Recall of spoken word lists in English and native Swedish presented at different signal-to-noise ratios and different reverberation times – A comparison between children aged 10-11 years and College students

S. Hygge[1], A. Nöstl[1], A. Hurtig[1], A. Haga[1], M. Holmgren[1]

[1] Environmental Psychology - Faculty of Engineering and Sustainable Development, University of Gävle, Gävle, SWEDEN. Staffan.Hygge@hig.se

ABSTRACT

Two experiments will be presented which assessed free recall of spoken words in Swedish (native tongue) and in English heard under different signal-to-noise ratios (SNR: +3 and +12 dB), and different reverberation times (RT: 0.3 and 1.2 s). All participants encountered these eight experimental conditions (Language*SNR*RT). The first experiment was run with college student (*N*=48), who were run individually. In the second experiment children in grade 4 (10-11 years, *N*=72) took part and they were run as a group in their regular classrooms.

Twelve wordlists in English and twelve wordlists in Swedish were generated. The words were chosen according to their ranks in category norms for the two languages. The number of words in each list was 12 for the college group and 8 for children in Grade 4. The 2 x 12 wordlists were presented in counter balanced presentation orders in three blocks (Blocks). To compare primacy and recency effects the word lists were divided into three parts (p3rd). After each wordlist the participants typed in or wrote down the words they could recall.

The basic hypotheses for the recall of the words were that working memory would be overloaded when the SNR was low and the RT was long, and that SNR and RT would interact with each other, with Language and with Study (Grade4/College). The analyses suggest that for both groups there were expected effects of language and of SNR, but the effect of RT was smaller and only showed up in interactions.

INTRODUCTION

The acoustical conditions in classrooms may severely impair listening, which in turn impairs learning. To safe-guard against inferior listening conditions in the classroom, government agencies and professional societies have established building codes, standards and recommendations for acceptable signal-to-noise ratios (SNR) and reverberation times (RT) in classrooms and other work places where it is important to apprehend auditory information (American National Standards Institute 2002; Swedish Standards Institute 2007; Swedish Work Environment Authority 2006, 2011; Vallet & Karabiber 2002). Such codes and standards have been founded on conditions required for speech intelligibility and correct identification of spoken words or isolated sentences.

There is, however more to learning than just listening. Correct identification of what was said is a necessary condition for memory and learning, but it is not a sufficient one. There is a gap between speech intelligibility and memory of the speech, and the size of that gap is basically a function of to which extent low intelligibility of the

spoken message has exhausted the limited working memory resources and left little, if any, left for elaboration, recoding, storing, and subsequent recall. Thus, it must be easy to identify what was said. Otherwise subsequent recall will suffer, even when speech intelligibility is at an acceptable level (Kjellberg 2004; Kjellberg et al. 2008; Ljung et al. 2013; Ljung & Kjellberg 2009; Ljung et al. 2009).

The emerging evidence for a gap between speech intelligibility and the memory of what was said, is an argument for that the building codes and recommendations for acoustics in classrooms should be evidence-based on how much is remembered, and learnt, and not only on speech intelligibility.

To learn more about memory of spoken words when presented under suboptimal listening conditions we designed the present experiment. The following variables were set up as within person factors: (1) Language (Swedish native tongue or English, (2) SNR (+3 or +12 dB), (3) RT (=1.2, 0.3 s), (4) the early, middle or last part of the word list (p3rd), and (5) the three presentation blocks (Blocks). Study (Grade4 vs College) was included as a between person factors

The basic hypotheses for the recall of the words were that working memory would be overloaded and impair recall of the words when the language was English, when the SNR was low, when the RT was long and more so in Grade4 than for the College students.

METHOD

There were 48 persons in the College student, with equal numbers of males and females. In grade 4 (10-11 years) a total of 72 children took part, slightly more girls than boys.

Twenty-four wordlists were presented, twelve lists in English and twelve in Swedish. For the children the lists had 8 words, for the college students 12 words. The words were taken from 24 different categories and chosen according to their rank orders in category norms for the two languages.

The lists were recorded by a female speaker in a sound-attenuated chamber and normalized to 66 dB(A). The words in the lists were read with approximately 3 s interval between each word. Broadband noise was added to the word lists to create the SNR-conditions of +12 and + 3 dB. RT (0.3 and 1.2 s) was manipulated with acoustical software. The lists were played back at a sound level of ~ 66 dB(A) in the listening positions in the children's classroom and at a comfortable level in the earphones for the college students.

