Influence of age on speech intelligibility in babble noise

Libor Černý, Jan Vokřál* and Olga Dlouhá

Department of Phoniatrics, First Faculty of Medicine, Charles University and the General University Hospital in Prague, the Czech Republic

* Email: vokral@lf1.cuni.cz

Problems with hearing aids, particularly with regard to speech intelligibility in the presence of noise, are commonly reported by older individuals in everyday practice. The main goal of this study was to measure differences in speech intelligibility between older and younger people and to establish how speech intelligibility in competitive noise differs between younger and older populations with similar hearing status. More than 400 persons were tested using the Czech Test of Sentence Intelligibility in Babble Noise and divided into younger (40–65 years) and older (66–85 years) age groups. Test performance was compared between age groups based on subgroups stratified by SRT values (speech reception threshold in word audiometry in silence). Results showed a significant correlation between older age and diminished sentence intelligibility in competitive noise. Evaluation using a nonparametric U-test showed a statistical difference between the younger and older groups in sentence intelligibility, with a speech signal presented at 65 dB sound pressure level (SPL) and competitive babble noise also presented at 65 dB SPL. Increased difficulty in the use of hearing aids in older users is related, among other things, to a reduced ability to discriminate speech not only in silence but particularly in competitive variable noise due both to their aging auditory functions and to a diminished capacity to differentiate the time factors of sounds. It is probably connected with the diminished function of inhibitory neurons.

Key words: aging, central auditory processing, speech intelligibility in noise

INTRODUCTION


The differentiation of sounds for speech comprehension is a function of the higher levels of the brain’s hearing pathway. Factors that affect one’s ability to understand speech include: hearing threshold of speech frequencies, time decoding, working memory, concentration, and the ability to make associations. All of these functions can be negatively affected by brain aging.

The most common problem reported by unsatisfied hearing-aid users is speech intelligibility in a noisy environment (McCormack and Fortnum 2013, Sprinzl and Riechelmann 2010). Diminished speech intelligibility performance in competitive noise is associated with an age-related decrease in time discrimination ability in the auditory cortex, as demonstrated by inferior noise gap detection (Hesse et al. 2014, Kishon-Rabin et al. 2013, Ozmeral et al. 2016, Walton 2010). Similarly, compared to younger subjects, elderly and subjects and centennians demonstrate poorer temporal order judgment of auditory stimuli (Kolodziejczyk and Szeg 2008). Other studies have examined age-related decreases in recognition of speech in silent conditions, and found that the
most rapid declines occur between 65 - 70 years (Dubno 2015), or after 80 years (Hoppe et al. 2016). In contrast, at least one study reported only a small difference between young and elderly people in speech intelligibility in a quiet environment, but there is a significant difference in competitive environment, especially in non-continual noise (Humes and Dubno 2010). In addition, experiments have documented an age-related decrease in the inability to differentiate sound signal time factors, connecting this phenomenon with diminished speech comprehension at the central level (Walton 2010).

Over the past few decades, a growing body of research has examined age-related changes in cortical and subcortical functions associated with presbycusis. These studies, conducted primarily in experimental animals models, have shown a decreased ability to discriminate in the time domain with aging, accompanied by changes in GABAergic neurons and receptors (Burianová et al. 2015, Caspary et al. 2008, 2013). A series of studies have examined aging in the central auditory system of the F344 rat, particularly in γ-aminobutyric acid (GABA) neurotransmitter function in an important auditory midbrain structure, the inferior colliculus (IC), (Syka 2010). These studies found: decreased numbers of GABA immunoreactive neurons; decreased basal levels (concentrations) of GABA; decreased GABA release; decreased glutamic acid decarboxylase activity; decreased GABAB receptor binding; decreased numbers of presynaptic terminals; and subtle GABAA receptor binding changes. Collectively, these age-related changes suggest blunted GABA neurotransmitter function in the IC (Caspary et al. 1995). However, more recent studies failed to find the expected decreased concentrations of GABA in the auditory cortex of older human brains (Profant et al. 2013).

