Acoustical conditions in the classroom: Recall of spoken words in English and Swedish heard at different signal-to-noise ratios

Anders Hurtig[1,2,3], Staffan Hygge[1,] Anders Kjellberg[1], Anatole Nöstl[1], Marijke Keus van de Poll[1], Robert Ljung[1], and Patrik Sörqvist[1,2]

[1] Department of Building, Energy, and Environmental Engineering, University of Gävle, SE-80176 Gävle, Sweden

[2] Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

[3] Department of Sports and Medicine, University of Dalarna, SE-791 88 Falun, Sweden

ABSTRACT

Speech intelligibility (SI) and free recall of spoken words in Swedish (native tongue) and English as second language were assessed in two signal-to-noise (S/N) conditions (3 and 12 dB), and with or without the spoken words being repeated back orally directly after presentation (shadowing). Twelve wordlists with 12 words each were presented in English as well as in Swedish. The words were chosen on the basis of their ranks in category norms for the two languages. After each wordlist, the participants wrote down the words they could recall. Pre-experimental measures of working memory capacity (WMC) were taken. The basic hypothesis for the recall of the words were that WM would be overloaded, thus interfering with storing of the information, when the S/N-ratio was low and when the language was in English. Further, the shadowing procedure was expected to improve recall. A low S/N ratio was expected to have more detrimental effects on the memory of the English words than for the Swedish words. A low score on WMC was expected to further enhance these effects. Shadowing had a more complex effect than expected, whereby it interacted with language and S/N-ratio, there was no interaction between S/N and language, and WMC did not interact with the other experimental factors. However, recall was better for Swedish (L1) words in comparison with English (L2) words and recall was better in the high S/N-ratio condition in comparison with the low S/N-ratio condition, even though S/N-ratio did not influence the participants' ability to identify the words correctly at presentation.

INTRODUCTION

The acoustical conditions in classrooms may severely impair speech comprehension, which in turn impairs learning (Kjellberg, Ljung, & Hallman, 2008). To safe-guard against inferior listening conditions in the classroom, government agencies have established building codes, standards and recommendations for acceptable S/N-ratios and reverberation time (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; Swedish Work Environment Authority, 2011; Vallet, & Karabiber, 2002). Such codes and standards have been based on what is required for correct identification of

spoken words or isolated sentences, *i.e*. speech intelligibility (SI).

Correct identification of what is said is a necessary condition for memory, but it is not a sufficient one. There are reasons to believe that memory of speech may be impaired even when SI is unaffected. The size of the gap between intelligibility and memory is likely to be a function of to which extent word identification has exhausted the limited WM-resources and, thereby, what resources are left for the elaboration, recoding and storing of information (Kjellberg *et al*., 2008). Thus, if acoustical conditions or other factors make listening more effortful, subsequent recall will suffer, even when SI is high.

This means that intelligibility should be a less sensitive indicator of the quality of listening conditions than memory of what has been said. The aim of the project to which the present study belongs, is to pin-point where acoustical conditions impair memory without affecting SI. Thereby we hope to contribute to the development of acoustical guidelines and building codes that are based on memory rather than intelligibility.

In recent papers, we have described and discussed relevant research on how background noise and poor acoustical conditions affect SI and memory (Kjellberg, *et al*., 2008; Ljung, Israelsson, & Hygge, 2013; Hygge, Ljung, & Israelsson, 2013; Ljung, & Kjellberg, 2009; Ljung *et al.*, 2009) and similar findings have been reported by others (Rabbitt, 1966; Rabbitt, 1968; Surprenant, 1999; Surprenant, 2007).

When discussing factors that make listening more effortful, a distinction can be made between critical physical (sensory) properties of the speech signal (*e.g*., S/N-ratio, RT, Speech Transmission Index (STI), articulation), the characteristics of the spoken material (*e.g*., redundancy, language, difficulty, word frequency), and the individual characteristics of the receiver (*e.g*., hearing impairment, knowledge of the language, age, cognitive capacities like working memory capacity, (WMC)). Although the characteristics of the spoken material and the individual characteristics, at least in theory, can be assessed independently of each other, there is arguably a functional equivalence between them. They all contribute to listening effort.

The present study

In the present study, we selected three variables as within person factors: (1) whether the spoken material (wordlists) was in Swedish (native tongue, L1) or in English (L2), (2) whether the words were heard under acceptable or less than acceptable S/N-ratios (12 or 3 dB) according to prevailing acoustic norms, and (3) whether the spoken words were shadowed orally directly after presentation or not. The last factor was included since shadowing was used in previous studies to control for SI and it is not known how such shadowing affects memory of the word lists. Thus, all participants encountered all experimental combinations of Language, S/N-ratio and Shadowing. In addition, WMC split by the median, was included as a between person variable.

