12th ICBEN Congress on
Noise as a Public Health Problem

When and why does a long reverberation time improve comprehension and recall?

Staffan Hygge[1]

[1] University of Gävle, Gävle, Sweden

Corresponding author's e-mail address: Staffan.Hygge@hig.se

ABSTRACT

In four recent experiments we have seen that a long reverberation time (RvT) may improve, rather than impair, comprehension and recall of spoken words or texts for participants who have limited language skills.

A long RvT improved, rather than impaired, comprehension for Swedish pupils with a low proficiency in English reading when taking a grade 9 English listening comprehension test in their classroom. For those who were good at reading English there was a better recall with a short RvT. This crossover antagonistic interaction was replicated with Swedish college students grouped by their English proficiency reading skills.

In two word list experiments with Swedish pupils in grade 4 and college students, English and Swedish words were presented with a long and short RvT and crossed with two signal-to-noise ratios. Also here there were indicators of a crossover interaction to the effects that a long reverberation time improved, rather than impaired, the recall of the words for students that were on the low side of English language proficiency.

Possible explanations will be discussed in the presentation.

BACKGROUND AND ISSUES

The acoustical conditions in classrooms may severely impair listening, which in turn impairs learning. To safe-guard against inferior listening conditions in the classroom, recommendations and standards have been imposed for acceptable background levels (BL), signal-to-noise ratios (SNR) and reverberation times (RvT) in classrooms and other work places where it is important to apprehend auditory information. These recommendations are based on conditions required for good speech intelligibility (SI), i.e. correct identification of spoken words or isolated sentences. WHO [1] stated that a RvT below 0.6 s is desirable for adequate speech intelligibility, and that an optimal RvT for speech intelligibility is 0.3-0.5 s for vulnerable groups. The Swedish Work Environment Authority [2] stated that in premises for children, the elderly and the hearing impaired a RvT of less than 0.5 s may be needed. The American National Standards Institute [3] also has published similar recommendations.

The BL and the SNR in the classroom are also crucial for what is understood and subsequently remembered of the spoken information. Recommendations about BL are as a rule in the region of 30-35 dBA [1, 2, 3]. For SNRs in a classroom the recommendation are that SNR should be at least 15 dB to yield a good speech intelligibility [1, 2, 3]. Applying a SNR criterion for an ordinary classroom where the teacher is standing at the front of the classroom and speaking at 66 dBA, the sound level of the speech will decrease to ≈ 52 dBA at a distance of ≈ 6 m from the speaker out into the classroom, resulting in a +12 dB signal-to-noise ratio (SNR) above a very low 40 dBA background noise level where the pupils' activity noise level is assumed to add only 5 dB to the recommended 35 dBA. Thus, more than 15 dB for the SNR in a classroom should be wanted.

However, correct identification of what was said is only a necessary condition for memory and learning, but it is not a sufficient condition. There is a gap between speech intelligibility and the memory of the information conveyed in the speech. If a low intelligibility of the spoken message has exhausted the limited working memory (WM) resources little, if any, are left for the elaboration, recoding, and storing, necessary subsequent recall of the information. This result was indicated in a study by Kjellberg, Ljung and Hallman [4] where a high and a low SNR did not affect speech intelligibility, but the lower SNR impaired recall. Thus, the SNR and RvT norms for e.g., classrooms therefore should be based not only on SI but also on studies of comprehension, memory and learning.

Additivity or interactions between RvT, BL, and SNR?

A pertinent question is whether a change of e.g., 6 dB SNR causes the same amount of change in speech intelligibility with a RvT = 0.3 sec as with RvT = 0.9 sec? From a WM theoretical point of view, such a synergistic interaction would be predicted. For instance, at a lowered SI of the spoken message, e.g., because of a long RvT, the limited working memory (WM) resources have been somewhat taxed. If the SNR then is lowered by e.g., 6 dB we would expect a greater loss of SI and memory, than if the 6 dB decrement was made when the RvT was set to a more favorable and shorter RvT.

However, in building standards and recommendations, BL, SNR and RvT are basically treated as independent and *additive* variables, and not much, if any, attention is given to the possibility of their interaction.