One animal model of aging in rats demonstrated an age-related decrease in the number of SMI-32 neurons (Burianová et al. 2009, Syka 2016) and inhibitory GABAergic interneurons that are associated with determining the time parameters of sounds (Syka 2016). These experiments suggest that the central auditory system is associated with a decreased capacity for inhibitory processes, which may negatively impact the ability to differentiate rapid changes in sound stimuli and to differentiate them from background noise (Syka 2016). Interestingly, however, the same GABAergic inhibitory system and the same types of neurons have also been shown to deteriorate in Alzheimer’s disease (Morrisson and Hof 2002, Oyelami et al. 2016). Thus, the observed changes may be a more general phenomenon of degenerative changes in aging.

Complex sound stimuli, such as speech, have been shown to activate non-auditory parts of the cortex more in elderly individuals than in younger individuals (Eckert et al. 2008). The frontal cortex in particular has been shown to be more active in elderly people as compared to younger individuals, when listening to speech, especially under conditions involving background noise (Wong et al. 2008). In another study by the same group, functional magnetic resonance imaging (fMRI) was used to examine cortical cerebral hemodynamics associated with spoken language processing in the aging brain. In this study, younger and older subjects identified single words in quiet and in two multi-talker babble noise conditions (signal-to-noise ratio, SNR=20 dB and -5 dB). Behaviorally, older and younger subjects did not show significant differences in the first two conditions but older adults performed less accurately in the SNR=-5 dB condition. The fMRI results showed reduced activation in the auditory cortex but an increase in working memory and attention-related cortical areas (prefrontal and pre-cuneus regions) in older compared to younger subjects, especially in the SNR=-5 dB condition. Increased cortical activation in general cortical regions was positively correlated with behavioral performance in older listeners, suggesting a compensatory neural response. Functional connectivity analyses revealed that, while younger subjects showed a more streamlined cortical network of auditory regions activated in response to spoken word processing in noise, older subjects showed a more diffuse network involving frontal and ventral brain regions. These results are consistent with the decline-compensation hypothesis of brain aging, and also suggest that this hypothesis applies to functions in the auditory domain (Wong et al. 2009).

The basic method for speech intelligibility testing is a word audiometry. In word audiometry several decades of single words are presented to a subject in a quiet booth, and each decade is presented with the same intensity level. The intensity of signal wherein the subject is able to repeat 50% of words correctly is marked as the speech reception threshold (SRT). The SRT value ranges between 15-20 dB sound pressure level (SPL) in normal hearing persons.

Based on the audiological literature, it is unclear as to whether single words or sentences are better for the examination of speech intelligibility (Benichov et al. 2012, Roesser 1996, Wagener and Brand 2005, Wilson and Cates 2008). Sentences are a more realistic type of signal for the evaluation of the perception of fluent speech. However, diminished cognitive function among elderly individuals could be a confound when sentences are longer, and particularly when examined in background noise (Tremblay and Ross 2007). On the other hand, the use of a set of single words is often criticized for its loss of the natural speech dynamic. In sentence test material there is the question of the predictability of sentence content due to syntactic, semantic, and prosodic
cues that could influence test performance (Hutcherson 1979). There is also a role for sentence topics. There are many tests of sentence recognition in noise developed with material based on conversational everyday speech (Dlouhá and Vokřál 2011, Hallgren et al. 2006). The use of competitive noise is another advantageous variation of tests, for evaluation in more realistic conditions. This is extremely useful for measuring the effect of hearing aids or for the study of auditory functions (Rawool 2016, Wong et al. 2009). Non-constant noise, such as multi-talker noise (babble noise), is more similar to the human voice and speech in terms of time factors. Thus, babble noise provides more masking at the central level than steady types of noise, such as white noise, pink noise, or speech-weighted noise matched for the tested sentence material (Kollmeier 1992, Krishnamurthy et al. 2008, Wilson et al. 2007). One novel test that includes babble noise is the Czech Test of Sentence Intelligibility in Babble Noise. The test was developed at the Department of Phoniatrics in Prague and is used for evaluating the effectiveness of hearing aids and cochlear implants (Dlouhá and Vokřál 2011, Dlouhá et al. 2012, Vokřál and Dlouhá 2009).

Using the Czech Sentence Intelligibility Test in Babble Noise, we tested the hypothesis that poorer speech intelligibility is associated with aging. We measured and controlled for effects related to severity of hearing impairment.

