Speech Intelligibility in Automobile Noise in Young and Middle-aged Adults

Aimée M. Surprenant[1], Patricia Davies[2], and Donald P. Gallant[2]
1-Psychology Department, Memorial University of NL, St. John's, NL, A1M 3T4
2-Ray W. Herrick Labs, 140 S Martin Jischke Drive, Purdue University, West Lafayette, IN 47907, USA

ABSTRACT

Two experiments investigated how automobile noise affects intelligibility of speech signals in both young and middle-aged individuals. In Experiment 1, the effect of automobile noise was compared to speech babble at a number of speech-to-noise ratios. In order to achieve the same intelligibility, the speech-to-noise ratio for the speech babble needed to be substantially greater than the automobile noise. In Experiment 2, middle-aged adults between the ages of 50 and 65 were given the sentences in automobile noise. Even though their hearing acuity was not severe enough to warrant a clinical diagnosis, their performance was significantly worse than the younger adults, particularly for sentences that had few contextual cues. In conclusion, although automobile noise is less damaging than speech babble at typical speech-to-noise ratios, speech understanding for individuals with even small amounts of hearing loss is significantly impacted by the noise. Automobile makers therefore should continue their efforts to reduce the noise levels in cars in order to increase speech intelligibility. [Work supported by Ford Motor Company].

SOMMAIRE

L'effet du bruit ambiant d'une voiture sur l'intelligibilité de la parole chez les jeunes adultes et les adultes d'âge moyen a été étudié à l'aide de deux études. L'expérience 1 comparait l'effet du bruit d'une automobile à un bruit de verbiage pour différent ratios parole-bruit. Les résultats ont révélé que pour atteindre le même niveau d'intelligibilité, le ratio parole-bruit dans le bruit de verbiage devait être considérablement plus grand que dans le bruit de l'automobile. Dans l'expérience 2, des adultes d'âge moyen de 50 à 65 ans écoutaient des phrases en présence du bruit d'automobile. Même si leur acuité auditive n'était pas assez affectée pour mériter un diagnostique clinique, leur performance était significativement plus faible que celle des jeunes adultes et ce, particulièrement pour les phrases avec un faible niveau de prédictibilité. Donc, même si le bruit d'une automobile affecte moins la perception auditive que le bruit de verbiage, la compréhension de la parole est significativement affectée chez les individus souffrant d'une perte auditive même légère. Les producteurs automobiles devraient poursuivre leurs efforts afin de réduire le niveau de bruit à l'intérieur des voitures et améliorer ainsi l'intelligibilité de la parole. [Financé par Ford Motor Company]

1. INTRODUCTION

In 2009, Canadians amassed about 500,000 passenger kilometers (Stats Can, 2010a[1]). Not surprisingly, almost 70% of that distance was accumulated by adults between the ages of 20-65 years. More and more of the time spent in the automobile is occupied with listening to and understanding speech; either instructions from in-vehicle navigation, traffic information or simple conversations among passengers. Misunderstanding directions and/or difficulty in following conversations can result in attention being pulled from the task of driving and can lead to taking wrong turns, annoyance, or even accidents. Thus, it has become increasingly important to measure the effect of automobile noise on speech intelligibility.

---

[1] Statistics Canada defines passenger-kilometres as the sum of the distances traveled by individual passengers (the driver being considered as one of the passengers).

1.1 Speech intelligibility measurements in automobile noise

Automobile engineers are understandably very interested in predicting how speech intelligibility is affected by automobile noise (Farina, Bozzoli and Strasser, 2003). A number of reports have measured intelligibility by calculating expected speech intelligibility in a background of automobile noise using already-existing metrics including the Speech Transmission Index (STI) and the Speech Intelligibility Index (SII).

The STI (Steeneken and Houtgast, 1980) models the reduction in intelligibility in noise as a result of a decrease in the intensity of modulation found in the speech signal caused by the noise mixing with the signal. It takes into consideration reverberation, separation of signal and noise and distance of the signals from the receiver. Thus, things like where the noise is coming from and the interior textures

of an automobile are important in calculating this measure. In contrast, the SII uses the spectrum of the speech and noise, as well as information about the listener's hearing threshold to predict intelligibility. Each signal is broken down into a number of bands (up to 20) and each band is weighted in terms of its importance. The importance functions vary depending on the content of the speech and the listener's hearing acuity. Thus, the power spectra of the noise and speech are large contributors to this measure. Both measures are highly correlated with intelligibility of speech under many listening conditions (see Steeneken and Houtgast, 1980 for the STI; ANSI, 1997 for SII).