Empirical studies of relevance for interactions between RvT, BL, and SNR

There is a lack of empirical studies of whether these variables interact or are additive. However, Klatte, Lachman and Meis [5] is one notable exceptions. They varied RvTs between 0.47 and 1.1 s, SNRs for three seat rows in the classroom, and different background noise conditions (silence, background speech and classroom noise without speech). Their results indicated that speech perception for children (1st and 3rd grade) was more impaired than for adults by background sounds both in speech perception and listening comprehension. RvT had no effect on speech perception in silence but strong effects against background noise in all age groups. RvT also interacted with age to the effect that the negative effect of a long RvT was more marked for the younger children than for the adults. There was also a significant interaction RvT and SNR (Seat row) on speech perception to the effects that decreased SNRs with sitting further back in the classroom was more pronounced in the room with a long RvT than in the room with a short RvT. However, listening comprehension was not studied under different RvTs.

For SI we a study by Payton, Uchanski and Bradia [6], reported a synergistic interaction between SNR and RvT (their fig 12 A) indicating a larger drop in speech intelligibility when the same SNR decrease is accompanied by an RvT of 0.60 s rather than with .18 and .00 s.

Valente, Plevinsky, Franco, Heinrichs-Graham and Lewis [7, Experiment 2] crossed SNR: 10 and 7 dB with RvT: 1.5 and 0.6 s., in a speech recognition task and a classroom learning task (comprehension of video recordings of a play). Subjects were children 8 - 11 years and adults. For the speech recognition task there were main effects of age, SNR and RvT, but no significant interactions between them were reported. The same was true for their classroom learning task: main effects of age, SNR and RvT, but no significant interactions between them. However, the authors stated that the differences for the sentence recognition task were minimal.

Thus, there is some evidence of interactions between RvT and SNR on SI, but not on recall/memory, and there is no report about improved SI or memory/recall from a prolonged RvT.

We have not found any studies on SI or on memory/recall that indicate interactions between RvT and e.g., language proficiency, or a positive effect of a long RvT.

ANALYSES AND REANALYSES

Two recent studies – listening and comprehension

In the first of our four recent studies with counterintuitive findings, Hurtig [8] studied pupils in grade 9 (N = 125) who took two versions of the English listening comprehension test, which is one of the Swedish National Tests for that grade. Two classrooms were used for the test, one with a short and one with a long RvT (0.33 and 1.07 s, means 250 Hz - 4 kHz). The listening test was played back from only one loudspeaker in the front of the classroom. All pupils took two counterbalanced versions of the test, one at the front and one in the middle of the classroom (1 and 6 m). We also had their scores on an English *reading* test. The results indicated that there was a significant effect of position, meaning that sitting at the front of the classroom resulted in higher scores on the English listening test. There was also an unexpected cross-over interaction between RvT and how good the pupils were at reading English. See Figure 1.The participants who scored low on the English *reading* test did unexpectedly benefit from a longer RvT, while those with a higher reading score did benefit from the shorter RvT.

In Hurtig et al. [9] Swedish college students were presented with 18 audio clips (1.5-2.0 min) in English. Half of the clips were presented with a RvT of 0.3 s and the other half (counterbalanced) with a RvT of 0.9 s. The students worked individually with a computer and the audio clips presented over headphones. Comprehension of the text was assessed with multiple-choice questions directly after each clip. All participants were also pre-tested on a WM updating test and on the Boston Naming Test to determine their baseline second-language (L2) naming ability. There was a significant synergistic between RvT and how good they were at the Boston test (median split). The high scorers did benefit from a short RvT, while those with a low score did not. See Figure 2! In contrast, there was no significant interaction between RvT and above/below median on the WM test. See Figure 3!

Also, there was no three way interaction RvT*Boston naming test*WM. This lack of interaction indicates that the effects of RvT primarily was mediated by or routed by language competence processes rather than by WM-processes.