The basic hypotheses for the recall of the words were that WM would be more likely to be overloaded and, thus, impair recall of the words when the language was English, when the S/N-ratio was low, and when the WMC for the participants was on the low side. The shadowing procedure was expected to improve recall. Interactions with WMC were also expected, such as that the participants with a high WMC, compared to those with a low WMC, will show better performance on the recall task in the conditions with lower S/N ratios and when the words are

presented in English but not necessarily in the other conditions.

METHOD

Participants and design

A sample of 48 undergraduate students all native Swedish speakers participated in the study. Self-reported normal hearing was an inclusion criterion and the subjects received a cinema ticket for their participation.

Materials, apparatus and procedure

A Swedish translation of the automated operation span (OSPAN) task (Unsworth, Heitz, Schrock, & Engle, 2005) was adopted to measure WMC. Mathematical operations [*e.g.*, "Is (5 + 3) × 3 = 24?"] were presented on a computer screen. The participant was told to respond "yes" or "no" to the operation, as quickly as possible, by pressing a button on the screen using the computer mouse. When a response was recorded, a letter was presented for 800 ms and the participant was told to remember it for later recall. When the to-be-remembered letter disappeared, a new mathematical operation was presented or the list ended. The list length varied from three to seven letters. A total of 15 lists were used (3 of each list length), and the length increased across the task. When a list ended, the participants were asked to recall the letters in order of presentation. The total number of letters recalled in the correct list position, across all lists, was used as a person specific measure of WMC, with a maximum score of 75.

Twenty-four wordlists with twelve words each were generated, twelve lists in English and twelve in Swedish. The words were taken from 24 categories and chosen from category norms for the two languages in which the words are ranked with respect to the strength of their association with the category. The distributions of the number of syllables were the same in all lists. Different categories were used for the Swedish and English lists. By means of Graeco-Latin squares the words from the 24 categories were organized into twelve Swedish and twelve English wordlists, where the average category rank order of the words was the same. Three blocks with eight lists each and counterbalanced presentation orders of language, S/N-ratios and shadowing were generated. The lists were recorded by a female speaker in a sound-attenuated chamber and normalized to 66 dB(A). The words were read with an approximately 3 s interval between the words. Pink noise was added to the word lists to create the S/N ratio conditions of 12 and 3 dB. The lists were presented to participants via Sennheiser HD-202 headphones.

Participants were instructed to memorize as many of the spoken words as possible. After each list had been presented the participants were given 1 min. to type the words they could remember from the most recent list. This procedure continued until all twenty-four lists had been presented. Participants were given one point for each correctly recalled word even if the spelling was not perfect. The scores for the correctly recalled words were entered separately for the twelve ordinal presentation positions in the lists. For the lists with shadowing the answer was recorded and each correct response was given a score of one point.

Between 1 and 3 participants were tested in each session. Each participant was seated in front of a laptop in a sound isolated test-room. The testing procedure was split into two parts. All participants started with the self-paced OSPAN task.

They were given oral instructions about the OSPAN task and to step out to the next room when finished.

The second part consisted of the spoken word lists heard under the eight experimental conditions. Each participant began with a training phase in which they listened to two lists each from the two languages, encountering one example each of the two S/N-levels combine with the two levels of shadowing. After the training phase the twenty-four word lists were presented. The duration of each word was approximately 1000 ms with a 3000 ms inter stimulus interval. The presentation order of the lists was pseudo randomized for each set of eight participants. The lists were further organized into three blocks with eight lists each, four from each language, and within each language the four combination of shadowing and S/N-ratio was counterbalanced. This response window remained open for 60 sec. and was followed by the playback of the next list. No feedback was given on either recall or identification of the words. The total session lasted 55-65 minutes depending on how fast participants completed the OSPAN task.

The WMC-scores were split by the median to form one group with high WMC-scores and one group with low scores. A split-plot ANOVA was performed with Language, Position in the word list (three parts of four words each), S/N-ratio and Shadowing as within subject factors and WMC as a between subject factor.

RESULTS

Correct identification of the words at presentation was slightly higher in the 12 dB S/N-ratio condition compared with the 3 dB S/N-ratio condition for Swedish words, but not for English words (Figure 1). However, neither the main effects nor the interaction were significant (note that the y-axis is truncated and exaggerates the visual impression of an effect).

