-administered hearing test. *Int J Audiol*, 51, 606-610.

Kaplan, H. & Pickett, J.M. 1981. Effects of dichotic/diotic versus monotic presentation on speech understanding in noise in elderly hearing-impaired listeners. *Ear Hear*, 2, 202-207.

Keefe, D.H. 2007. Influence of middle-ear function and pathology on otoacoustic emissions. *In*: Robinette, M.S. & Glattke, T.J. (eds) *Otoacoustic Emissions: Clinical Applications*. 3rd ed. New York. Thieme Medical Publishers, pp. 163-196.

Kemp, D.T. 1978. Stimulated acoustic emissions from within the human auditory system. *J Acoust Soc Am*, 64, 1386-1391.

Kemp, D.T. 1986. Otoacoustic emissions, travelling waves and cochlear mechanisms. *Hear Res*, 22, 95-104.

Kemp, D.T., Ryan, S. & Bray, P. 1990. A guide to the effective use of otoacoustic emissions. *Ear Hear*, 11, 93-105.

Kemp, D.T. 2007. The basics, the science, and the future potential of otoacoustic emissions. *In*: Robinette, M.S. & Glattke, T.J. (eds) *Otoacoustic Emissions: Clinical Applications*. 3rd ed. New York, Thieme Medical Publishers, pp. 7-42.

Keppler, H., Dhooge, I., Maes, L., D'haenens, W., Bockstael, A., Philips, B., Swinnen, F. & Vinck, B. 2010. Transient-evoked and distortion product otoacoustic emissions: A short-term test-retest reliability study. *Int J Audiol*, 49, 99-109.

Kimura, D. 1961. Some effects of temporal-lobe damage on auditory perception. *Can J Psychol*, 15, 156-65.

Konopka, W., Pawlaczyk-Luszczynska, M., Sliwinska-Kowalska, M., Grzanka, A. & Zalewski, P. 2005. Effects of impulse noise on transiently evoked otoacoustic emission in soldiers. *Int J Audiol*, 44, 3-7.

Koopman, J., Davey, E., Thomas, N., Wittkop, T. & Verschuure, H. 2008. How should hearing screening tests be offered? *Int J Audiol*, 47, 230-237.

Kramer, S.E., Kapteyn, T.S., Kuik, D.J. & Deeg D.J.H. 2002. The association of hearing impairment and chronic diseases with psychosocial health status in older age. *J Aging Health*, 14, 122-137.

Kuipers, S.M. 2007. *Bedrijfsoorcheck, een spraak in ruis test via internet speciaal voor werknemers*. Leiden, Internal report, LUMC.

Kumar, U.A., Ameenudin, S. & Sangamanatha, A.V. 2012. Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. *Noise Health*, 14, 100-105.

Kurmis, A.P. & Apps, S.A. 2007. Occupationally-acquired noise-induced hearing loss: a senseless workplace hazard. *Int J Occup Med Environ Health*, 20, 127-136.

Lapsley Miller, J.A., Marshall, L. & Heller, L.M. 2004. A longitudinal study of changes in evoked otoacoutic emissions and pure-tone thresholds as measured in a hearing conservation program. *Int J Audiol*, 43, 307-322.

Lapsley Miller, J.A., Marshall, L., Heller, L.M. & Hughes L.M. 2006. Low-level otoacoustic emissions may predict susceptibility to noise-induced hearing loss. *J Acoust Soc Am*, 120, 280-296.

Lapsley Miller, J.A. & Marshall, L. 2007. Otoacoustic emissions as a preclinical measure of noise-induced hearing loss and susceptibility to noise-induced hearing loss. *In*: Robinette, M.S. & Glattke, T.J. (eds) *Otoacoustic Emissions: Clinical Applications*. 3rd ed. New York, Thieme Medical Publishers, pp. 321-341.

Leensen M.C.J., Helleman, H.W., Jansen, E.J.M. & Dreschler, W.A. 2011. *De toepasbaarheid van otoakoestische emissies als screenings- en monitoringsinstrument voor lawaaischade in de bouwnijverheid.* Harderwijk, Stichting Arbouw.

LePage, E.L. & Murray, N.M. 1993. Click-evoked otoacoustic emissions: comparing emissions strengths with pure tone audiometric thresholds. *Aust J Audiol*, 15, 9-22.

Le Prell, C.G., Yamashita, D., Minami, S.B., Yamasoba, T. & Miller, J.M. 2007. Mechanisms of noise-induced hearing loss indicate multiple methods of prevention. *Hear Res*, 226, 22-43.

Lonsbury-Martin, B.L., Martin, G.K., McCoy, M.J. & Whitehead, M.L. 1995. New approaches to the evaluation of the auditory system and a current analysis of otoacoustic emissions. *Otolaryngol Head Neck Surg*, 12, 50-63.

Lusk S.L, Ronis D.L. & Hogan M. 1997. Test of the Health Promotion Model as a causal model of construction workers ' use of hearing protection. *Res Nurs Health*, 20, 183-194.

Lusk, S.L., Kerr, M.J. & Kauffman, S.A. 1998. Use of hearing protection and perceptions of noise exposure and hearing loss among construction workers. *Am Ind Hyg Assoc J*, 59, 466-470.

Lusk S.L., Hong O.S., Ronis D.L., Eakin B.L., Kerr M.J. & Early, M.R. 1999. Effectiveness of an intervention to increase construction workers ' use of hearing protection. *Hum Factors*, 41, 487-494.

Marshall, L., Heller, L.M. & Westhusin, L.J. 1997. Effect of negative middle-ear pressure on transient-evoked otoacoustic emissions. *Ear Hear*, 18, 218–226.

Marshall, L., Lapsley Miller, J.A., Heller, L.M., Wolgemuth, K.S., Hughes, L.M., Smith, S.D. & Kopke, R.D. 2009. Detecting incipient inner-ear damage from impulse noise with otoacoustic emissions. *J Acoust Soc Am*, 125, 995-1013.

Martens, M.K., Perenboom, R.J.M., van der Ploeg, C.P.B. & Dreschler, W.A. 2005. *Oorcheck: analyse van gehoortesten voor jongeren over de periode april 2004-oktober 2004.* Leiden, TNO.

Mathers, C.D., Smith, A. & Concha, M. 2003. *Global burden of hearing loss in the year 2000.* Geneva: World Health Organization.

May, J.J. 2000. Occupational hearing loss. *Am J Ind Med,* 37, 112-120.

McArdle, R.A., Killion, M., Mennite, M.A. & Chisolm, T.H. 2012. Are two ears not better than one? *J Am Acad Audiol*, 23, 171-181.

Melamed, S., Rabinowitz S., Feiner M., Weisberg E. & Ribak J. 1996. Usefulness of the protection motivation theory in explaining hearing protection device use among male industrial workers. *Health Psychol*, 15, 209-215.

Melnick, W. 1991. Human temporary threshold shift (TTS) and damage risk. *J Acoust Soc Am*, 90, 147-154.

Meyer-Bisch, C. 1996. Epidemiological evaluation of hearing damage related to strongly amplified music (personal cassette players, discotheques, rock concerts)--high-definition audiometric survey on 1364 subjects. *Audiology*, 35, 121-142.