At the end of each wordlist the College participants were asked to type the words they could recall and the children wrote them down on paper.

RESULTS

Table 1 shows the results for the significant main effects and interactions up to 3-way interactions. (There were also significant 4-6-way interactions but the space restrictions for the present conference papers demand simplifications). The decimals in the degrees of freedom for the *F*-tests indicate that a Greenhouse-Geisser correction was made because of violations of the sphericity assumption.

Table 1: Significant variance sources for all effects up to 3-way interactions.

Study	$F(1, 118) = 99.30$	***
Lang	$F(1, 118) = 295.79$	***
Lang*Study	$F(1, 118) = 67.24$	***
SNR	$F(1, 118) = 318.58$	***
SNR*Study	$F(1, 118) = 11.55$	**
p3rd	$F(1.56, 184.3) = 81.95$	***
p3rd*Study	$F(1.57, 184.3) = 19.15$	***
Block*Lang	$F(1.98, 233.8) = 5.87$	**
Block*p3rd	$F(3.72, 438.9) = 5.46$	***
Block*p3rd*Study	$F(3.72, 438.9) = 3.01$	*
Lang*SNR*Study	$F(1, 118) = 6.60$	*
Lang*p3rd	$F(1.97, 232.7) = 41.32$	***
Lang*p3rd*Study	$F(1.97, 232.7) = 3.28$	*
RT*SNR	$F(1, 118) = 8.74$	**
RT*SNR*Study	$F(1, 118) = 2.92$	(*)
SNR*p3rd	$F(1.96, 231.2) = 21.62$	***
Block*Lang*p3rd	$F(3.57, 421.3) = 10.17$	***
Lang*SNR*p3rd	$F(1.99, 234.4) = 9.81$	***

Note: (*) = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

As can be seen there were strong and reliable main effects of Study, Language, SNR and p3rd. There were no significant main effects of RT and Block, but they showed up in interactions.

Four pairs of nested interactions will be looked into in more detail.

1. RT*SNR, RT*SNR*Study (trend)
2. Lang*Study, Lang*SNR*Study
3. Block*p3rd, Block*p3rd*Study
4. SNR*p3rd, Lang*SNR* p3rd

The first three pairs have specifications by Study, and they say something about the differences between Grade4 and College. The fourth has relevance for understanding how SNR and the different parts of the word lists are selectively affected by Language. All these four pairs of nested interactions are shown in Figures 1-4.