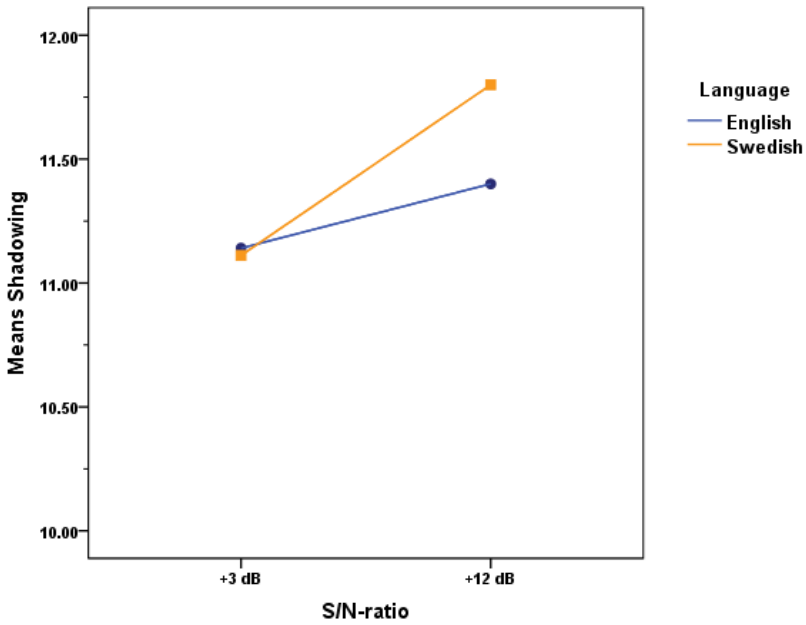

Figure 1. Mean number of correctly shadowed words with low and high S/N ratio for Swedish and English words.

As can be seen in Figure 2 (with and without shadowing), overall recall was better for Swedish in comparison with English words, and better for 12 dB S/N-ratio than for 3 dB S/N-ratio, but the two variables did not interact. This conclusion is supported by an analysis of variance with Language and S/N-ratio as independent variables. There was a main effect of language, $F(1, 47) = 24.74$, $MSE = 0.003$, $p < .001$, a main effect of S/N-ratio, $F(1, 47) = 49.40$, $MSE = 0.002$, $p < .001$, but no interaction, $F < 1$. Moreover, when shadowing (yes versus no) was added as an independent variable, the shadowing procedure interacted with language and S/N-ratio, $F(1, 47) = 9.23$, $MSE = 0.005$, $p = .004$. As can be seen in Figure 2, there was a larger difference between the two languages in the 12 dB S/N-ratio condition when the participants did not shadow the words at presentation, whereas the difference was larger in the 3 dB S/N-ratio condition when the participants did shadow the words at presentation. Adding individual differences in WMC as a further independent variable with participants divided into two groups (above versus below the median), did not add any significant main effect or interactions with the other variables.

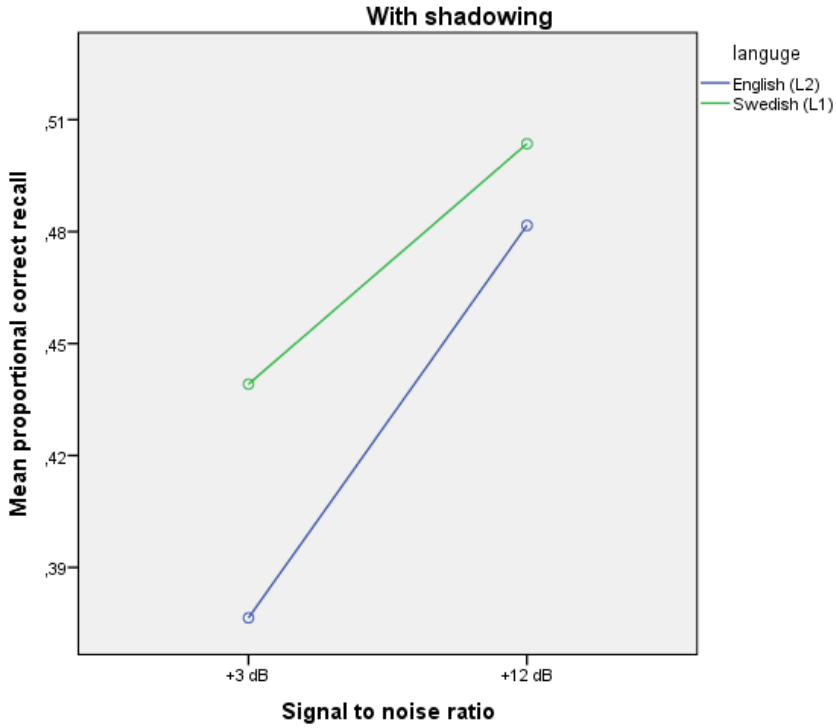

Figure 2. Recall of Swedish and English words heard at two different signal-to-noise ratio conditions.