Mizoue, T., Miyamoto, T. & Shimizu, T. 2003. Combined effect of smoking and occupational exposure to noise on hearing loss in steel factory workers. *Occup Environ Med*, 60, 56-59.

Mostafapour, S.P., Lahargoue, K. & Gates, G.A. 1998. Noise-induced hearing loss in young adults: the role of personal listening devices and other sources of leisure noise. *Laryngoscope*, 108, 1832-1839.

Mulrow, C.D., Aguilar, C., Endicott, J.E., Velez, R., Tuley, M.R., Charlip, W.S. & Hill, J.A. 1990. Association between hearing impairment and the quality of life of elderly individuals. *J Am Geriatr Soc*, 38, 45-50.

Nachtegaal, J., Smit, J.H., Smits, C., Bezemer, P.D., van Beek, J.H.M., Festen, J.M. & Kramer, S.E. 2009. The association between hearing status and psychosocial health before the age of 70 years: Results from an Internet-based National Survey on Hearing. *Ear Hear*, 30, 302-312.

National Institute for Occupational Safety and Health (NIOSH). 1998. *Criteria for a recommended standard: occupational noise exposure, revised criteria. Pub. No. 98-126*. Cincinnati, NIOSH.

Nederlands Centrum van Beroepsziekten (NCvB). 1999. *Registratierichtlijn B001. Beroepslechthorendheid 503: hardhorendheid of doofheid ten gevolge van lawaai.* Amsterdam, Nederlands Centrum voor Beroepsziekten, Coronel Instituut voor Arbeid en Gezondheid, AMC, UvA.

Nederlands Centrum van Beroepsziekten (NCvB). 2011. *Registratierichtlijn B001 - Gehoorverlies door beroepsmatige blootstelling aan lawaai.* Amsterdam, Nederlands Centrum voor Beroepsziekten, Coronel Instituut voor Arbeid en Gezondheid, AMC, UvA.

Neitzel, R.L., Seixas, N., Goldman, B. & Daniell, W. 2004. Contributions of non-occupational activities to total noise exposure of construction workers. *Ann Occup Hyg*, 48, 463-473.

Neitzel, R.L. & Seixas, N. 2005. The effectiveness of hearing protection among construction workers. *J Occup Environ Hyg*, 2, 227-238.

Neitzel, R.L., Stover, B. & Seixas, N.S. 2011. Longitudinal assessment of noise exposure in a cohort of construction workers. *Ann Occup Hyg*, 55, 906-916.

Nelson, D.I., Nelson, R.Y, Concha-Barrientos, M. & Fingerhut M. 2005. The Global Burden of Occupational Noise-Induced Hearing Loss. *Am J Ind Med*, 48, 446–458.

Niskar, A.S., Kieszak, S.M., Holmes, A.E., Esteban, E., Rubin, C. & Brody, D.J. 2001. Estimated prevalence of noise-induced hearing threshold shifts among children 6 to 19 years of age: the Third National Health and Nutrition Examination Survey, 1988-1994, United States. *Pediatrics*, 108, 40-43.

Nordmann, A.S., Bohne, B.A. & Harding, G.W. 2000. Histopathological differences between temporary and permanent threshold shift. *Hear Res*, 139, 13-30.

Norton, S.J. 1992. Cochlear function and otoacoustic emissions. *Seminars in Hearing*, 13, 1-14.

NVAB. 2006. *Multidisciplinaire richtlijn Preventie beroepsslechthorendheid door een effectief gehoorbeschermings-programma*. Utrecht, NVAB.

Ozimek, E., Kutzner, D., Sek, A. & Wicher, A. 2009. Polish sentence tests for measuring the intelligibility of speech in interfering noise. *Int J Audiol*, 48, 433-443.

Passchier-Vermeer, W. 1986. The effects of age, otological factors and occupational noise exposure on hearing threshold levels of various populations. *In*: Salvi, R.J., Henderson, D., Hamernik, P. & Coletti, V. (eds). *Basic and applied aspects of noise-induced hearing loss*. New York, Plenum Press, pp 571-581.

Passchier-Vermeer, W., van Hof, W. & Rovekamp, A.J.M. 1991. *Het gehoor van werknemers in de bouwnijverheid*. Leiden, TNO.

Pekkarinen, E., Salmivalli, A. & Suonpaa, J. 1990. Effect of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing. *Scand Audiol*, 19, 31-36.

Peters, R.W., Moore, B.C. & Baer, T. 1998. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J Acoust Soc Am*, 103, 577-587.

Phillips, D.P., Rappaport, J.M. & Gulliver, J.M. 1994. Impaired word recognition in noise by patients with noise-induced cochlear hearing loss: contribution of temporal resolution defect. *Am J Otol*, 15, 679-686.

Plack, C.J. 2005. *The sense of hearing*. New York, Lawrence Erlbaum Associates Publishers.

Plomp, R. & Mimpen, A.M. 1979a. Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, 18, 43-52.

Plomp, R. & Mimpen, A.M. 1979b. Speech-reception threshold for sentences as a function of age and noise level. *J Acoust Soc Am*, 66, 1333-1342.

Plomp, R. 1986. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *J Speech Hear Res*, 29, 146-154.

Prince, M.M. 2002. Distribution of risk factors for hearing loss: implications for evaluating risk of occupational noise-induced hearing loss. *J Acoust Soc Am*, 112, 557-567.

Prince, M.M., Gilbert, S.J., Smith, R.J. & Stayner, L.T. 2003. Evaluation of the risk of noise-induced hearing loss among unscreened male industrial workers. *J Acoust Soc Am*, 113, 871-880.

Purdy, S. & Williams, W. 2002. Development of the Noise at Work Questionnaire to assess perceptions of noise in the workplace. *J Occup Health safety - Aust NZ*, 18, 77 - 83.

Quist-Hanssen, S., Thorud, E. & Aasand, G. 1979. Noise-induced hearing loss and the comprehension of speech in noise. *Acta Otolaryngol Suppl*, 360, 90-95.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rabinowitz, P.M. 2000. Noise-Induced Hearing Loss. *Am Fam Physician*, 61, 2749-2756.

Rabinowitz, P.M., Slade, M.D., Galusha, D., Dixon-Ernst, C. & Cullen, M.R. 2006. Trends in the prevalence of hearing loss among young adults entering an industrial workforce 1985-2004. *Ear Hear*, 27, 369-375.

Rabinowitz, P.M., Galusha, D., Dixon-Ernst, C., Slade, M.D. & Cullen, M.R. 2007. Do ambient noise exposure levels predict hearing loss in a modern industrial cohort? *Occup Environ Med,* 64, 53-59.

Rabinowitz, P.M., Galusha, D., Kirsche, S.R., Cullen, M.R., Slade, M.D. & Dixon-Ernst, C. 2011. Effect of daily noise exposure monitoring on annual rates of hearing loss in industrial workers. *Occup Environ Med*, 68, 414-418.