**METHODS**

**Subjects**

423 adults participated in this study. Subjects were divided into two groups based on their age (older > 65 years; younger ≤ 65 years). The older group consisted of 232 persons, ages 66 - 85 years ($M=75.4$ years). The younger group consisted of 191 persons, between the ages of 40 - 65 years ($M=55.8$ years). Subjects were eligible with or without hearing loss. Exclusion criteria were as follows: unilateral hearing loss or more pronounced hearing loss asymmetry, inability to achieve at least 50% understanding in a word audiometry test up to an intensity of 70 dB SPL. Also excluded were those with a relevant conductive component of hearing loss, with any indication of a dominant suprachlear (e.g., neurinoma of an acoustic nerve) or central hearing loss or with a neurologic disease or cognitive impairment.

All subjects were interviewed and we obtained their detailed medical history and all subjects were either long time follow-up patients of our clinic or family members of clinic staff.

**Test materials**

The Czech Test of Sentence Intelligibility in Babble Noise was developed at the Department of Phoniatrics, First Faculty of Medicine, Charles University and the General University Hospital in Prague as the first test of this type for the Czech language. The test consists of 100 sentences, ranging in length from 4 to 10 words, and topics were general. For this study, we used 3 groups of 10 sentences that were previously validated in persons with normal hearing. The length of sentences ranged from 1.6 – 3.5 sec, with a 6.5 sec pause between sentences. The recording was prepared under studio conditions and recorded digitally using the voice of a male professional speaker. The sentences were organized in decades and the average intensity of all the decades of sentences was regulated for the same level as per the norm EN ISO 8253-3: 1998. This standard allows for a maximum deviation in sentence intensity of above or below 3 dB from the average of sentences in a decade. For this study, we shortened the intensive interval to above or below 2 dB. In previous studies, investigators balanced the intensities of the separate sentences in such a way that all the sentences have the same root-mean-square (RMS) value. Of note, the RMS value is not related to intelligibility of the sentence, because it reflects only the energetic point of view and is agnostic to sentence contents. Thus, the RMS does not contain the “correct” value, of how the intensity of sentences should be balanced and set up. Here, all the average intensities of the sentence decades have the same average RMS value.

Calibration was performed in an audiometric test booth by a sound level meter and a calibration signal, which was presented via compact disk (CD) with speech material. The values measured on the sound level meter in dB SPL are the same as levels of presented speech material in dB SPL. We used the same calibration levels and RMS signal values as in Czech word audiometry.

The test evaluation is expressed as a percentage for each decade, taking into account the correctness of even just a part of a sentence (i.e., 25%, 50%, 75%).

In a 2011 study, Dlouhá and Vokřál constructed and published the audiological norms applicable to an 18 to 25 year-old normal hearing group. They found that the subjects understood 100% of sentences at SNR=0 dB. At SNR=-5 dB, sentence intelligibility in babble noise was 98.1% ±/− 2.3%.

Competitive noise was provided by babble noise, which consisted of the speech of 8 speakers (4 male, 4 female), also developed at the Department of Phoniatrics. We used this type of competitive noise for its
maximal interference with speech signal, both in spectral and time domains.

**Procedures**

All subjects underwent a basic ear canal investigation followed by tympanometry. Following this, subjects were tested in an audiometric booth using pure tone audiometry in a classical frequency range of 125-8000 Hz, with air conduction and bone conduction with a masking as per audiological rules using the Madsen – Orbiter 922 version 2 clinical audiometer.

An identical word audiometry test was given to subjects without a hearing aid and without any competitive noise in a silent booth using the Czech word audiometry test using a Technics SL-PG 580A CD player and Madsen – Orbiter 922 version 2 clinical audiometer. The test consisted of 1, 2, 3 and 4-syllable words, the most frequently-used in Czech, phonetically balanced, and presented in groups of 10 words at the same level (Seeman 1960). The speech signal was presented in a free field from Jamo A320 speakers in front of the listener at a distance of 1 m, 45 degree to the left and right. SRT values were determined from the word audiometry results, defined as the lowest level at which the person correctly repeated 50% of the words.