A recent series of studies has evaluated the effectiveness of these measures of speech intelligibility in different driving environments (Samardzic and Novak, 2011a, 2011b). Samardzic and Novak (2011a) used the STI to measure the effect of different road surfaces, as well as the talker and listener position in the car. They found that the overall sound level did not always predict speech intelligibility. Of most interest here, the frequency content of the background noise, not simply the absolute level of the noise, was an important factor in calculated intelligibility.

In a further study, Samardzic and Novak (2011b) used the SII to generate predictions of speech understanding in an automobile for individuals with typical configurations of age-related hearing loss. The index predicted poor intelligibility for all conditions for the hearing impaired compared to normal hearing listeners. However, for the normal hearing conditions, the SII was lower (worse) for the smooth road condition (which resulted in a great deal of high frequency components in the noise) than the rough road conditions (which resulted in more low frequency energy)–even though the rough road had had higher overall sound pressure levels. The opposite was true for the predictions for hearing-impaired listeners. This was because the SII gives more weight to the important frequencies in speech, which, in turn, is cancelled out by the loss of acuity at those higher frequencies for the hearing impaired. Essentially, the noise is masking frequencies that are already inaudible to the hearing-impaired listeners.

Although these studies show that existing speech intelligibility metrics predict that speech should be harder to understand in an automobile than in quiet, there are few reports directly measuring intelligibility in automobile noise using real human listeners. Although the metrics are quite good, their predictions should be tested with real human listeners.

1.2 Speech Intelligibility in Noise

It has long been known that speech reception performance in noise cannot be predicted from either pure tone thresholds or speech understanding in quiet (see, e.g., Beattie, Barr, and Roup, 1997). In addition, even mild amounts of hearing impairment, as is common with increasing age, magnify this difficulty (Dubno, Dirks and Morgan, 1984). One of the biggest complaints of older adults is that hearing in noisy situations like restaurants and

automobiles is difficult, even though they often report little difficulty in quiet conditions.

In order to measure speech intelligibility in noise, Kalikow, Stevens, and Elliott (1977; later revised by Bilger, 1994; Bilger, Nuetzel, Rabinowitz and Rzeczkowski 1984) developed the Speech Perception in Noise test (SPIN; later revised and renamed R-SPIN). We chose to use the SPIN test, rather than one of the many other tests of speech in noise because we were interested in how context might interact with the different types of noise. The test was developed as a screening measure to assess speech perception in noise from both a perceptual (bottom-up) and cognitive (top-down) perspective. In this test, individuals listen to sentences and are asked to report the final word in each sentence. The sentences are presented in a background of multi-talker babble at a single speech-to-noise (S/N) ratio. Half of the sentences in each list have clear contextual cues that allow the listener predict the final word (high-predictability; HP) and half do not (low-predictability; LP). Thus, each individual's performance can be measured when only bottom-up or perceptual information is available, as in the low-predictability sentences, and when both bottom-up and top-down information is available, as in the high predictability sentences. Typically there is a large difference in performance for the two types of lists (Humes, Watson, Christensen, Cokely, Halling and Lee, 1994). This difference illustrates the power of context and top-down processing. Although the test was validated using the single S/N level, it can be transformed into a paradigm using multiple S/N ratios without substantially affecting its validity (Wilson, McArdle, Watts and Smith, 2012).

1.3 Age and speech intelligibility

As we grow older, we experience progressively more difficulty in understanding speech, particularly in situations with background noise (CHABA, 1988; Dubno, Dirks and Morgan, 1984; Sperry, Wiley, and Chial, 1997). Speech understanding in noise starts to decline in the fourth decade even before loss of hearing sensitivity becomes clinically significant (Bergman, 1980). Thus, individuals in their forties and fifties are already experiencing some difficulty in noisy environments. Typically, age-related hearing loss begins in the high frequencies and progressively moves downward to affect lower frequencies. Because of this, low-frequency noises become a greater problem as a person ages and hearing loss progresses, in part because the noise starts to mask the frequencies that do remain audible. Since the noise made by an internal combustion engine is generally loudest in the lower frequencies, it is important to determine whether automobile noise particularly affects speech understanding in individuals with mild hearing losses.