Rabinowitz, P.M., Galusha, D., McTague, M.F., Slade, M.D., Wesdock, J.C. & Dixon-Ernst, C. 2012. Tracking occupational hearing loss across global industries: a comparative analysis of metrics. *Noise Health*, 14, 21-27.

Rhebergen, K.S., Versfeld, N.J. & Dreschler, W.A. 2006. Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *J Acoust Soc Am*, 120, 3988-3997.

Rhebergen, K.S., Versfeld, N.J. & Dreschler, W.A. 2008 Learning effect observed for the speech reception threshold in interrupted noise with normal hearing listeners. *Int J Audiol*, 47, 185-188.

Rhebergen, K.S., Lyzenga, J., Dreschler, W.A. & Festen, J.M. 2010a. Modelling speech intelligibility in quiet and noise in listeners with normal and impaired hearing. *J Acoust Soc Am*, 127, 1570-1583.

Rhebergen, K.S., Versfeld, N.J., de Laat, J.A.P.M. & Dreschler, W.A. 2010b. Modelling the speech reception threshold in non-stationary noise in hearing-impaired listeners as a function of level. *Int J Audiol*, 49, 856-865.

Robinette, M.S., Cervette, M.J. & Probst, R. 2007. Otoacoustic emissions and audiometric outcomes across cochlear and retrocohlear pathology. *In*: Robinette, M.S. & Glattke, T.J. (eds) *Otoacoustic Emissions: Clinical Applications*. 3rd ed. New York, Thieme Medical Publishers, pp. 197-275.

van Rooij, J.C.G.M. & Plomp, R. 1992. Auditive and cognitive factors in speech perception by elderly listeners. III. Additional data and final discussion. *J Acoust Soc Am*, 91, 1028-1033.

Rösler, G. 1994. Progression of hearing loss caused by occupational noise. *Scand Audiol*, 23, 13-37.

Royster, L.H., Lilley, D.T. & Thomas, W.G. 1980. Recommended criteria for evaluating the effectiveness of hearing conservation programs. *Am Ind Hyg Assoc J*, 41, 40-48.

Royster, J.D. & Royster, L.H. 1986. Using audiometric database analysis. *J Occup Med*, 28, 1055-1068.

Royster, J.D. 2003. Audiometric monitoring phase of the HCP. *In*: Berger, E.H, Royster, L.H., Royster, J.D., Driscoll, D.P. & Layne M (eds). *The noise manual, revised 5th edition*. Fairfax, Virginia, USA: American Industrial Hygiene Association, pp. 455-516.

Sataloff, T.F. & Sataloff, J. 1993. *Hearing loss; third edition, revised and expanded*. 3rd ed. New York, Taylor & Francis.

Sbihi, H., Teschke, K., MacNab, Y.C. & Davies, H.W. 2010. An investigation of the adjustment of retrospective noise exposure for use of hearing protection devices. *Ann Occup Hyg*, 54, 329-339.

Schlauch, R.S. & Carney, E. 2012. The challenge of detecting minimal hearing loss in audiometric surveys. *Am J Audiol*, 21, 106-119.

Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR). 2008. *Potential health risks of exposure to noise from personal music players and mobile phones including a music playing function*. Brussels, Belgium; European Commission.

Seixas, N.S., Kujawa, S.G., Norton, S., Sheppard, L., Neitzel, R. & Slee, A. 2004. Predictors of hearing threshold levels and distortion product otoacoustic emissions among noise exposed young adults. *Occup Environ Med*, 61, 899-907.

Seixas, N.S., Goldman, B., Sheppard, L., Neitzel, R., Norton, S. & Kujawa, S.G. 2005. Prospective noise induced changes to hearing among construction industry apprentices. *Occup Environ Med,* 62, 309-317.

Seixas, N.S., Neitzel, R., Stover, B., Sheppard, L., Daniell, B., Edelson, J. & Meischke, H. 2011. A multi-component intervention to promote hearing protector use among construction workers. *Int J Audiol*, 50 Suppl 1, S46-56.

Seixas, N.S., Neitzel, R., Stover, B., Sheppard, L., Feeney, P., Mills, D. & Kujawa, S.G. 2012. 10-Year prospective study of noise exposure and hearing damage among construction workers. *Occup Environ Med*, 69, 643-650.

Seren, E. 2009. Web-based hearing screening test. *Telemed J E Health*, 15, 678-681

Shah, S., Gopal, B., Reis, J. & Novak, M. 2009. Hear today, gone tomorrow: an assessment of portable entertainment player use and hearing acuity in a community sample. *J Am Board Fam Med*, 22, 17-23.

Shargorodsky, J., Curhan, S.G., Curhan, G.C. & Eavey, R. 2010. Change in prevalence of hearing loss in US adolescents. *JAMA*, 304, 772-778.

Sliwińska-Kowalska, M. & Jedlinska, U. 1998. Prolonged exposure to industrial noise: cochlear pathology does not correlate with the degree of permanent threshold shift, but is related to duration of exposure. *J Occup Health*, 40, 123-131.

Sliwińska-Kowalska, M., Dudarewicz, A., Kotylo, P., Zamyslowska-Szmytke, E., Pawlaczyk-luszczyńska, M. & Gajda-Szadkowska, A. 2006. Individual susceptibility to noise-induced hearing loss: choosing an optimal method of retrospective classification of workers into noise-susceptible and noise-resistant groups. *Int J Occup Med Environ Health*, 19, 235-245.

Sliwińska-Kowalska, M. & Kotylo, P. 2007. Evaluation of individuals with known or suspected noise damage to hearing. *Audiological Medicine*, 5, 54-65.

Sliwińska-Kowalska, M. & Davis, A. 2012. Noise-induced hearing loss. *Noise Health*, 14, 274-280.

Smits, C., Kapteyn, T. S., Houtgast, T. 2004. Development and validation of an automatic speech-in-noise screening test by telephone. *Int J Audiol*, 43, 15-28.

Smits, C. & Houtgast, T. 2005. Results from the Dutch speech-in-noise screening test by telephone. *Ear Hear*, 26, 89-95.

Smits, C., Merkus, P. & Houtgast, T. 2006a. How we do it: The Dutch functional hearing-screening tests by telephone and internet. *Clin Otolaryngol*, 31, 436-440.

Smits, C., Kramer, S.E. & Houtgast, T. 2006b. Speech reception thresholds in noise and self-reported hearing disability in a general adult population. *Ear Hear*, 27, 538-549.

Smits, C. & Houtgast, T. 2006. Measurements and calculations on the simple up-down adaptive procedure for speech-in-noise tests. *J Acoust Soc Am*, 120, 1608-1621.

Smits, C. & Houtgast, T. 2007. Recognition of digits in different types of noise by normal-hearing and hearing-impaired listeners. *Int J Audiol*, 46, 134-144.

Smits, C. & Festen, J.M. 2011. The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: steady-state noise. *J Acoust Soc Am*, 130, 2987-2998.

Smoorenburg, G.F., de Laat, J.A.P.M. & Plomp, R. 1982. The effect of noise-induced hearing loss on the intelligibility of speech in noise. *Scand Audiol Suppl*, 16, 123-133.

Smoorenburg, G.F. 1990. On the limited transfer of information with noise-induced hearing loss. *Acta Otolaryngol Suppl*, 469, 38-46.