All subjects were examined using the Czech Test of Sentence Intelligibility in Babble Noise. The sentences were presented in a free field in a silent booth at 65 dB SPL, with competitive noise in two steps at 65 dB SPL and 70 dB SPL. The same technical equipment with the same calibration was used for all the tests, as described above.

For statistical evaluation, the non-parametric Mann-Whitney U test, two-tailed (software Statistica 12 by StatSoft, Inc.) was used.

**RESULTS**

Fig. 1 shows a histogram of subject age stratification, while Fig. 2 illustrates the relationship between age and results of word audiometry in a free field. With age, there was a trend for an increase in SRT as demonstrated by a straight line fit (approximation was done by least square method, line parameters y=0.4596x+13.91, correlation coefficient 0.418). This trend led us to divide patients into subgroups according to SRT, to eliminate the dependency of SRT with age. Thus, only subgroups with no statistically significant difference of median SRT values were subsequently compared. To achieve this, we divided older and younger age groups into subgroups, stratified according to SRT values (i.e., results of word audiometry in quiet) in 10 dB steps: subgroup 1 (SRT 21 – 30 dB SPL), subgroup 2 (SRT 31 – 40 dB SPL), subgroup 3 (SRT 41 – 50 dB SPL), subgroup 4 (SRT 51 – 60 dB SPL), subgroup 5 (SRT 61 – 70 dB SPL).

Fig. 3 shows results of younger vs. older age groups in the Czech Test of Sentence Intelligibility in Babble Noise, as a boxplot. Sentences and babble noise were presented at the same level (65 dB SPL). The boxplot shows the main statistical parameters of subgroups: arithmetic mean, median, quartiles, and range of values without outliers.

Fig. 4 shows a relationship between the results of the Test of Sentence Intelligibility in Babble Noise (as a percentage of correctness) and the word audiometry results. A decreasing trend was approximated by a straight line (least square method, line parameters y=-1.695x+145.1, correlation coefficient -0.614).

The main results are aggregated into Table I. Subgroups (determined according to SRT values) are presented as lines in the table with information about the number of persons, their mean age, mean SRT, and results of the Test of Sentence Intelligibility in Babble Noise (mean and median values), for both older and younger subjects. Interestingly, results are relatively consistent across both groups. The results of the
Test of Sentence Intelligibility in Babble Noise were evaluated statistically by means of a non-parametrical U-test and the significance of difference between groups of younger and older persons is shown in the last column.

Across subgroups (except subgroup 1 with SRT<30 dB SPL), median sentence intelligibility was lower for older compared to younger subjects (P<0.05). There were no significant differences in results when the babble noise was presented at 70 dB SPL (see Table II).

The median percentages of correctness in both age groups did not statistically differ (Table II). Even among the younger group, correctness was low in higher noise conditions, making it difficult to differentiate between older and younger groups. Also, in cases of a higher severity of hearing loss, the subjects were unable to understand common conversation without hearing aids.

### DISCUSSION

#### Influence of test materials on test results

Previous studies have documented age-related declines in auditory temporal resolution and working memory, which negatively affect the discrimination of speech in background noises (Mills et al. 2006, Pichora-Fuller et al. 1995, 2006). Most studies in this area have utilized the Speech Perception in Noise (SPIN), which consists of both high-predictable and low-predictable sentences. The test used consists only of high-predictable sentences, because lower pre-
dictability has been shown to negatively influence speech intelligibility even in quiet conditions (Kirk et al. 1997). Matrix sentence tests are advantageous both in testing non-native listeners and for its lower cognitive function dependency (Kollmeier et al. 2015, Warzybok et al. 2015), but this type of test is not yet available in Czech. Moreover there was no need to use this material in our study as there were no non-native listeners examined.

In regards to multi-talker noise, previous studies have demonstrated that fluctuating noise has a similar effect on speech reception thresholds in different language-mutations of tests, tested in matrix sentence tests (Hochmuth et al. 2015). In addition, new variations of multi-talker babble, such as the use of competing coordinate response measure sentences, have masking properties similar to other types of multi-talker babble (Humes et al. 2017).