It is well established that, among different types of background noise, meaningful speech noise causes the most disruption in speech intelligibility for both normal hearing and hearing-impaired individuals (Sperry et al., 1997). However, the differential effect of automobile noises on low and high predictable context sentences (reflecting

differential contributions of bottom-up and top-down processing) is unknown. Driving is an effortful process and so may take away from cognitive resources that are necessary to use context in understanding speech. Even small amounts of increased effort can have measurable effects on comprehension (Stanley, Tun, Brownell, & Wingfield, in press). In addition, most of the speech heard in the automobile is contextually appropriate; thus allowing us to more comfortably generalize our results to real-world situations.

Thus, the current project measured performance on the Speech Perception in Noise (Revised; R-SPIN; Bilger, et al., 1984) test with the original background noise of multi-talker speech babble and compared it to performance on the same test in a background of automobile noise. In a further experiment, younger and middle-aged drivers were compared to see how age and age-related hearing deficits affect speech intelligibility in the automobile.

2. EXPERIMENT 1

The purpose of Experiment 1 was to compare the SPIN test with the usual speech babble noise to a test using the same sentences but presented -in an automobile noise. We recorded the noise of a Ford Motor Company SUV at 80 mph in the passenger seat of the car. We separated the two channels of the SPIN test and replaced the noise channel with the automobile noise. The question is whether the same level of automobile noise is as disruptive as speech babble for speech understanding even though the frequencies of the automobile noise do not overlap with the speech signal to the same extent as the babble noise.

2.1 Method and Materials

2.1.1 Subjects

Thirty-six normal-hearing college-age Purdue University students between the ages of 18 and 22 participated in this experiment in exchange for partial fulfillment of one of their course requirements. Half were randomly assigned to the babble and half to the automobile noise condition. The design was between subjects because, although the R-SPIN has 8 forms, each pair of forms has the same word in either high or low-context. We tried to minimize the amount of priming of the word for each participant. All individuals tested within the normal range (e.g., ≤25 dB HL at octave frequencies from .25 to 8 kHz (inclusive) on a brief pure-tone hearing screening).

2.1.2 Materials and Design

The materials were adapted from the R-SPIN test (Bilger, et al., 1984). This test contains 200 words distributed as the last words in 200 low-predictability (LP) and 200 high-predictability (HP) sentences. The HP sentences give clear contextual cues about the identity of the final word in the sentences whereas the LP sentences do not. An example of a high predictability sentence is, "The watchdog gave a warning growl." An example of a low predictability sentence is, "I had not thought about the growl." Listeners are asked to repeat back the final word in each sentence. Four list pairs, each consisting of two 50-sentence lists, contain the same target word in either a LP or HP sentence. Normally, the R-SPIN is presented with a background of multi-talker speech babble composed of twelve simultaneous voices at a S/N ratio of 8 dB. The validity and reliability of the R-SPIN test have been solidly established (Bilger, et al., 1984; Kalikow, Stevens and Elliott, 1977).

In the present experiment, the R-SPIN test sentences were presented at 70 dB SPL. The background noise was either the original babble or automobile noise recorded in an SUV moving on a road at 80 miles per hour. Automobile noise contains much higher levels of low frequency (< 200 Hz) noise than speech babble. The noise is primarily caused by road-tire interaction and at higher speeds, such as 80 mph, wind noise can also be a problem. Sound transmission into the passenger compartment is controlled and there is sound absorption within the car due to headliners, seats, and carpeting. The noise reduction is more effective at higher frequencies resulting in the spectrum shown in Figure 1 (gray line). Several recordings of automobile noise were examined, all had similar spectral characteristics but there were differences in level. Pilot tests showed that in order to obtain about the same levels of performance we needed to adjust the noise level differently for the two types of background noise. The automobile noise level was varied to create S/N ratios of -10, -8, -5, 0, 5, and 10 dB whereas the speech babble was varied to create S/N ratios of 0, 3, 5, 7, 10, and 12 dB. To be consistent with ratios typically reported with the SPIN test, the S/N ratios were calculated from the unweighted sound pressure levels of the signal (the final word in each sentence) and the background noise (speech babble or automobile noise):

$$SNR = 10 \log_{10} \left[\frac{\overline{p_{signal}^2}}{\overline{p_{noise}^2}} \right],$$

where the overbar represents a time average of the squared pressure. All eight versions with 50 sentences each of the R-SPIN test were used in this study. Sentences were presented monaurally in the right ear through headphones in a sound-isolated acoustic chamber.