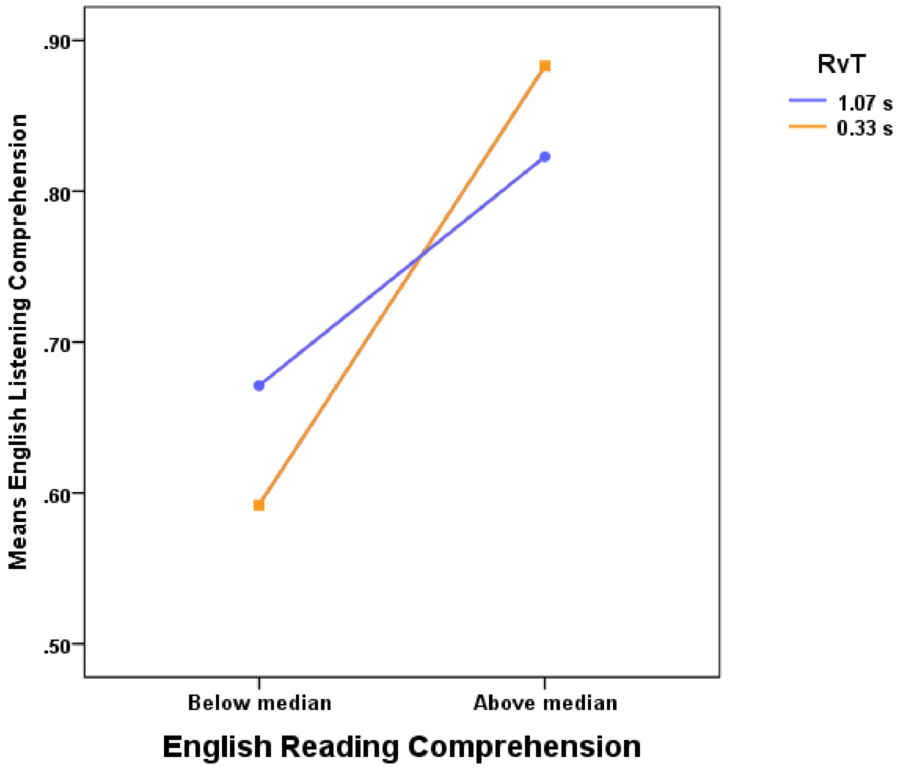

Figure 1: Means English Listening Comprehension for Grade 9 by their English reading comprehension at RvTs of 0.33 and 1.07 s. Interaction significant.