**Tested subjects**

Our cohort of subjects examined by the Czech Test of Sentences Intelligibility in Babble Noise is strongly evaluable due to the relatively large sample size (N=423). Data were gathered over the period 2012–2017, following the creation of the Test by our Department. Its validity had previously been demonstrated (Dlouhá and Vokřál 2011, Dlouhá et al. 2012). Identical test conditions were maintained throughout all data collection procedure. The internal structure of the cohort in terms of parameters beyond the audiological criteria, which was very strictly ensured in the exclusion criteria, was given by the structure of available probands. These were mostly Phoniatric Department patients, i.e., persons undergoing treatment for communication problems. Hearing-aid users in receipt of regular care are strongly represented in this group, along with a smaller number of patients with voice disorders and other conditions. However, patients with any indication of a central neurological problem, such as a cerebral stroke or a transient ischemic attack, or patients with a cognitive problem, were strictly excluded.

Only a minimal proportion of examined persons were specifically selected (for example family members and acquaintances of Phoniatric Department staff). Elderly people with no desire or ability to solve their hearing and other problems who thus have not sought any care, are absent from our cohort. In particular, seniors with mild hearing loss tend not to seek out care, rendering it difficult to acquire data from a sufficient number of subjects from this population. One should also take into account, when considering the exact degree of difficulty in a particular age group that our measured population were mainly inhabitants of the capital city, that is, Prague and its hinterland. This population, as per the populations of other large cities, generally exhibits better health status parameters than provincial populations.

One interesting direction for further research could be the question of the differences in the age-related determination of acoustic stimuli time factors between subjects with tonal and non-tonal languages. Indications suggested specific mechanisms regarding neuronal plasticity for perceiving the order of frequency-related auditory stimuli for tonal language speakers with or without a secondary non-tonal language experience (Bao et al. 2014). For example, there is a relatively large bilingual population of native tonal-language Vietnamese speakers in the Czech Republic.

**Evaluation of results**

Our main results showed a stratification of younger vs. older subjects based on single word perception in quiet conditions. Of note, we found a strong correlation between word audiometry results and results in tests of sentence intelligibility, where the higher the value of...
the SRT (word audiometry result) the poorer the sentence intelligibility in noise. Therefore, we subsequently tested only subjects with similar SRT values when evaluating whether older and younger subjects differ in their performance in the Test of Sentence Intelligibility according to their age.

To address this, we divided subjects into subgroups based on their SRT value (word audiometry results), and compared younger vs. older subjects with the same SRT values (resp. with SRT values in 10 dB intervals). This approach ensured that the significant differences between age groups in the Czech Test of Sentence Intelligibility in Babble Noise were not influenced by hearing status.

Interestingly, we found no significant differences between younger and older ages groups among those with good hearing (SRT up to 30 dB). We speculate that this null effect is due to the fact that hearing loss accelerates degeneration of central hearing pathways that are connected with a poorer performance in the Czech Test of Sentences Intelligibility in Babble Noise. There may not be a significant reduction in the number of functional neurons in older persons with good hearing to detect an impairment in the test.

Our results are largely consistent with previous studies by other groups (Humes and Dubno 2010, Tremblay and Ross 2007). Our study extends these important previous studies by utilizing sentence material built for the Czech language. Thus, this is the first study to examine age-related changes in sentence intelligibility in noise in the Czech language.

**CONCLUSION**

A significant difference in speech intelligibility in competitive inconstant noise (babble noise) was found between younger and older adults with comparable hearing status. A total number of 423 investigated persons was divided to two groups according to age (the borderline at 65 years) and stratified within the groups according to the results of a word audiometry test in a free field in silence. Evaluation using a nonparametric U-test showed a statistical difference (P<0.05) between the younger group (40 - 65 years, mean age 55.8 years) and the older group (66-85 years, mean age 75.4 years) in the intelligibility of sentences with both the speech signal and competitive babble noise presented at 65 dB SPL in a free field.

The influence of the hearing threshold and a simple decoding of speech sounds were excluded by comparing subgroups with the same score in the word audiometry in quiet. The main factor influencing sentence intelligibility in noise was a capacity for differentiating the time domain of speech and competitive noise: this was proved to be worse in the older group, which had a mean age 20 years higher than the younger group.

**REFERENCES**


Aging and speech intelligibility in babble noise