The average power spectrum of the babble noise and the automobile noise is presented in Figure 1; the un-weighted sound pressure level of both noises is 72.4 dB. This level was chosen for illustration because it was the overall noise level required to meet, on average, a S/N ratio of 0 dB on the last word in the sentences. The playback system was calibrated so that the average level across the entire sentences was 70 dB. As is evident in Figure 1 the majority of the energy in the automobile noise is at low frequencies (below 100 Hz) and in the speech range (200 Hz to 2000 Hz), the babble noise spectrum is always above the automobile noise. Because of people's lower sensitivity to

noise at low frequencies, the automobile noise is perceived as being quieter than the babble noise even though the unweighted sound pressure levels are the same.
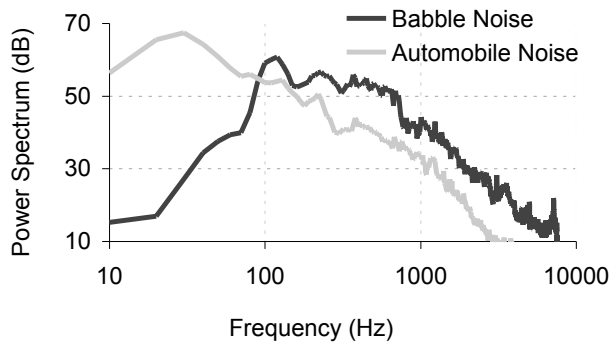

Figure 1. Power spectra of the babble noise (black line), and the automobile noise (gray line). The unweighted sound pressure level for both signals is 72.4 dB.

2.1.3 Procedure

Each subject listened to eight versions of the R-SPIN test (one of the sets of 50 sentences at each S/N ratio) and recorded the last word of each sentence on a form provided for them. The eight lists and six S/N ratios were counterbalanced across subjects in an incomplete Latin square design. No subjects heard the same sentences more than once.

2.2 Results

The percent correct score for the high- and low-predictability sentences with babble noise compared with automobile noise is shown in Figure 2. Each point represents the average score for 18 participants. Error bars indicate standard error of the mean. The small size of the error bars reflects the homogenous performance of the younger adults.

First, the functions look very regular with fairly linear increases in percent correct as a function of S/N ratio in areas of the curve off the floor or ceiling. Second, as expected, there was a substantial difference between the high and low predictability sentences at every S/N level with the low predictability sentences averaging 57% correct (collapsed across conditions) and the high predictability sentences averaging 82% correct. Stated in terms of the S/N ratio, there about a 4-6 dB difference increase in S/N ratio was needed to achieve 50% correct. This is about what has been shown in the past (Pichora-Fuller, 2008). Looking at the differential effect of context for the automobile noise and babble in (for example, at 60, 70 and 80% correct), we see about a 5 dB effect in the babble and the automobile noise for those conditions in which performance is not on the ceiling or the floor.

Finally, it is clear that the babble noise had a greater effect on overall performance than the automobile noise. Comparing performance at 0 dB S/N, there was a difference

of 53% for the low-predictability and 44% for the high predictability sentences between the babble noise and the automobile noise. . Thus, at equal unweighted intensities, the babble was having a much greater effect on performance than the automobile noise. Looking at it another way, the S/N ratio needed in order to reach 50% correct is about -8 dB for the automobile noise and about +3 for the speech babble (collapsing across predictability).
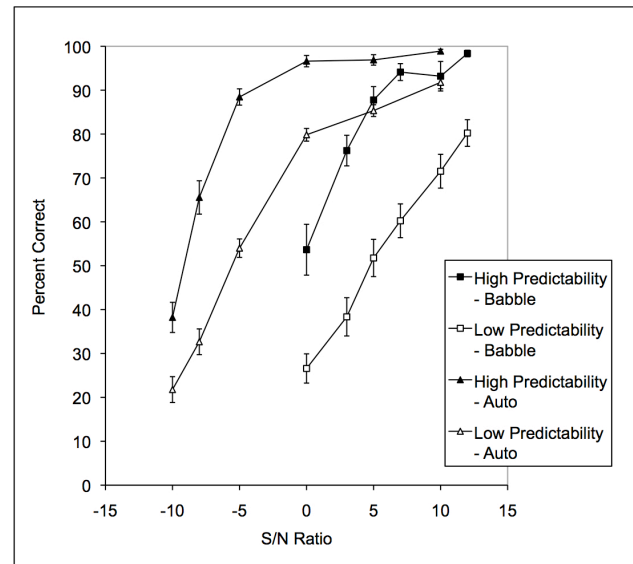

Figure 2. Percent correct word identification as a function of S/N ratio in dB for high and low predictability SPIN sentences in speech babble and automobile noises. Error bars are standard error of the mean.

