Noise induced hearing loss: Screening with pure-tone audiometry and speech-in-noise testing
Leensen, M.C.J.

Citation for published version (APA):
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>
<thead>
<tr>
<th>Type of Earcheck</th>
<th>Study type (Chapter)</th>
<th>Learning effect (dB)</th>
<th>SEM (dB)</th>
<th>ICC*</th>
<th>Correlation PTA_{3A6}</th>
<th>Se for NIHL (%)</th>
<th>Sp for NIHL (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>Lab (4)</td>
<td>0.82</td>
<td>1.24</td>
<td>0.75</td>
<td>0.62</td>
<td>51</td>
<td>90</td>
</tr>
<tr>
<td>Homogenized</td>
<td>Lab (5)</td>
<td>1.62</td>
<td>1.17</td>
<td>0.84</td>
<td>0.76</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LP</td>
<td>Lab (5)</td>
<td>0.71</td>
<td>1.25</td>
<td>0.93</td>
<td>0.92</td>
<td>95</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Lab (6)*</td>
<td>0.54</td>
<td>1.62</td>
<td>0.81</td>
<td>0.83†</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Field (7)</td>
<td>1.40</td>
<td>1.63</td>
<td>0.65</td>
<td>0.56</td>
<td>68</td>
<td>71</td>
</tr>
<tr>
<td>LPmod</td>
<td>Lab (5)</td>
<td>0.48</td>
<td>1.39</td>
<td>0.87</td>
<td>0.88</td>
<td>94</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Lab (6)*</td>
<td>0.54</td>
<td>1.70</td>
<td>0.68</td>
<td>0.73†</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Field (7)</td>
<td>1.80</td>
<td>1.68</td>
<td>0.54</td>
<td>0.50</td>
<td>74</td>
<td>59</td>
</tr>
</tbody>
</table>

* 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 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 monotonic 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 parameters, 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 testing are 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.