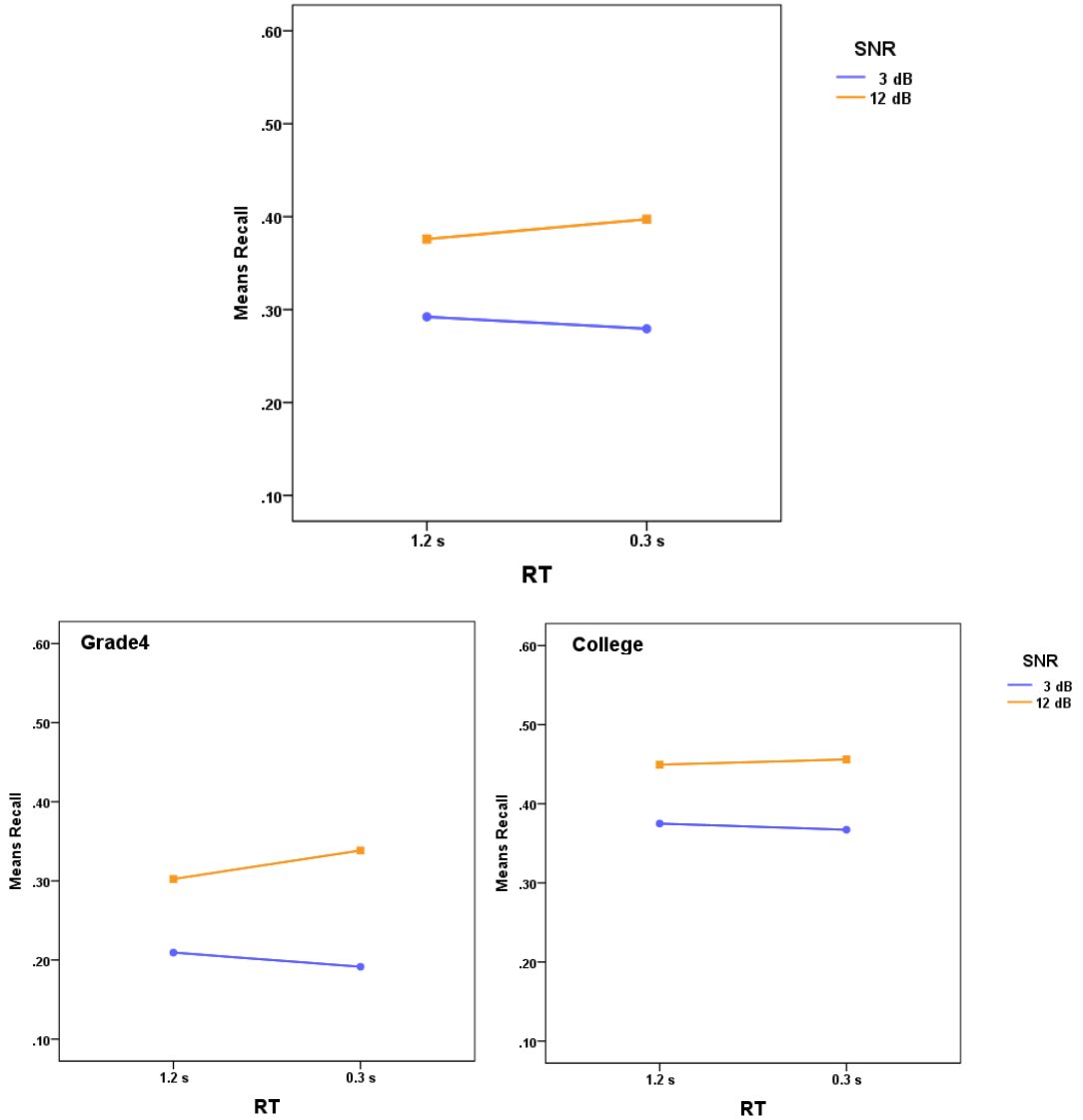

Figure1: Means recall RT*SNR (upper panel), and RT*SNR*Study (trend., lower panel)

There is a significant interaction between SNR and RT indicating that the net effect of improvement between SNR 12 and 3 dB is larger when the RT = 0.3 s. The interaction RT*SNR*Study is of borderline significance but indicates that this net effect of improvement between SNR 12 and 3 dB at RT = 0.3 s is true only for Grade4, but not for College students.