As mentioned above, in order to achieve the same level of intelligibility as the automobile noise, the S/N ratio based on unweighted sound pressure level needed to be substantially higher for the speech babble. However, we re-calculated the same data in terms of a S/N ratio based on an estimate of loudness (Zwicker and Fastl, 2007). Loudness was calculated over the whole sentences by using the Zwicker time-varying loudness algorithm in the Brüel and Kjaer Sound Quality Software Module. This is based on the German DIN 45631 (2010). The mean value of the loudness over the duration of the last word in the sentence is used in the S/N (Loudness) calculation. Sample average loudness spectra for the three types of signals all normalized to 20 sones are shown in Figure 3. Note that in the he speech range (200 to 2000-3000 Hz) which corresponds to 2.5 to 14-16 Bark, the babble and the speech signal are substantially louder than the automobile noise.
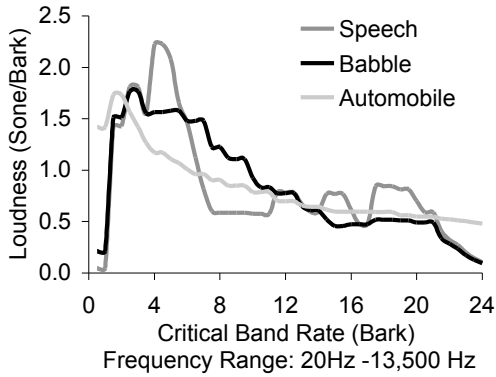
Figure 3. Samples of predicted loudness spectra for speech (dark gray line), babble (black line) and automobile noise (light gray). All of these sounds have been normalized to have a total loudness of 20 sones.

The results of the reanalysis with percent correct as a function of Zwicker loudness is represented in Figure 4. When the data are plotted this way, they demonstrate an interaction of perceived loudness and context (see Figure 4). At equal perceived loudness, the noise has the same effect regardless of the composition of the background noise, provided the context is predictable. However, when context is not present, the physical similarity of the babble to the signal makes the signal more difficult to understand. Thus when top-down information is available, the differences in the physical masking of the stimulus are not relevant—the automobile noise is as detrimental as the babble. However, when the listener must rely solely on the bottom-up perceptual information, the overlap in frequency range of the signal and noise becomes more important.
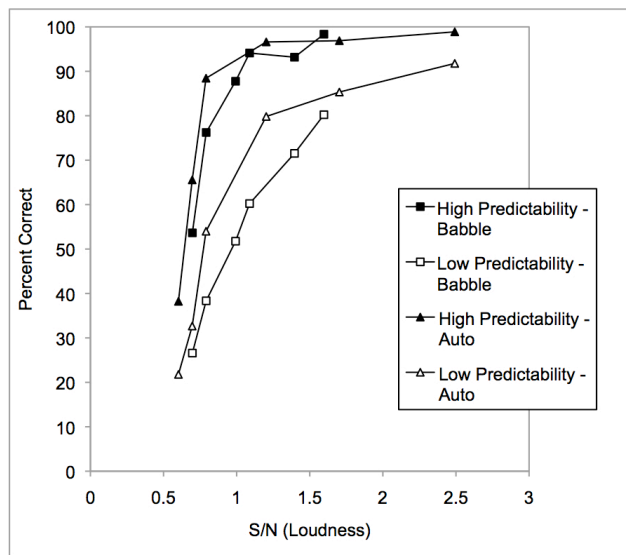

Figure 4. Percent correct as a function of S/N (loudness) for high and low predictability SPIN sentences in speech babble or automobile noise. S/N (loudness) is the ratio of the average of the estimated loudness (Zwicker and Fastl, 1997) over the last words of the sentences to the average estimated loudness of the noise that was playing at that time.