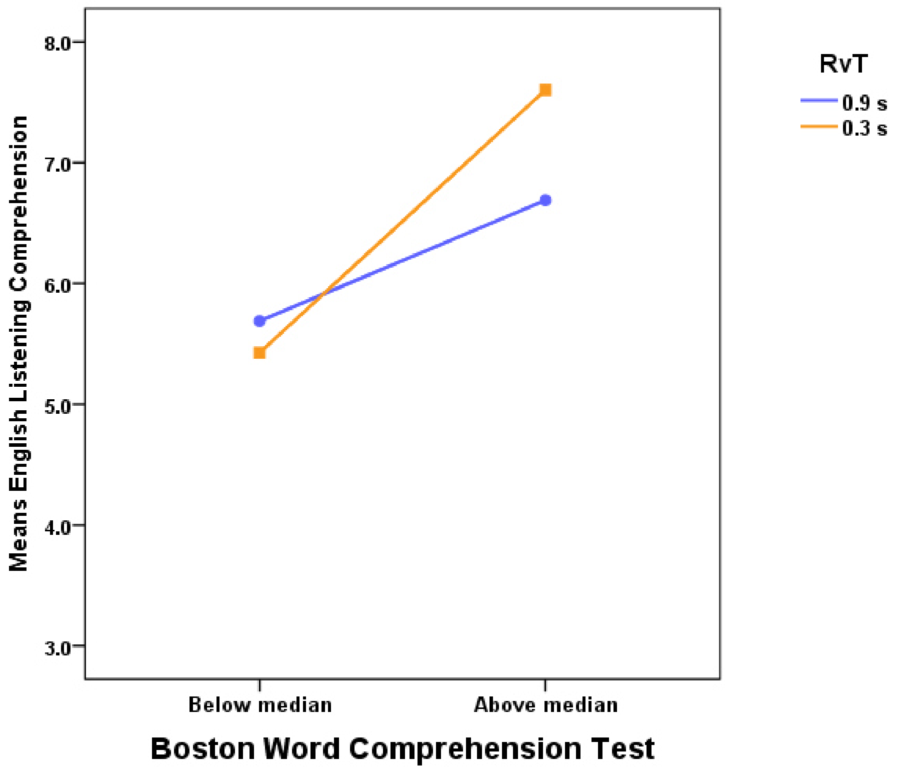

Figure 2: Means English Listening Comprehension for College students by their scores on the Boston Word Comprehension test at RvTs of 0.9 and 0.3s. Interaction significant.

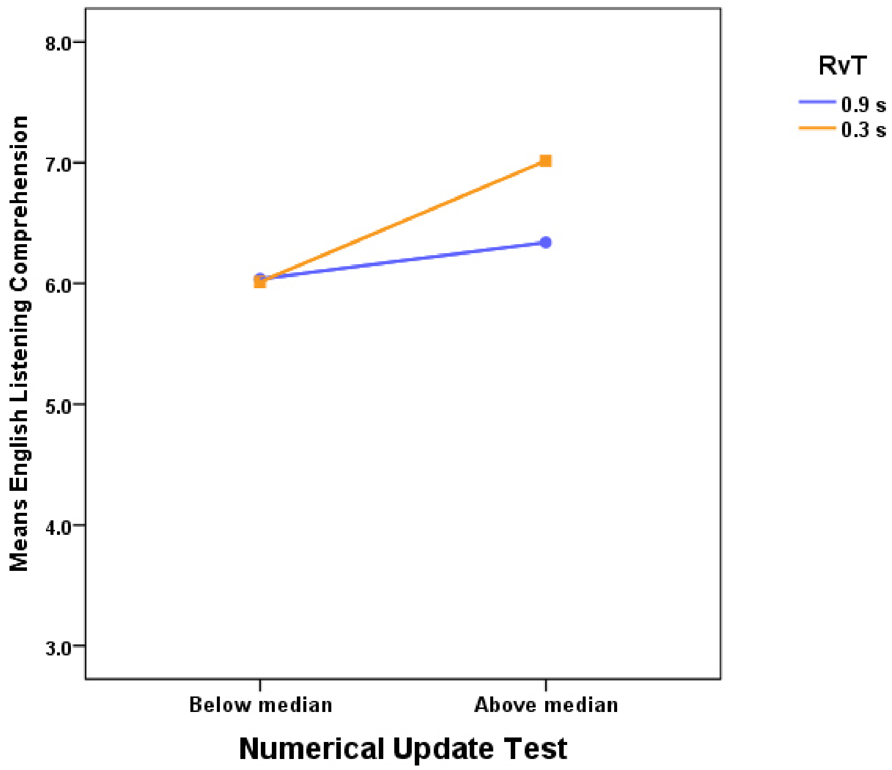

Figure 3: Means English Listening Comprehension for College students by their scores on the Numerical Update test at RvTs of 0.9 and 0.3s. Interaction not significant.

Two further recent studies – word lists and recall

In the two others of our four studies, free recall of spoken Swedish and English wordlists was assessed. The lists were presented to college students [10] (*N* = 48) and to children in grade 4 [11] (*N* = 72). The 24 word lists had 12 words each for the college students. For the children the 8 easiest words of the 12 words in the college students' lists were presented. The acoustic manipulations of the words were identical in the two groups, SNR: +3 and +12 dB, RvT: 0.3 and 1.2 s. The college students worked individually with headphones and typed the words they recalled into a computer. The children in grade 4 listened to the words lists over eight loudspeakers (uniformly ≈ 66 dBA) in their regular classroom and wrote down their replies on paper. In addition, the participants in both groups were pre-tested with working memory (WM) tests, the college students with an operation span test and the children with a size ordering task.