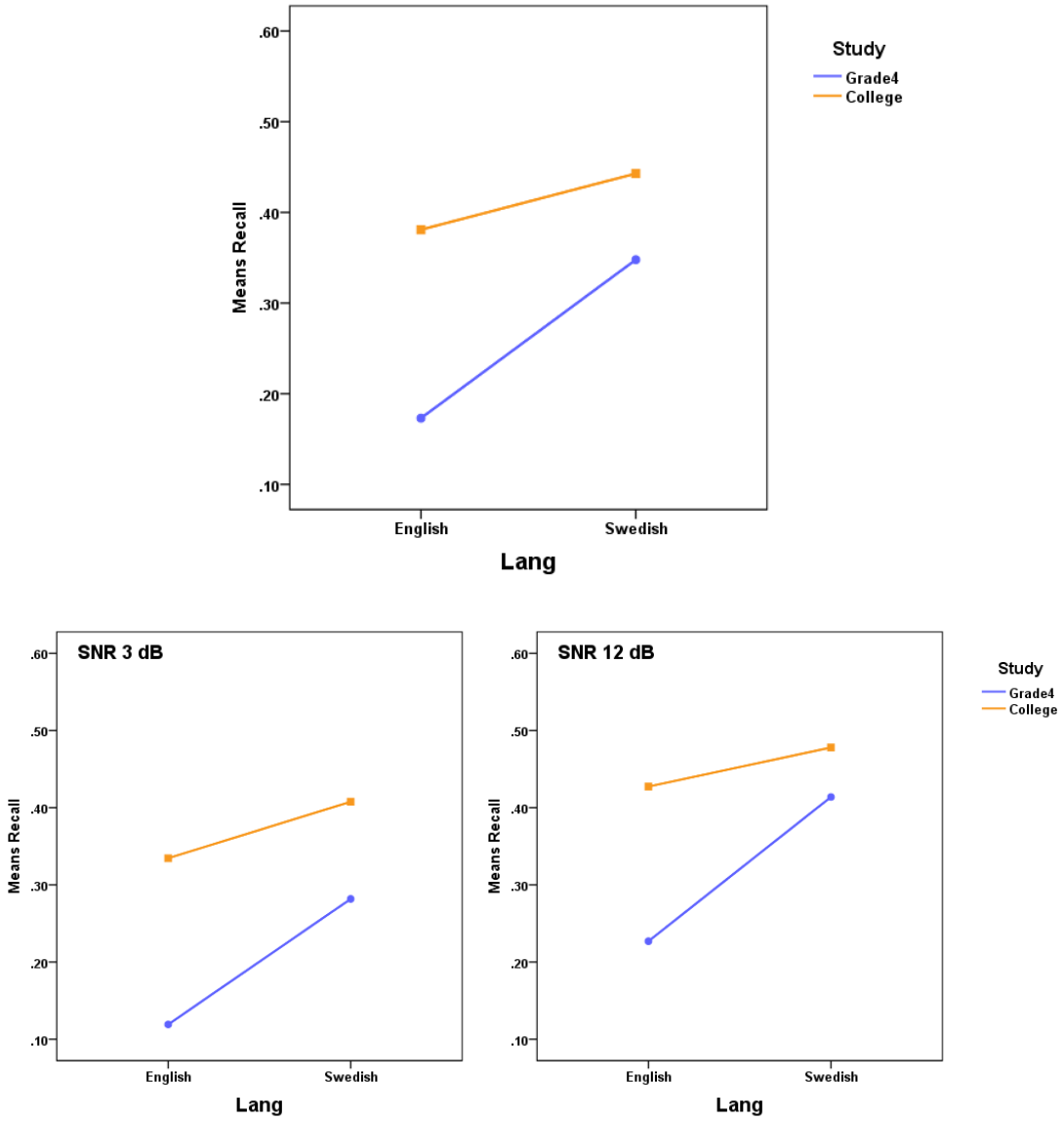

Figure 2: Means recall by Lang*Study (upper panel), and Lang*SNR*Study (lower panel)

As expected the College students perform much better than Grade4, particularly on the English words. Specifying this interaction for the two SNR conditions separately shows that the net advantage in recall for the College students compared to Grade4 is smallest for the Swedish words at SNR 12 dB. For SNR = 3 dB and English there is not a difference between Grade4 and the College group.