Note also, even though the low frequency components are attenuated in the loudness calculations (which reflects characteristics of the human hearing system), there are still some contributions to the overall loudness from the low frequency critical bands and these contributions are more prominent in the automobile noise. We also looked at other measures of noise level that either attenuate or do not use the low frequency energy, e.g., A-weighted sound pressure level or just the levels in the speech bands. However, use of Zwicker's model to predict loudness and calculate S/N (loudness) yielded the most consistent results when comparing the effects of babble and automobile noise. This indicates that more accurate models of loudness, which include frequency- and level-dependent weighting and masking effects, should be used when examining the effects of noises that have spectral energy distributions that differ to each other and from those of speech signals. For tests with speech and babble noise only, use of unweighted sound pressure levels to determine S/N ratios is appropriate because the signals have very similar spectral shapes, though use of loudness should be considered when the levels of the sentence and the babble noise are very different.

The intelligibility results for the two background noises when plotted against a loudness S/N ratio (Figure 4), rather than over the speech frequency bands (Figure 2) is interesting and requires further investigation.

3. EXPERIMENT 2

Now that we have some sense of how different the automobile noise is from the babble, we can determine how age and mild hearing loss change the effects of automobile noise on understanding speech. In this study, we replicated the automobile noise conditions of Experiment 1 with middle-aged individuals ranging in age from 50-65 years old. Typically, individuals in these age ranges have mild age-related losses that are not severe enough to warrant remediation. However, as mentioned above, there is a substantial amount of data showing that even small amounts of hearing loss affect speech understanding, especially in noise (Dubno et al., 1984; Surprenant, 2007). Given that these individuals make up about 30% of the hours spent in an automobile (Statistics Canada, 2010), it is important to confirm that this holds true in automobile noise, rather than the usual speech babble noise.

3.1 Methods and Materials

3.1.1 Subjects

Eighteen individuals ranging in age from 50-65 years old (13 female; mean: 56.28 years) from the Purdue University community volunteered to participate in exchange for a small honorarium. None of the participants reported taking medication that affected cognitive functioning.

3.1.2 Materials and Design

The materials and design were identical to the automobile noise condition described in Experiment 1.

3.1.3 Procedure

Subjects were first given a brief pure-tone hearing screening at 250, 500, 1000, 2000, 4000, and 8000 Hz. The rest of the procedure was identical to that of the automobile noise condition described in Experiment 1.

3.2 Results

The mean audiometric thresholds are shown in Figure 5. Although none of the participants would qualify as clinically hearing impaired, they all showed some deviation from normal hearing, particularly at the higher frequencies, as is typically found in older adults (Bergman, 1980).
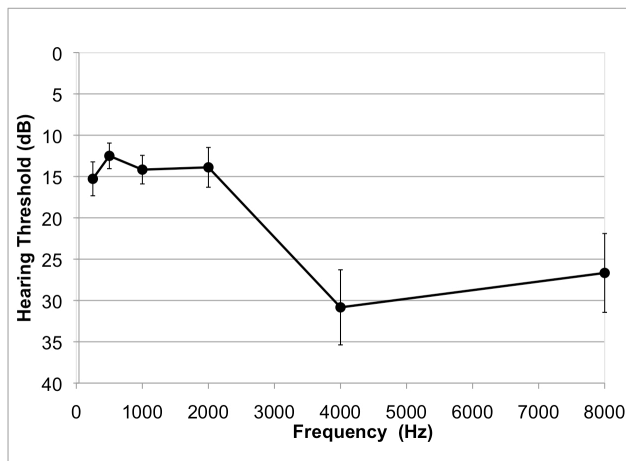

Figure 5. Average hearing threshold (dB HL) as a function of frequency for individuals in Experiment 2. Error bars are standard error of the mean.

The performance-intensity functions of high- and low predictability sentences with automobile noise as the background are illustrated in Figure 6 The data from the younger subjects in Experiment 1 are re-presented for comparison. Each point represents the average score for eighteen participants. Error bars indicate standard error of the mean.

As can be seen in Figure 6, performance for the older group is worse than the younger group, even though their hearing loss is minimal. Because the conditions were run between subjects we can consider them to be different conditions in the same experiment and perform statistical tests (ANOVA) to verify the visual inspection. This observation was confirmed by a 2 (young or middle-aged group) x 2 (predictability) x 6 (S/N level) mixed ANOVA that was performed on the data. There was a main effect of group ($F(1,34)$=1613, mean squared error (MSe)=0.71) with the younger (M=0.71) outperforming the older (M=0.63) group.