Smoorenburg, G.F. 1992. Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *J Acoust Soc Am*, 91, 421-437.

Soli, S.D. & Wong, L.L. 2008. Assessment of speech intelligibility in noise with the Hearing in Noise Test. *Int J Audiol*, 47, 356-361.

Staatsblad 56. 2006. *Besluit van 25 januari 2006 tot wijziging van het Arbeidsomstandighedenbesluit, houdende regels met betrekking tot de blootstelling van werknemers aan de risico's van lawaai*. Staatsblad van het koninkrijk der Nederlanden.

Sorgdrager, B. & Dreschler, W.A. 2010. Lawaaislechthorendheid. *In:* van der Laan, G., Pal, T.M. & Bruynzeel, D.P. (eds) *Beroepsziekten in de praktijk*. 2nd ed. Amsterdam, Elsevier Gezondheidszorg, pp 229-244.

Stenfelt, S., Janssen, T., Schirkonyer, V. & Grandori, F. 2011. E-health technologies for adult hearing screening. *Audiology Research*, 1, 55-57.

Strasser, H., Irle, H. & Legler, R. 2003. Temporary hearing threshold shifts and restitution after energy-equivalent exposures to industrial noise and classical music. *Noise Health*, 5, 75-84.

Stuart, A. & Phillips, D.P. 1996. Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners. *Ear Hear*, 17, 478-489.

Stuart, A. & Phillips, D.P. 1997. Word recognition in continuous noise, interrupted noise, and in quiet by normal-hearing listeners at two sensation levels. *Scand Audiol*, 26, 112-116.

Studebaker, G.A., Sherbecoe, R.L., McDaniel, D.M. & Gwaltney, C.A. 1999. Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am*, 105, 2431-2444.

Summers, V. & Molis, M.R. 2004. Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. *J Speech Lang Hear Res*, 47, 245-256.

Summers, V. & Cord, M.T. 2007. Intelligibility of speech in noise at high presentation levels: effects of hearing loss and frequency region. *J Acoust Soc Am*, 122, 1130-1137.

Suter, A.H. 2002. Construction noise: exposure, effects, and the potential for remediation; a review and analysis. *AIHA J (Fairfax, Va),* 63:768-789.

Swanepoel, D.W. & Hall III, J.W. 2010. A systematic review of telehealth applications in audiology. *Telemed J E Health*, 16, 181-200.

Tak, S. & Calvert, G. 2008. Hearing difficulty attributable to employment by industry and occupation: an analysis of the National Health Interview Survey – United States, 1997 to 2003. *J Occup Environ Med*, 50, 46-56.

Talaska, A.E. & Schacht, J. 2007. Mechanisms of noise damage to the cochlea. *Audiological Medicine*, 5, 3-9.

Taylor, W., Pearson, J., Mair, A. & Burns, W. 1965. Study of noise and hearing in jute weaving. *J Acoust Soc Am*, 38, 113-120.

Theunissen, M., Swanepoel, D.W. & Hanekom, J. 2009. Sentence recognition in noise: Variables in compilation and interpretation of tests. *Int J Audiol*, 48, 743-757.

Torre III, P. 2008. Young adults' use and output level settings of personal music systems. *Ear Hear*, 29, 791–799.

Toppila, E., Pyykko, I., Starck, J., Kaksonen, R. & Ishizaki, H. 2000. Individual risk factors in the development of noise-induced hearing loss. *Noise Health*, 2, 59-70.

Tufts, J.B., Weathersby, P.K. & Marshall, L. 2009. Estimation of equivalent noise exposure level using hearing threshold levels of a population. *Ear Hear*, 30, 287-290.

Vaillancourt, V., Laroche, C., Mayer, C., Basque, C., Nali, M., Eriks-Brophy, A., Soli, S.D. & Giguère C. 2005. Adaptation of the HINT (hearing in noise test) for adult Canadian Francophone populations. *Int J Audiol*, 44, 358-369.

van Hoesel, R.J.M. & Litovsky, R.Y. 2011. Statistical bias in the assessment of binaural benefit relative to the better ear. *J Acoust Soc. Am*, 130, 4082-4088.

van der Molen, H., van der Laan, G., Lenderink, A & Spreeuwers D. 2009. *Beroepsziekten in cijfers 2009*. Amsterdam, Nederlands Centrum voor Beroepsziekten, Coronel Instituut voor Arbeid en Gezondheid, AMC, UvA.

van der Molen, H., van der Laan, G., Lenderink, A & Spreeuwers D. 2010. *Beroepsziekten in cijfers 2010*. Amsterdam, Nederlands Centrum voor Beroepsziekten, Coronel Instituut voor Arbeid en Gezondheid, AMC, UvA.

van der Molen, H., van der Laan, G., Lenderink, A & Spreeuwers D. 2011. *Beroepsziekten in cijfers 2011.* Amsterdam, Nederlands Centrum voor Beroepsziekten, Coronel Instituut voor Arbeid en Gezondheid, AMC, UvA.

van der Molen, H. & Lenderink, A. (Eds.). 2012. *Beroepsziekten in cijfers 2012.* Amsterdam:, Nederlands Centrum voor Beroepsziekten, Coronel Instituut voor Arbeid en Gezondheid, AMC, UvA.

van Son, N. & Jellema, I.J. 2011. *Analyse van de hoortesten van de Nationale Hoorstichting 2010.* Leiden, Nationale Hoorstichting.

van Wieringen, A. & Wouters, J. 2008. LIST and LINT: sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. *Int J Audiol*, 47, 348-355.

Versfeld, N.J., Daalder, L., Festen, J.M. & Houtgast, T. 2000. Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J Acoust Soc Am*, 107, 1671-1684.

Versfeld, N.J. & Dreschler, W.A. 2002. The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *J Acoust Soc Am*, 111, 401-408.

Vogel, I., Brug, J., van der Ploeg, C.P. & Raat, H. 2007. Young people's exposure to loud music: a summary of the literature. *Am J Prev Med*, 33, 124-133.

Vogel, I., Brug, J., Hosli, E.J., van der Ploeg, C.P. & Raat, H. 2008. MP3 players and hearing loss: adolescents' perceptions of loud music and hearing conservation. *J Pediatr*, 152, 400-404.

Vogel, I., Brug, J., van der Ploeg, C.P. & Raat, H. 2009. Strategies for the prevention of MP3-induced hearing loss among adolescents: expert opinions from a Delphi study. *Pediatrics*, 123, 1257-1262.

Vogel, I., Verschuure, H., van der Ploeg, C.P., Brug, J. & Raat, H. 2010. Estimating adolescent risk for hearing loss based on data from a large school-based survey. *Am J Public Health*, 100, 1095-1100.

Wagener, K.C. & Brand, T. 2005. Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: influence of measurement procedure and masking parameters. *Int J Audiol*, 44, 144-156.

Watson, C.S., Kidd, G.R., Miller, J.D., Smits, C. & Humes, L.E. 2012. Telephone screening tests for functionally impaired hearing: current use in seven countries and development of a US version. *J Am Acad Audiol*, 23, 757-767.