The children in grade 4 recalled more words when the spoken words were presented in Swedish than in English. Words that were presented with a high SNR (+12 dB) showed improved recall compared to a low SNR (+3 dB). There was also a significant unexpected crossover interaction between RvT and SNR to the effect that at +12 dB the shorter RvT improved recall, but at +3 dB it impaired recall. See Figure 4! There was also a main effect of the pre-experimental WM size-ordering task, but no interactions between this variable and other variables.

For the college students SNR and language had, as expected, strong main effects on recall and substantial effect sizes. Actually, the effects size for the recall of the words were substantially larger for recall than it was for speech intelligibility of the very same words in another experiment, also with college students [12]. There was also an interaction between language and SNR, indicating that the disadvantage in recall for English was more pronounced at +3 dB than it is at +12 dB. There was no main effect of RvT, nor of any of its interactions with the other variables. There was no main effect of the pre-experimental WM operation span task, but some interactions with other variables, but not with SNR or RvT.

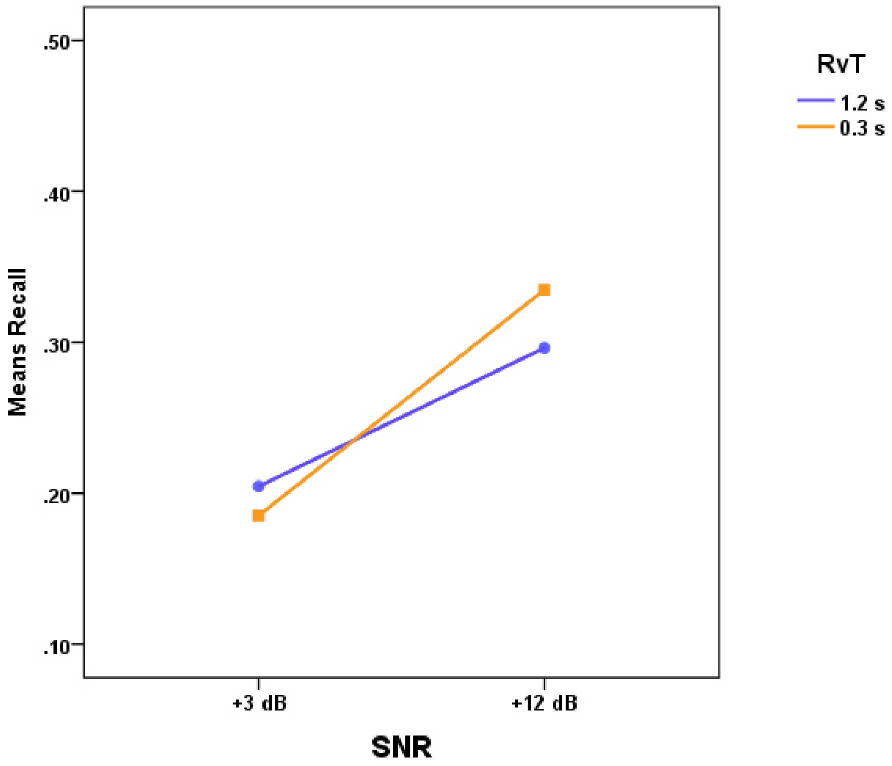

Figure 4: Recall in Grade 4 by SNR and RvT. Interaction significant.

Follow up analyses of the word lists and recall studies

Inspired by the findings from the two experiments on English listening comprehension, we first made a follow up analysis of the results for the college students. A trichotomized score of English skill was construed by taking the mean scores of each person's average scores for the English words (cf [5]). This trichotomized score showed a significant antagonistic crossover interaction with RvT, meaning that for the high scorers the short RvT yielded more recall than the long RvT, while the opposite was true for the low scorers who showed better recall with the long RvT. See Figure 5! It can be noted that splitting Figure 5 in one figure for recall of the Swedish (L1) words and recall of the English (L2) words did not change the pattern of the interaction. Only the average levels of the scores changed, but not the relative orders of the RvT-levels, which reflects that there was no significant interaction between Engskill*RvT*Language.