There were also main effects of predictability ($F(1,34)$=510, MSe=.01), and S/N level ($F(5,170)$=247,

MSe=5.2). There was no interaction of predictability by group ($F(1,34\_$=3.68, MSe=0.037) and no interaction of level by group ($F(5,170)$=1.09, MSe=0.02) However, there was a significant three-way interaction ($F(5,170)$=4.26, MSe=0.42). The three way interaction is due to the finding that the difference between the older and younger group is larger in the low predictability condition than in the high predictability condition but only for the middle (-5, 0, 5) S/N ratios. As mentioned above, performance on the low predictability sentences is considered to be influenced more heavily by bottom-up perceptual information. Thus, even though their hearing would be considered to be clinically normal, they were still more affected by the noise than the younger group. However, they were able to use context to make up for some of that difficulty in the HP condition.
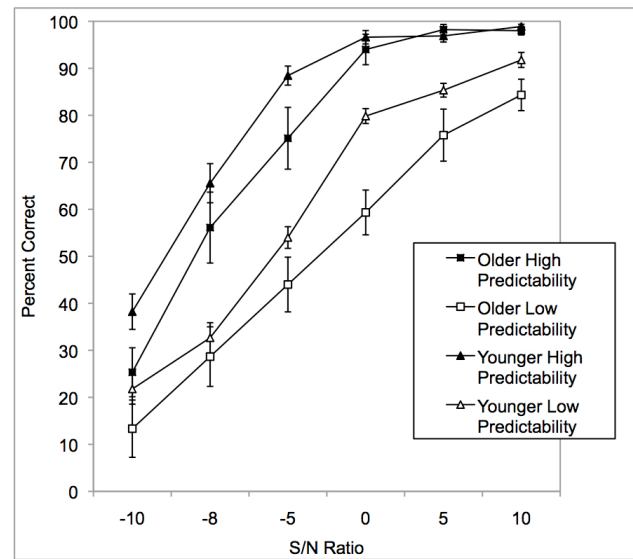

Figure 6. Percent correct word identification as a function of S/N ratio in dB for high and low predictability SPIN sentences in older and younger adults. Error bars are standard error of the mean.

4. DISCUSSION

The data reported here showed that, although speech babble has a greater effect on speech understanding than automobile noise at the same unweighted sound pressure level, the context effect is about the same, indicating that the use of context in a babble condition is no better or worse than it is in a background of automobile noise. When examined in terms of Zwicker loudness rather than unweighted signal to noise ratio, intelligibility is identical in the two noise conditions with predictable context but in the unpredictable context the speech babble continues to be more detrimental than the automobile noise. In Experiment 2, we showed that even a small amount of hearing loss has an impact on speech recognition in noise, particularly when there is little or no top-down or contextual information to support perception.

Recall that Samardzic and Novak (2011b) showed that the SII predicted less of an effect of some road noises for individuals with high frequency hearing loss because the

energetic masking of the high frequencies in that noise is having an effect on frequencies that are inaudible to them anyway. The noise used here had more of a low-frequency component to it. Thus, it would be interesting to see whether changes in the automobile noise to include more high frequency components would have as much of an effect on individuals with hearing loss than on those without. The SII would predict that it would not have an effect.

It should also be noted that our participants were merely listening to the stimuli over headphones rather than engaging in a dual task like driving and listening. Samardzic, Novak, and Gaspar (2012) showed that adding the driving task (in a simulator) required an increase of the signal of 3 dB for equivalent performance. Thus, it would be very interesting to see whether the difference between the low and high predictability sentences changes as more higher-level resources are being occupied by the second task of driving. We do know that listening takes away from driving (e.g., Horrey and Wickens, 2006); how much does driving take away from listening?

The current project showed that, although automobile noise is less damaging than speech babble at typical S/N ratios, speech understanding for individuals with even small amounts of hearing loss is significantly impacted by the noise. Automobile makers therefore should continue their efforts to reduce the noise levels in cars in order to increase speech intelligibility.

ACKNOWLEDGEMENTS

REFERENCES

American National Standards Institute (1997). "Methods for calculation of the Speech Intelligibility Index." (ANSI S3.5-1997). New York: ANSI.