Weir, J.P. 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, 19, 231–240.

Widen, S.E., Holmes, A.E., Johnson, T., Bohlin, M. & Erlandsson, S.I. 2009. Hearing, use of hearing protection, and attitudes towards noise among young American adults. *Int J Audiol*, 48, 537-545.

Wild, D.C., Brewster, M.J. & Banerjee, A.R. 2005. Noise-induced hearing loss is exacerbated by long-term smoking. *Clin Otolaryngol*, 30, 517-520.

Williams, W, Purdy, S.C., Murray, N., Dillon, H., LePage, E., Challinor, K. & Storey, L. 2004. Does the presentation of audiometric test data have a positive effect on the perceptions of workplace noise and noise exposure avoidance? *Noise Health*, 24, 75-84.

World Health Organization (WHO). 2006. *Primary ear and hearing care. Training resource. Advanced level* . Geneva: WHO.

World Health Organization (WHO). 2012. *Factsheet no 300: deafness and hearing impairment.* Geneva: WHO.

Zhao, F., Manchaiah, V.K., French, D. & Price, S.M. 2010. Music exposure and hearing disorders: an overview. *Int J Audiol*, 49, 54-64.

Zohar, D., Cohen, A. & Azar, N. 1980. Promoting increased use of ear protection in noise through information feedback. *Hum Factors*, 22, 69-79.

Zokoll, M.A., Wagener, K.C., Brand, T., Buschermohle, M. & Kollmeier, B. 2012. Internationally comparable screening tests for listening in noise in several European languages: The German digit triplet test as an optimization prototype. *Int J Audiol,* 51, 697-707.

List of abbreviations

ARHL	age-related hearing loss
AUC	area under the curve
CI	confidence interval
CVC	consonant vowel consonant
dB	decibel
dBA	decibel A-weighted
EC	Earcheck
EOAE	evoked otoacoustic emissions
HI	hearing-impaired
HP	high-pass
HPD	hearing protection device
HTL	hearing threshold level
ICC	intraclass correlation coefficient
ISO	international organization for standardization
IHC	inner hair cells
kHz	kilohertz
LP	low-pass
MDD	minimum detectable difference
NF	noise floor
NH	normal-hearing
NHF	national hearing foundation
NHT	national hearing test
NIHL	noise-induced hearing loss
NIPTS	noise-induced permanent threshold shift
OAE	otoacoustic emissions
OEC	occupational Earcheck
OHC	outer hair cells
PI	performance intensity
POHE	periodic occupational health examination
PTA	pure-tone average
PTS	permanent threshold shift
RMS	root mean square
ROC	receiver operating characteristics
ROS	reactive oxygen species
SD	standard deviation
SEM	standard error of measurement
SII	speech intelligibility index
SNR	signal-to-noise ratio
SRT	speech reception threshold
TTS	temporary threshold shift
TWA	time-weighted average

Summary

Noise-induced hearing loss (NIHL) is a highly prevalent public health problem, caused by exposure to loud noises both during leisure time, e.g. by listening to loud music, and during work. In the past years NIHL was the most commonly reported occupational disease in the Netherlands. Hearing damage caused by noise is irreversible, but largely preventable. The early detection of hearing loss is of great importance, and is applied by preventative testing of hearing ability. This thesis investigates methods of screening for hearing impairment that can be applied in occupational medicine.

Chapter 1 provides a general introduction regarding hearing loss (prevention), hearing protection programs in occupational health, and possible techniques that can be used for screening and monitoring of NIHL.

Part 1 of this thesis focuses on analysing (the development of) hearing loss in a very large group of noise exposed workers in the construction industry, measured by regular screening audiometry, often performed during a periodic occupational health examination.

Chapter 2 describes a cross-sectional analysis of pure-tone audiometric data from approximately 30,000 employees. They are frequently exposed to noise levels above 80 dBA during their work, and the results show that these construction workers have poorer hearing threshold levels than peers who are not exposed to noise. Noise-induced hearing damage increases with increasing exposure duration, but relates to the (estimated) noise intensity to a lesser extent. Additionally, the relationship between noise exposure and hearing loss is compared to model predictions (ISO-1999, 1990), showing that particularly the younger workers exhibit a greater deterioration in hearing than predicted.

A longitudinal analysis of follow-up data from about half of this cohort of workers is described in **Chapter 3,** and shows that hearing further deteriorates with continuous noise exposure. This worsening progresses with increasing exposure level. A further analysis of the data shows that the quality of the screening audiometry is highly variable, due to errors in data processing and/or varying testing conditions. For accurate pure-tone audiometry a quiet environment and good calibration are essential. In addition, a noise-free period prior to testing is required to eliminate temporary effects of noise. Because in practice these requirements are not always met, the sensitivity of screening audiometry to detect small differences in individual cases is limited.

Part 2 of this thesis, therefore, presents and investigates an alternative or additional approach for hearing screening. Online speech-in-noise tests are able to reach a broad public, and offer the possibility of an easily accessible self-test that can be

performed at any given time. The online speech-in-noise test Earcheck (www.oorcheck.nl) determines the speech reception in noise by presenting monosyllables in a stationary noise. Because of its supra-threshold stimuli and its relative insensitivity to test conditions, Earcheck seems to be a suitable screening tool for hearing impairment.

An evaluation study described in **Chapter 4** examines the value of this test for early detection of NIHL. Results of the screening test of normal-hearing and hearing-impaired listeners are compared to pure-tone audiometry and the standard Dutch sentence SRT test Plomp and Mimpen (1979a). These results show that the original Earcheck is not sensitive enough to adequately distinguish between normal-hearing listeners and hearing-impaired subjects with mild NIHL, and therefore is not suitable to be used for NIHL screening.

Chapter 5 studies ways to improve this sensitivity, in order to employ Earcheck as a reliable screening tool. First, the intelligibility of the stimuli is homogenized, which leads to a slightly better test reliability. In addition, the broadband stationary masking noise of the original test is replaced by five modulated noises with different properties. The use of a low-pass filtered masking noise emphasizes the use of high-frequency speech information. By using this type of noise instead of a broadband noise the sensitivity of Earcheck is significantly improved from 51% to 95%, while maintaining a high specificity (98%) and reliability.

Part 3 of this thesis focuses on the application of this online screening test of hearing in occupational medicine. When the self-test is performed at home, uncontrollable parameters, such as the presentation level and the characteristics of the equipment used, can influence the test outcomes due to the noise filtering.

Chapter 6 describes an implementation study that investigates the correspondence between field results of Earcheck with low-pass filtered noise and results obtained in a controlled laboratory environment. Earcheck is conducted in different test conditions both at home and in the lab, and results show that hearing-impaired listeners perform better at home than in a controlled lab situation, probably because they perform the domestic Earcheck at higher presentation levels. Normal-hearing performance was poorer at home. As a result, the discriminative power of the domestic self-test is reduced compared to results obtained in the lab. Based on the results of this study we recommend conducting the online Earcheck monotically using headphones.

Chapter 7 presents the results of the validation of the online Earcheck at home. A study among 210 occupationally noise-exposed workers shows a sensitivity of 68% and a specificity of 71% of Earcheck for detecting NIHL compared to the regular screening audiogram. Also, the reliability of domestic testing was somewhat poorer than of lab testing.