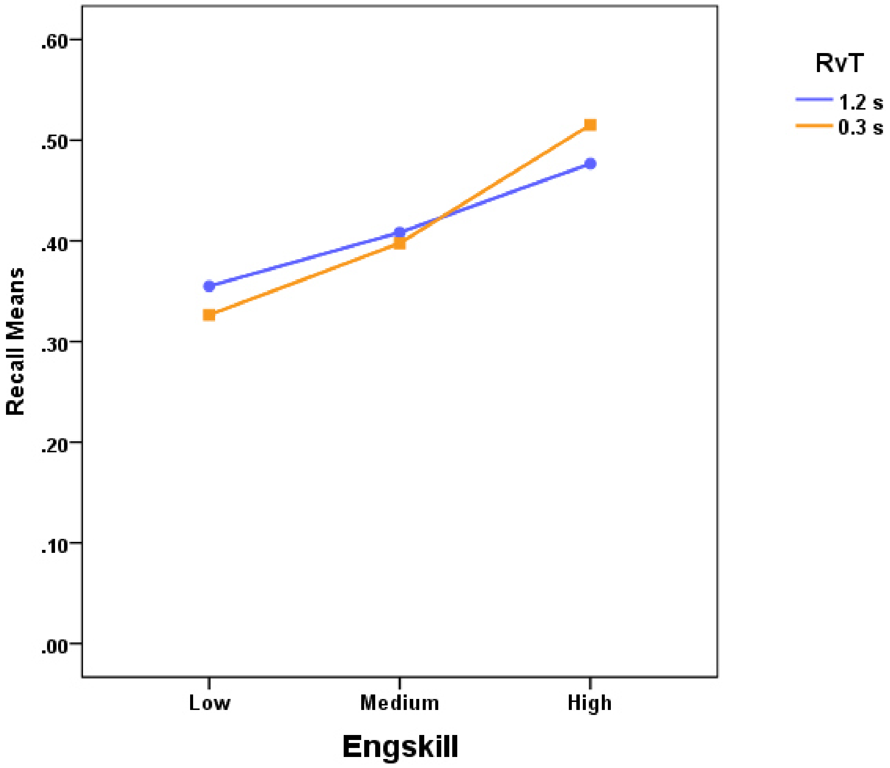

Figure 5: Recall for College students by their English skills by RvT. Interaction significant.

This result replicates the crossover interaction from the two studies above on English language comprehension with college students and with grade 9 pupils.

For the pupils in grade 4 a similar trichotomized score of skills in English indicated a difference in recall at the two RvT levels as shown in Figure 6. This interaction was not significant, but the advantage for the high skill group at the shorter RvT was significant ($p < .05$). Also here splitting Figure 6 into one figure for recall of the Swedish (L1) words and recall of the English (L2) words did not change the pattern of the interaction. Only the average levels of the scores changed, but not the relative orders of the RvT-levels, which reflects that there was no significant interaction between Engskill*RvT*Language.

For the college students, there was also trend ($p < .10$) towards a significant crossover interaction between SNR and RvT for the group with a medium score of the trichotomized score of English skill. See Figure 7! This interaction replicates the corresponding crossover interaction in the study with grade 4 in Figure 4.

Thus, within each of the four studies there are fairly consistent and recurrent interactions that question the conventional, empirically based conclusion that a short RvT improves comprehension and memory, that RvT and SNR interact, and that the advantage of a short RvT is restricted to those with a good mastery of the language in question.