Beattie, R.C., Barr, T., and Roup. C. (1997). "Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise." Br. J. Audiol. 31, 153–164.

Bergman, M. (1980). Aging and the perception of speech. Baltimore: University Park.

Bilger, R.C. (1994). *Authorized version revised SPIN test (Revised Speech Perception in Noise Test)*." Champaign, IL: The University of Illinois Press.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (1984). "Standardization of a test of speech perception in noise." J. Speech Hear. Res. 27, 32-48.

Committee on Hearing, Bioacoustics, and Biomechanics (CHABA) – Working Group on Speech Understanding and Aging (1988). "Speech understanding and aging." J. Acoust. Soc. Am. 83, 859-895.

DIN 45631 German Institute for Standardization (Deutsches Institut für Normung), Calculation of loudness level and loudness from the sound spectrum – Zwicker Method, Revision/Edition 1991; Amendment 1 March 2010; Calculation of the loudness of time-variant sound.

Dubno, J.R., Dirks. D.D., and Morgan, D.E. (1984). "Effects of age and mild hearing loss on speech recognition in noise." J. Acoust. Soc. Am. 76, 87-96.

Farina, A., Bozzoli , F., and Strasser, P. (2003). "Comparative Study of Speech Intelligibility Inside Cars." EuroNoise03, (2003).

Horrey, W. J. and Wickens, C. D. (2006). "Examining the impact of cell phone conversations on driving using meta-analytic techniques." Human Factors: The Journal of the Human Factors and Ergonomics Society 48, 196-205,

Humes, L. E. and Roberts, L. (1991). "Speech-recognition difficulties of the hearing-impaired elderly: The contributions of audibility." J. Speech Hear. Res. 33, 726-735.

Humes L.E., Watson, B.U., Christensen, L.A., Cokely. C.G., Halling, D.C., and Lee L. (1994). "Factors associated with individual differences in clinical measures of speech recognition among the elderly." J. Speech Hear. Res. 37, 465–474.

Kalikow, DN, Stevens KN, and Elliott LL. (1977). "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability." J. Acoust. Soc. Am. 61, 1337–1351.

Pichora-Fuller, M. (2008). "Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing." Int. J. Aud., 47(S2), S72-82.

Samardzic, N., and Novak, C. (2011a). "In-vehicle speech intelligibility for different driving conditions using the Speech Transmission Index." Noise Control Eng. J. 59, 397-407.

Samardzic, N. and Novak, C., (2011b). "In-vehicle speech intelligibility for the hearing impaired using speech intelligibility index." SAE Technical Paper 2011-01-1681.

Samardzic, N., Novak, C. and Gaspar, R. (2012) "The evaluation of speech intelligibility in a simulated driving environment using the hearing in noise test (HINT)." Int. J. Vehicle Noise and Vibration, 8, 318-336.

Sperry, J.L., Wiley, T.L., and Chial, M.R. (1997). "Word recognition performance in various background competitors." J. Amer. Acad. Audiology 8, 71-80.

Stanley, R., Tun, P.A., Brownell, H., and Wingfield, A. (In press). "Hidden costs of effortful listening on speech comprehension." In F. Columbus (Ed.). Speech processing. Hauppauge, NY: Nova Science Publishers.

Statistics Canada (2010). "Canadian Vehicle Survey: Annual." Catalogue no. 53-223-X, 2009. Document downloaded on December 1, 2012 from http://www.statcan.gc.ca/pub/53-223-x/53-223-x2009000-eng.pdf.

Steeneken, H.J.M. and Houtgast, T. (1980) ''A physical method for measuring speech-transmission quality'', J. Acoust. Soc. Am., 67, 318 – 326.

Surprenant, A. M. (2007). "Effects of noise on identification and serial recall of nonsense syllables in older and younger adults." Aging, Neuropsychology, and Cognition, 14, 126-143.

Wilson, R. H., McArdle, R., Watts, K. L., and Smith, S. L. (2012). "The Revised Speech Perception in Noise Test (R-SPIN) in a multiple signal-to-noise ratio paradigm." J. Amer. Acad. Aud 23, 590-605.

Zwicker, E. and Fastl, H. (2007) *Psychoacoustics: Facts and Models,* 3rd Edition, Springer-Verlag, Berlin Heidelberg.