Therefore, **Chapter 8** presents recommendations to improve Earcheck's sensitivity and reliability, based on analysis and our experiences, for example by the inclusion of a number of practice lists and adaptations in the testing procedure. Future research should confirm this.

At this point, Earcheck can not yet replace the screening audiogram, but this easily accessible online test provides a good complementary screening method, which can be implemented and used in different ways.

Samenvatting

Lawaaislechthorendheid is een groot sociaal gezondheidsprobleem, dat wordt veroorzaakt door blootstelling aan harde geluiden zowel tijdens de vrije tijd, bijvoorbeeld door luisteren naar luide muziek, als tijdens werk. De laatste jaren is beroepsslechthorendheid dan ook de meest gemelde beroepsziekte. Slechthorendheid door lawaai is irreversibel, maar wel grotendeels te voorkomen. Het vroegtijdig ontdekken van gehoorverlies is dan ook van groot belang en is mogelijk door het gehoor preventief te testen. Deze thesis onderzoekt screeningsmethoden voor lawaaislechthorendheid die kunnen worden toegepast in de arbeidsgeneeskunde.

Hoofdstuk 1 geeft een algemene inleiding met betrekking tot (preventie van) lawaaislechthorendheid, gehoorbeschermingsprogramma's in de arbeidsgeneeskunde en mogelijke instrumenten die ingezet kunnen worden voor screening en monitoring van gehoorverlies door lawaai.

Het eerste deel van deze thesis richt zich op de analyse van (de ontwikkeling van) gehoorschade in een zeer grote groep beroepsmatig blootgestelde werknemers in de bouwnijverheid, gemeten door middel van de reguliere screeningsaudiometrie, zoals dat vaak wordt uitgevoerd tijdens een periodiek arbeidsgeneeskundig onderzoek.

Hoofdstuk 2 beschrijft een cross-sectionele analyse van data van ongeveer 30.000 werknemers verkregen door middel van een toonaudiogram. Zij worden tijdens hun werk veelvuldig blootgesteld aan geluidniveaus boven 80 dBA, en de resultaten laten dan ook zien dat werknemers in de bouwnijverheid een slechter gehoor hebben dan leeftijdsgenoten die niet in lawaai werken. De lawaaischade neemt vooral toe bij een langere blootstellingduur, en is in mindere mate gerelateerd aan de (geschatte) hoogte van het expositieniveau. Ook wordt de relatie tussen lawaaiblootstelling en gehoorverlies vergeleken met model voorspellingen (ISO-1999, 1990), waaruit blijkt dat vooral de jongere werknemers meer gehoorverlies vertonen dan voorspeld.

Een longitudinale analyse van follow-up data van ongeveer de helft van dit cohort werknemers, beschreven in **Hoofdstuk 3**, laat zien dat het gehoor verder verslechtert bij voortdurende lawaaiblootstelling, en dat deze verslechtering groter is naarmate de intensiteit van de blootstelling toeneemt. Een nadere analyse van de data laat zien dat de kwaliteit van het screeningsaudiogram door registratiefouten en verschillen in testomstandigheden veel variatie vertoont. Voor accurate audiometrie zijn een stille meetomgeving en goede kalibratie vereist. Daarnaast is een geluidsvrije periode voorafgaand aan de test noodzakelijk om tijdelijke lawaai-effecten uit te sluiten. Omdat hier in de praktijk niet altijd aan kan worden voldaan, blijkt de gevoeligheid van screeningaudiometrie om in individuele gevallen kleine verschillen zichtbaar te maken beperkt.

In het tweede deel van deze thesis wordt daarom naar een alternatief dan wel aanvullend screeningsinstrument gezocht. Online spraak-in-ruistesten kunnen een groot publiek bereiken, en bieden de mogelijkheid om op een laagdrempelige manier zelf het gehoor te testen, op elk gewenst tijdstip. De online spraak-in-ruistest Oorcheck (www.oorcheck.nl) bepaalt het spraakverstaan in ruis door monosyllaben aan te bieden in een stationaire ruis. Door zijn bovendrempelige karakter en relatieve ongevoeligheid voor testomstandigheden lijkt Oorcheck geschikt als screeningsinstrument voor lawaaislechthorendheid.

Een evaluatieonderzoek beschreven in **Hoofdstuk 4** onderzoekt de waarde van deze test voor het vroegtijdig ontdekken van lawaaislechthorendheid. Resultaten van normaal- en slechthorenden op de screeningtesten zijn vergeleken met het toonaudiogram en de standaard spraak-in-ruis test van Plomp en Mimpen (1979a). Uit de resultaten blijkt dat de originele Oorcheck niet gevoelig genoeg is om onderscheid te maken tussen normaalhorenden en slechthorenden met een licht lawaai gerelateerd hoge-tonen verlies, en dus niet geschikt is om te gebruiken om te screenen op lawaaislechthorendheid.

In **Hoofdstuk 5** worden manieren onderzocht om deze gevoeligheid te verbeteren, zodat Oorcheck gebruikt kan worden als een betrouwbaar screeningsinstrument. Allereerst wordt de verstaanbaarheid van de stimuli gehomogeniseerd, wat leidt tot een iets grotere test betrouwbaarheid. Daarnaast wordt de continue breedbandige maskeerruis van de originele test vervangen door een vijftal gemoduleerde ruizen met verschillende eigenschappen. Het gebruik van een maskeerruis die is gefilterd met een laagdoorlaatfilter legt de nadruk op het gebruik van hoogfrequente spraakinformatie. Door een dergelijke ruis te gebruiken in plaats van een breedbandige ruis, werd de sensitiviteit van Oorcheck sterk verbeterd: van 51% naar 95%, met behoud van een hoge specificiteit (98%) en betrouwbaarheid.

Het derde deel van deze thesis richt zich op de toepassing van deze gehoorscreening in de arbeidsgeneeskunde. Wanneer de zelf-test thuis wordt uitgevoerd kunnen oncontroleerbare parameters, zoals bijvoorbeeld het afspeelniveau en de eigenschappen van de gebruikte apparatuur de testresultaten beinvloeden door de filtering van de ruis.

Om na te gaan in hoeverre de resultaten van Oorcheck met gefilterde ruis uit de praktijk overeenkomen met de resultaten verkregen in een gecontroleerde lab omgeving, wordt in **Hoofdstuk 6** een implementatieonderzoek uitgevoerd. Oorcheck wordt in verschillende condities zowel thuis als in het lab uitgevoerd en resultaten laten zien dat slechthorenden thuis beter scoren dan in een gecontroleerde lab-situatie, waarschijnlijk door het gebruik van hogere afspeelniveaus. Normaalhorenden presteren thuis juist slechter. Hierdoor is het discriminerend vermogen van de zelf-test in de thuissituatie verminderd ten opzichte van resultaten in het lab. Op basis van de resultaten van dit onderzoek wordt geadviseerd om de Oorcheck via internet monotisch met een hoofdtelefoon af te nemen.