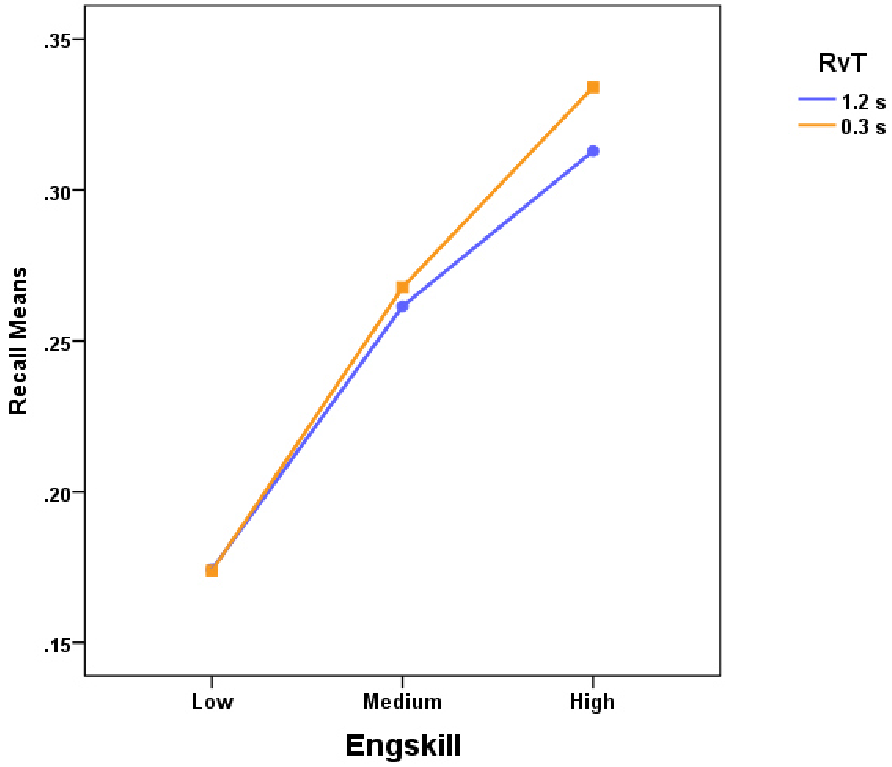

Figure 6: Recall for Grade 4 pupils by their English skills by RvT. Interaction not significant, but there is a significant difference between RvTs at the high Engskill level.

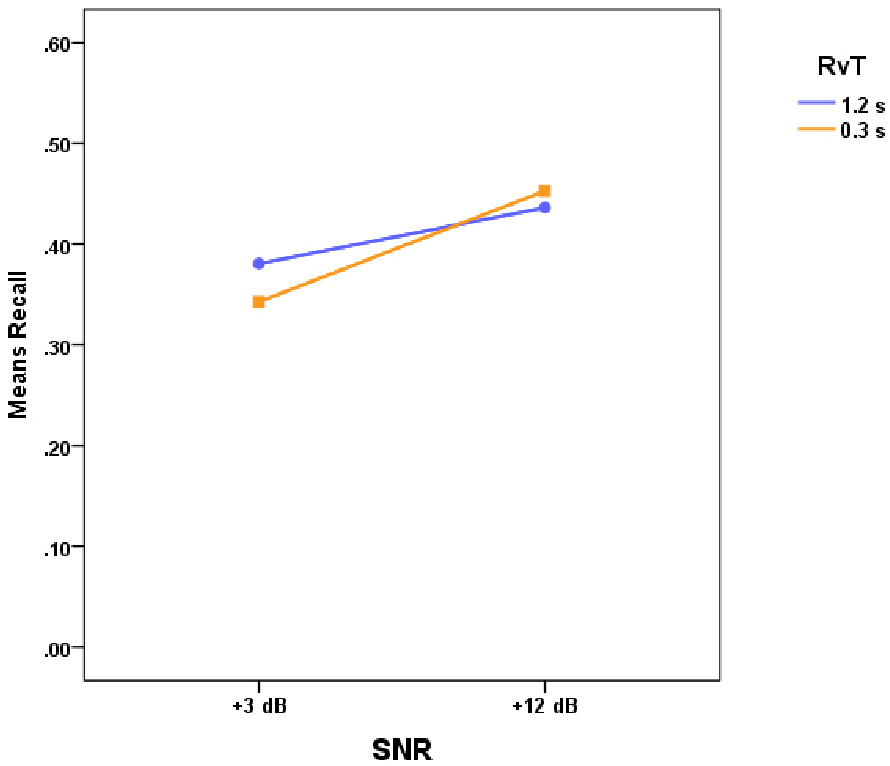

Figure 7: Recall for College students with a medium score on their English skills by RvT and by SNR. Interaction marginally significant, $p < 0.10$.