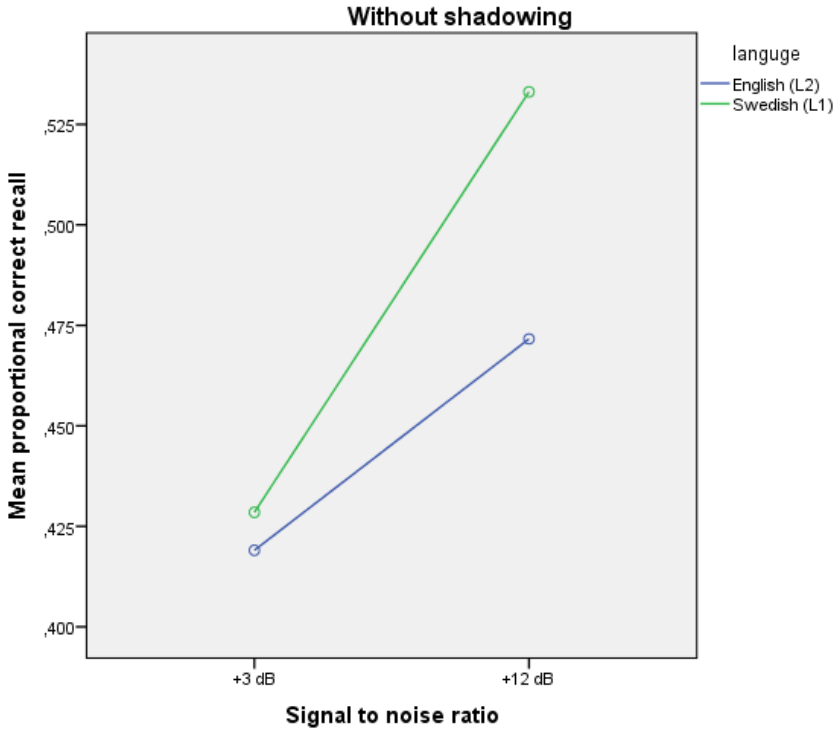

Figure 2. Recall of Swedish and English words heard at two different signal-to-noise ratio conditions.

CONCLUSIONS

Expected differences in recall were obtained between the native tongue (Swedish) and a L2 (English) as well as between the 3 and 12 S/N, but in neither case was SI affected. Thus, these results support our basic hypothesis that recall of what has been said is a more sensitive indicator than intelligibility of the acceptability of the acoustic conditions in a room.

Learning word lists is, of course, a rare task outside the laboratory. However, similar effects have been demonstrated on memory of lectures listened to in different acoustic conditions (Ljung *et al.,* 2009).

Further studies are necessary to make it possible to use the results for acoustic recommendations. As of now we can only conclude that memory seems to be a better indicator than SI of the acoustic conditions. However, we do not know what is the highest S/N or longest RT that has a measurable effect on learning. In the near future we hope to be able to report studies that give us a better basis for formulating acoustic recommendation based on the effect of acoustic conditions on learning and memory.

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.

Hygge, S., Ljung, R & Israelsson, K. (2013) “Acoustical conditions in the classroom I - Speech intelligibility and recall of spoken material heard at different signal-to-noise ratios.” Paper presented at Internoise 2013, Innsbruck, Austria, Sept 15-18, 2013.

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 decrease recall of spoken information. Building *Acoustics*, 16, 301-312.doi: 10.1260/135101009790291273

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

Rabbitt, P. (1966). “Recognition memory for words correctly heard in noise.” Psychonomic Science, 6, 383-384.

Rabbitt, P. (1968). “Channel-capacity, intelligibility and immediate memory.” Quarterly Journal of Experimental Psychology, 20, 241-248. doi: 10.1080/14640746808400158.

Surprenant, A. M. (1999). “The effect of noise on memory for spoken syllables.” International Journal of Psychology, 34, 328-333.

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.

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 (2011). Södersten, M & Lindhe, C: “Kunskapsöversikt - Yrkesrelaterade röststörningar och röstergonomi” [Review – Work related voice problems and voice ergonomics]. Report 2011:6.

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

Unsworth, N., Heitz, R.P., Schrock, J.C., & Engle, R.W. (2005). “An automated version of the operation span task.” Behavior Research Methods, 37, 498-505. doi: 10.3758/BF03192720.

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