Hoofdstuk 7 geeft de resultaten van de validatie van de online Oorcheck in de thuissituatie weer. Een onderzoek onder 210 beroepsmatig aan lawaai blootgestelde werknemers toont een sensitiviteit van 68% en een specificiteit van 71% voor Oorcheck voor detectie van lawaaislechthorendheid in vergelijking met het reguliere screeningsaudiogram. Ook is de betrouwbaarheid van de thuistest wat minder goed dan in het lab.

Op basis van ervaringen en nadere analyses worden in **Hoofdstuk 8** dan ook aanbevelingen gedaan om de testgevoeligheid en betrouwbaarheid te vergroten, bijvoorbeeld door het opnemen van een aantal oefenlijsten en aanpassingen in de testprocedure. Toekomstig onderzoek moet dit nader onderbouwen.

Vooralsnog kan Oorcheck nog niet als vervanging voor het audiogram dienen, maar deze gemakkelijke en laagdrempelige online test biedt wel een goede aanvullende screeningsmethode, welke op verschillende manieren kan worden geimplementeerd en benut.

Dankwoord

Het is zover, mijn proefschrift is af! En zoals veel mensen al voor mij op deze plaats hebben gezegd; promoveren doe je niet alleen. Zonder de begeleiding, medewerking en steun van een aantal mensen was me dit niet gelukt. Daarvoor wil ik hen hier heel graag bedanken!

Allereerst natuurlijk mijn promotor, Wouter. Bedankt voor de kans die je me hebt geboden om als onderzoeker deze weg in de slaan, en voor het vertrouwen dat je me hebt gegeven het pad via verschillende onderzoeksprojecten te volgen naar deze daadwerkelijke promotie. Je had het al over een boekje toen dat voor mij nog een vaag begrip was in plaats van een eindstation, en ik dank je voor dit enthousiaste (en blijkbaar terechte) optimisme. Ik heb veel van je geleerd, en je kritische feedback op mijn werk heeft de kwaliteit ervan zeker verbeterd.

Mijn copromotor Koen, hoewel je betrokkenheid bij mijn projecten varieerde, was je interesse er altijd. Jouw hulp bij de opzet en realisatie van de eerste Oorcheck experimenten was onmisbaar, en je literatuurkennis en deskundigheid met betrekking tot spraak-in-ruis testen hebben me ontzettend goed op weg geholpen. Bedankt dat je in de laatste fase van mijn promotie de planning en praktische zaken goed in de gaten hield.

Verder wil ik de leden van de leescommissie, Prof. dr. S. van der Baan, Prof. dr. F.J.H. van Dijk, Prof. dr. ir. J.M. Festen, Dr. ir. J.A.P.M. de Laat, en Dr. B. Sorgdrager, bedanken voor het zorgvuldig lezen en beoordelen van mijn proefschrift.

Cor van Duivenbooden wil ik bedanken voor de prettige samenwerking en praktische input tijdens de verschillende Arbouw projecten. Veel dank aan Frans Meijer voor de snelle en adequate hulp bij de aanlevering van de verschillende databestanden. Je inzet bij de online dataverzameling was onmisbaar. Je hebt me veel werk uit handen genomen, en dankzij jouw hulp heb ik gemakkelijk vele bouwvakkers kunnen testen.

Het doen van wetenschappelijk onderzoek wordt zeker bevorderd wanneer het serieuze denkwerk op zijn tijd wordt afgewisseld met wat minder serieus geklets. Dank dan ook aan alle collega's van het AC die door de jaren heen gekomen, gegaan of gebleven zijn, voor de goede sfeer en de gezellige praatjes bij de koffieautomaat of tijdens pauzes. Miranda, ondanks het zeer vroege tijdstip waren onze meetdagen in de bouwkeet op locatie erg gezellig!

Natuurlijk wil ik mijn directe research collega's hartelijk bedanken voor hun hulp bij allerhande praktische zaken, voor hun feedback en luisterend oor. Maar ik wil

Femke, Jelmer, Jordy, Maj, Maaike van D., Maaike de V., Monique, Rolph en Thijs vooral bedanken voor de gezelligheid op en rond de afdeling, tijdens de dagelijkse treinreis of tijdens verschillende borrels en etentjes. Speciale dank aan Inge voor het audiometreren van bijna al mijn proefpersonen, en natuurlijk voor haar uitstekende koffie-service!

Door de jaren heen heb ik mijn paar vierkante meter AMC met veel kamergenoten gedeeld. Jullie waren allen fijn gezelschap, altijd bereid tot eerste hulp bij vragen of frustrerende (computer) zaken, maar ook om met allerlei verhalen voor welkome afleiding te zorgen.

Bastiaan, bedankt voor het warme welkome op de afdeling, en Noortje voor het op weg helpen in het preventieteam en je levendige verhalen. Thamar, zeer bedankt voor de gezelligheid op en naast de werkvloer. Het was fijn om met iemand allerlei promotie perikelen te delen. Ik wens je heel veel succes met jouw laatste loodjes! Marya, de tweede gezondheidswetenschapper op de afdeling, ondanks dat we maar kort kamergenoten waren vond ik dat gezellig en was het erg prettig om ideeën met je uit te wisselen.

En Hiske, van mede-preventist werd je mijn kamergenoot maar vooral een goede vriendin. Ik vond het ontzettend fijn om met je samen te werken, al gingen onze goed bedoelde discussies over een onderzoeksopzet of analysemethode vaak snel over in gesprekken over verhuizingen, vakantietips of kinderfoto's. Hoogtepunt was natuurlijk ons bezoek aan Orlando, en daarom blijft het erg jammer dat onze gezamenlijke trip naar Italië niet door kon gaan. Ik ben erg blij dat je nu weer met je onderzoek bezig bent, en vind het erg fijn dat je vandaag mijn paranimf wilt zijn om deze periode in stijl af te sluiten!

Natuurlijk mogen in dit dankwoord de mensen die het leven aangenaam maken niet ontbreken.

Dank aan al mijn lieve vrienden; ondanks dat jullie niet allemaal even goed wisten wat mijn promotie onderzoek nu precies inhield, waren jullie wel altijd geïnteresseerd. Veel belangrijker nog waren, én zijn, alle gezellige momenten die we samen beleven! Bedankt voor jullie gezelligheid, gesprekken, grappen en eindeloos herhaalde anekdotes waardoor alle etentjes, feestjes en weekendjes samen een gezellige en geslaagde afleiding vormden.

Robbin, Bram en Gijs, dankzij mijn test op afstand kon ik jullie gehoor ook testen, dank dat jullie aan mijn onderzoek mee wilden werken. Martijn, ontzettend bedankt dat je naar Amsterdam bent gekomen om in een paar uur in mijn onderzeeër te zitten!

Marja, Marijn en Marlieke, ik ben blij dat ik het rijtje doctors nu kan completeren. Ik vind het bijzonder en vooral heel leuk dat we al bijna 15 jaar vriendinnen zijn en hoop dan ook dat er nog vele M-dates zullen volgen!

Jessica, bedankt voor je interesse, goede gesprekken, maar vooral voor de lol die we vaak samen hebben. Ik vind het echt geweldig dat je me vandaag als paranimf wil bijstaan!