IMPLICATIONS

Theoretical

To predict the outcomes for recall in the four studies, we have relied on WM-theory and the depletion of the limited WM memory resources with an increased degradation of the speech signal with a low SNR and a long RvT. The WM-model can handle synergistic interactions, but it does not have antagonistic crossover interactions involving RvT as a logical possibility. So, if the antagonistic crossover interactions stand up for a replication, a fresh theoretical start should be considered.

In addition, the synergistic interactions we have seen involving language proficiency, rather than WM-processes, opens up the possibility that a more valid explanatory model should be sought in the theoretical area of speech and language competence, rather that within a traditional WM-theory. The fact that there was no interaction between language proficiency and WM-skills in Figures 2 and 3 above [9] also hints to the possibility that language proficiency and competence are more important that WM-processes.

Methodological

In our four studies, a traditional and blunt measure of RvT was employed, that does not make a distinction between early and late reflections of the speech signal. As the early reflections that reach the listener within 50 ms are integrated with the direct sound, they have a positive effect on speech clarity. An alternative to the traditional RvT is the Clarity index (C_{50}) which is controlled by the ratio of the energy arriving in the first 50 ms to the energy arriving later. A high value is positive for speech clarity.

It may well be the case that our unexpected findings of no positive effects in memory and recall for those with a low language proficiency, is confounded or muddled with the RvT-measure we have been using and things would look somewhat different with a Clarity index (C_{50}).

Practical implication

Our results also have some fairly immediate practical implications:

1. It is within practical reach starting to substitute the speech intelligibility criteria for good acoustic conditions with research based criteria for good memory and learning.
2. This is most urgent for learning situations involving a secondary language, or children, or immigrants, or the elderly. These are all increasing in number in our Western societies.

Acknowledgements

Both the Swedish Research Council (PI: S Hygge - A0204201) and the Swedish Research Council Formas (PI:S Hygge - 242-2010-1006) contributed to the funding of studies [4, 8-12].

REFERENCES

[1] WHO. (1999). Guidelines for Community Noise. Eds. B. Berglund, T. Lindvall & D.H. Schwela. *WHO, Geneva*. Eds.

[2] Swedish Work Environment Authority. (2005). Buller Föreskrifter [In Swedish] [Noise Regulations]. Arbetsmiljöverket 2005:26.

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

[4] Kjellberg, A., Ljung, R., & Hallman, D. (2008). Recall of words heard in noise**.** *Applied Cognitive Psychology,* 22, 1088-1098. doi:10.1002/acp.1422

[5] Klatte, M., Lachmann, T., & Meis, M. (2010). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise & Health,* 12, 270–282. doi:10.4103/1463-1741. 70506

[6] Payton, K.L., Uchanski, R.M., & Bradia, L.D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America,* 95, 1581-1592.

[7] Valente, D. L., Plevinsky, H. M., Franco, J. M., Heinrichs-Graham, E. C., Lewis, D. E. (2012). Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children a). *Journal of the Acoustical Society of America,* 131, 232-246.

[8] Hurtig, A., Sörqvist, P., Ljung, R., Hygge, S., & Rönnberg, J. (2016). Student's second-language grade may depend on classroom listening position. *PlosOne*, PONE-D-15-41873R2

[9] Hurtig, A. (2017). Comprehension of second language audio segments presented with short and long reverberations times. (Manuscript).

[10] Hygge, S. (2017). Recall of words spoken in the first and second language: Effects of signal-to-noise ratio and reverberation time. (Manuscript)

[11] Hurtig, A., Keus van de Poll, M., Pekkola, E., Hygge, S., Ljung, R., & Sörqvist, P. (2015). Children's recall of words spoken in their first and second language: Effects of signal-to-noise ratio and reverberation time. *Frontiers in Psychology,* 6:2029. doi: 10.3389/fpsyg.2015.02029

[12] Hygge, S, Kjellberg, A., & Nöstl, A. (2015). Speech intelligibility and recall of first and second language words heard at different signal-to-noise ratios. *Frontiers in Psychology,* 6:1390. doi: 10.3389/fpsyg.2015.01390