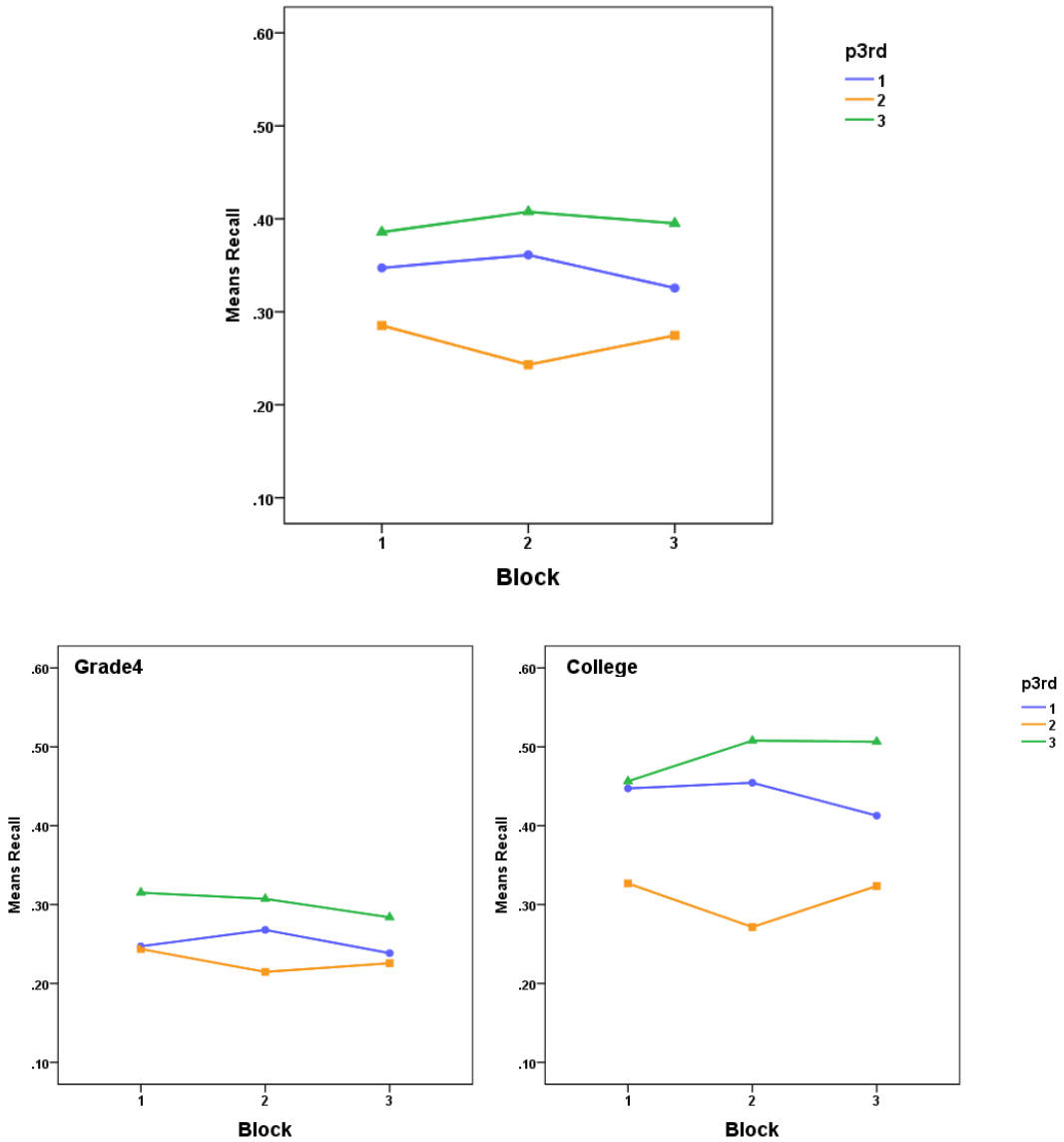

Figure 3: Means recall Block*p3rd (upper panel) and Block*p3rd*Study (lower panel)

Both Grade4 and the College group have their worst performance in the middle part of the wordlist in Block2. For the 3rd part of the word lists there is a decrease with Blocks in performance for Grade4 but an improvement form Block 1 for the College students.