Ook de altijd aanwezige steun van mijn familie was heel fijn. Marieke en Mathieu, bedankt voor jullie interesse. Het is fijn te weten dat jullie altijd voor ons klaar staan, bijvoorbeeld als we weer eens gaan verhuizen. Lieve Matthijs en Maarten, ik ben er trots op jullie tante te zijn! Ik hoop dat jullie nog vaak langskomen om te bouwen, racen of logeren.

Lieve Ans en Willie, ik voel me altijd welkom bij jullie. Jullie aandacht, gezelligheid en lieve woorden op allerlei gebied zijn fantastisch. En Willie, bedankt voor de vele ritjes, vaak zwijgzaam maar toch prettig, ’s ochtends vroeg naar het station!

Lieve papa en mama, dank voor jullie liefdevolle steun en onvoorwaardelijke vertrouwen. Jullie hebben me altijd alle kansen geboden en mij ondersteund in mijn keuzes. Ik vind het erg fijn dat ik nog altijd voor allerlei (on)gevraagde adviezen bij jullie aan kan kloppen! Jullie zijn fantastische ouders, en daarom draag ik met liefde dit boekje op aan jullie.

En tenslotte Joost, mijn lief. Dankjewel dat je er voor me bent. Met jou is alles zoveel leuker! Je geloofde soms meer in mij dan ik dat zelf deed, en je goede zorgen , aansporing of afleiding kwamen vaak op het juiste moment. De periode van mijn promotieonderzoek is voor jou ook niet altijd even makkelijk geweest, niet alleen door de laatste stressvolle fase waarin ik opgesloten was in mijn ‘bubbel’, maar ook door mijn vijf jaar lange doordeweekse verblijf in Utrecht. Toch was jij altijd mijn echte thuis, en ik ben dan ook erg blij dat we daar nu ook in de letterlijke zin invulling aan kunnen geven. Op naar een geweldige toekomst samen!

Curriculum Vitae

Monique Leensen was born on July 27th 1981 in Nijmegen. After finishing secondary school at Stedelijk Gymnasium Nijmegen in 1999, she studied Biomedical Health Science at the Radboud University Nijmegen.

First, she finished a master in Human movement science, with a minor in Epidemiology; statistics and research methodology. In her first internship at Department of Physiology of the UMC St. Radboud she investigated peripheral muscle characteristics and blood flow in patients with Chronic Obstructive Pulmonary Disease. She finished her first MSc thesis at Sint Maartenskliniek research, during which she studied force regulation in one- and two-handed tasks in the elderly. Afterwards, she conducted a second master in Occupational and environmental health, and wrote her final thesis at Seneca, the research centre of Hogeschool Arnhem en Nijmegen. She described and evaluated factors affecting occupational rehabilitation of disabled persons.

Since 2007 she has worked as a scientific researcher at the Department of Clinical and Experimental Audiology at the Academic Medical Centre (AMC) Amsterdam. She conducted several research projects concerning the prevention and early detection of occupational noise-induced hearing loss. Most of these projects resulted in international publications which form the basis for this dissertation.

PhD Portfolio - AMC Graduate School for Medical Sciences

Summary of PhD training and teaching activities

Name PhD student:	M.C.J. Leensen
PhD period:	2007-2013
Promotor:	Prof. dr. ir. W.A. Dreschler

1. PhD training

	Year	Workload (ECTS)
General courses		
- Good Clinical Practice (GCP)	2007	0.3
- Career development	2012	0.8
Specific courses		
- Sensory systems - Helmholtz institute Utrecht	2007	1.4
- Advanced Topics in Biostatistics	2011	2.1
- Statistical computing in R	2011	0.4

Seminars, workshops and master classes

- Weekly department seminars	2007-2013	7.5
- Diverse master classes by external experts	2007-2012	1.0
- Workshop on Speech in Noise: Intelligibility and Quality	2010	0.4
- Attending meetings Werkgroep Auditief systeem (WAS)	2009-2012	0.5
- Attending meetings Nederlandse vereniging van Audiologie	2008-2012	1.0

Presentations

- '*Evaluatie van het gehoor*' – Gehoor en arbeid, nascholingsdag arbodeskundigen.	2008	0.5
- *'Screening van gehoorschade door lawaai met behulp van internettesten'* – Wintervergadering, Nederlandse vereniging voor Audiologie.	2009	0.5
- *'De Nederlandse ervaringen met de 'Nationale Hoortest': Onderzoek ten behoeven van een verhoging van de gevoeligheid van de test voor hogetonenverlies.'* – Congres Fonds voor beroepsziekten, België.	2009	0.5
- *'Speech-in-noise screening tests by internet; improving test sensitivity for noise-induced hearing loss.'* (poster) - ARHES meeting.	2009	0.5
- *'Speech-in-noise screening tests by internet; improving test sensitivity for noise-induced hearing loss'* – Workshop Speech-in-noise.	2010	0.5
- *'Hearing loss in construction industry: comparisons to ISO-1999 predictions'* - Annual conference National Hearing Conservation Association, USA.	2010	0.5
- *'Speech-in-noise screening tests by internet; improving test sensitivity for noise-induced hearing loss'* - Annual conference National Hearing Conservation Association, USA.	2010	0.5
- *'Toepassen van een spraak-in-ruis screeningstest in de arbeidsgeneeskunde'* – Wintervergadering, Nederlandse vereniging voor Audiologie.	2012	0.5
- *'The applicability of an internet-based speech-in-noise screening test in occupational hearing conservation.'* - Adult Hearing Screening Conference.	2012	0.5
- *'Toepassing van een online spraak-in-ruis screeningtest voor lawaaislechthorendheid'* - Arbouw deskundigendag.	2012	0.5

(Inter)national conferences

- Arbouw deskundigendag – November 2008, Soesterberg	2008	0.2
- Congres Fonds voor beroepsziekten – May 2009, Brussel, België	2009	0.3
- ARHES meeting – November 2009, Nottingham UK	2009	0.5
- 35th Annual conference National Hearing Conservation Association – February 2010, Orlando, USA	2010	0.8
- Adult Hearing Screening Conference, - June 2012, Cernobbio, Italy	2012	0.5

Other

- Journal clubs	2008-2012	1.0
- PhD Retreat Helmholtz institute	2007	1.0

2. Teaching

Tutoring / Mentoring

- Contact person students Spinoza Lyceum, Amsterdam during their writing assignment: *'The prevalence of noise induced hearing loss in adolescents, using an internet based speech-in-noise test.'*	2010	0.7

Supervising

- Supervising student Logopaedic and Audiological Sciences, 4th year KU Leuven. Research traineeship: *'Evaluation of online hearing screening tests for noise-induced hearing loss.'*	2009	1.0
- Supervising medical student, 4th year UvA Amsterdam. Bachelor thesis: *'The effect of earphone style used when listening to personal listening devices on the music exposure of adolescents.'*	2011-2012	1.5
- Supervising medical student, 4th year UvA Amsterdam. Bachelor thesis: *'Influence of loud music on speech reception thresholds of disk jockeys.'*	2011-2012	1.5