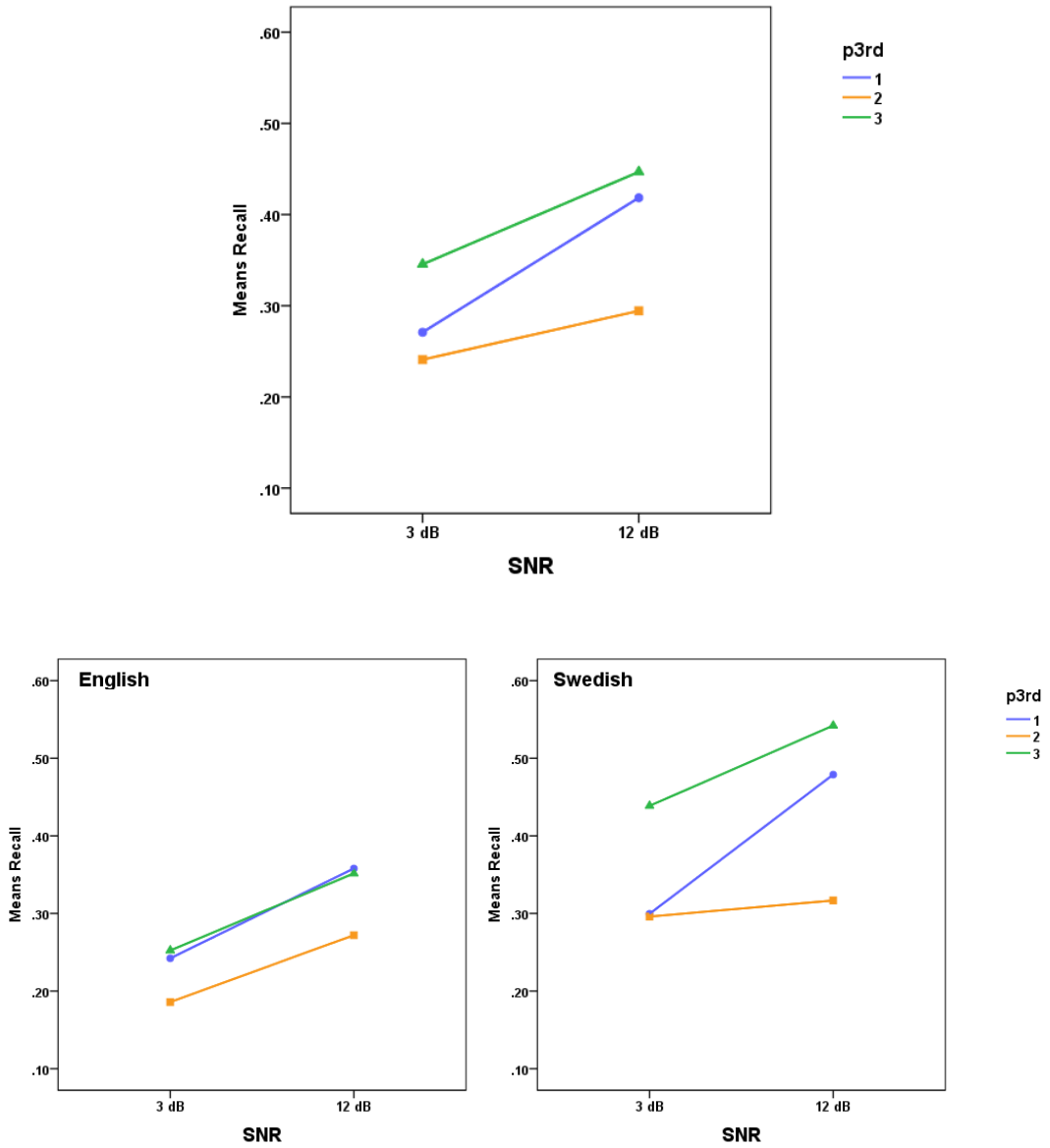

Figure 4: Means recall SNR*p3rd (upper panel) and Lang*SNR*p3rd (lower panel)

The advantage of the SNR difference between 12 and 3 dB is most marked for the first part of the word lists. The interaction with Language however restricts that effect to Swedish wordlists only.

It is also interesting to note that the smallest difference between the memory performance for the College students and Grade4 is for the Swedish words in Block2 in the middle of the list (p3rd = 2), where the mean is the same (0.32). This says something interesting about equal base performance for the two groups when the interference between long-term memory and short term memory is at its maximum.

CONCLUSIONS

There were strong differences between Grade4 and College students in recall of the word lists. There were also marked main effects of SNR, Language and the three parts of the word list (p3rd). However, there was no strong effect of RT. It showed up

only in interactions with other variables, indicating that the net effect of improvement between SNR 12 and 3 dB is larger for RT = 0.3 s than for RT = 1.2 s., and that this net effect may be restricted to Grade4 only.

This lack of a strong effect from RT is probably confounded by the fact that the words in the word list were presented as single words, not as a sentence where words directly follow each other. Thus, when there is a distinctive pause after the word, the effect of a prolonged RT may well be that a long RT does not matter much. But it matters more when there is no such pause.

ACKNOWLEDGEMENTS

The research reported here was supported by The Swedish Research Council Formas.

REFERENCES

American National Standards Institute (2002). ANSI S12.60-2002. Acoustical performance criteria, design requirements, and guidelines for schools: Standards Secretariat Acoustical Society of America.

Kjellberg A (2004). Effects of reverberation time on the cognitive load in speech communication: Theoretical considerations. Noise&Health 7: 11-21.

Kjellberg A, Ljung R, Hallman D (2008). Recall of words heard in noise. Applied Cognitive Psychology 22: 1088-1098.

Ljung R, Israelsson K, Hygge S (2013). Speech intelligibility and recall of spoken material heard at different signal-to-noise ratios and the role played by working memory capacity. Applied Cognitive Psychology 27: 198-203.

Ljung R, Kjellberg A (2009). Long reverberation time decreases recall of spoken information. Building Acoustics 16: 301-312.

Ljung R, Sörqvist P, Kjellberg A, et al. (2009). Poor listening conditions impair memory for intelligible lectures: Implications for acoustic classroom standards. Building Acoustics 16: 257-265.

Swedish Standards Institute (2007). SS 25268:2007 Acoustics - Sound classification of spaces in buildings - Institutional premises, rooms for education, preschools and leisure-time centers, rooms for office work and hotels [In Swedish]. Swedish Standards Institute.

Swedish Work Environment Authority (2006). Rapport – Verksgemensamt projekt – Buller i skolmiljö 2003 [Report – Interdepartmental project – Noise in the school environment 2003]. CTO 2003/482

Swedish Work Environment Authority (2011). Kunskapsöversikt - Yrkesrelaterade röststörningar och röstergonomi [Review – Work related voice problems and voice ergonomics]. Report 2011:6.

Vallet M, Karabiber Z (2002). Some European policies regarding acoustical comfort in educational buildings. Noise Control Engineering Journal 50: 58-62.