Journal of Cognitive Psychology

ISSN: 2044-5911 (Print) 2044-592X (Online) Journal homepage: https://www.tandfonline.com/loi/pecp21

Second language vocabulary level is related to benefits for second language listening comprehension under lower reverberation time conditions

Douglas MacCutcheon, Anders Hurtig, Florian Pausch, Staffan Hygge, Janina Fels & Robert Ljung

To cite this article: Douglas MacCutcheon, Anders Hurtig, Florian Pausch, Staffan Hygge, Janina Fels & Robert Ljung (2019) Second language vocabulary level is related to benefits for second language listening comprehension under lower reverberation time conditions, Journal of Cognitive Psychology, 31:2, 175-185, DOI: 10.1080/20445911.2019.1575387

To link to this article: https://doi.org/10.1080/20445911.2019.1575387

Published online: 30 Jan 2019.

JOURNAL OF COGNITIVE PSYCHOLOGY
2019, VOL. 31, NO. 2, 175–185
https://doi.org/10.1080/20445911.2019.1575387

Second language vocabulary level is related to benefits for second language listening comprehension under lower reverberation time conditions

Douglas MacCutcheon[a], Anders Hurtig[a], Florian Pausch[b], Staffan Hygge[a], Janina Fels [b] and Robert Ljung[a]

[a]Environmental Psychology, Faculty of Engineering and Sustainable Development, Department of Building, Energy and Environmental Engineering, University of Gävle, Gävle, Sweden; [b]Institute of Technical Acoustics, Teaching and Research Area of Medical Acoustics, RWTH, Aachen University, Aachen, Germany

ABSTRACT
The acoustic qualities of a room can have a deleterious effect on the quality of speech signals. The acoustic measurement of reverberation time (RT) has shown to impact second language (L2) speech comprehension positively when lower due to release from spectral and temporal masking effects as well as top-down processing factors. This auralization experiment investigated the benefits of better L2 vocabulary and executive function (updating) skills during L2 listening comprehension tests under shorter versus longer RT conditions (0.3 and 0.9 s). 57 bilingual university students undertook L2 vocabulary, number updating and L2 listening comprehension tests. After splitting groups into high/low vocabulary and updating groups, a mixed ANOVA was conducted. The high number updating group showed no significant differences or interactions in L2 listening comprehension than the lower number updating group across RT conditions. The high vocabulary group had 22% better L2 listening comprehension than the low vocabulary group in long RT, and 9% better in short RT. A significant benefit in L2 listening comprehension due to release from reverberation was only evident in the high vocabulary group. Results indicate that the benefit of good room acoustics for listening comprehension is greatest for those with better language (vocabulary) ability.

ARTICLE HISTORY
Received 7 March 2018
Accepted 16 January 2019

KEYWORDS
Reverberation time; listening comprehension; second language learners; classroom acoustics

Introduction

Background

Why do second language (L2) learners struggle to hear what people are saying in rooms with poor acoustics, and which aspects of cognition support L2 listening comprehension? This study investigates whether higher L2 vocabulary and better executive functions improve L2 listening comprehension under normal acoustic conditions as well as those made more cognitively demanding by the manipulation of reverberation time (RT).

Reverberation time and second language listening comprehension

Speech is a major channel through which learning and remembering occur. Accurate speech comprehension depends on a clear and distinct sound signal so that the speech sounds can activate corresponding phonological representations stored in long-term memory (Rönnberg et al., 2013). Under environmental conditions in which the sound signal is degraded or distorted, a number of past experiments have shown that greater demands are placed on mediating cognitive processes. For example, poor acoustic conditions such as long RT and additional background noise have been shown to impair listening comprehension (Klatte, Lachmann, & Meis, 2010; Peng & Wang, 2016; Sörqvist, Hurtig, Ljung, & Rönnberg, 2014), memory of spoken words (Hurtig et al., 2016; Ljung & Kjellberg, 2009), memory of spoken lectures (Ljung, Sörqvist, Kjellberg, & Green, 2009) and learning (Hygge, 2003; Hygge, Evans, & Bullinger, 2002). Therefore, acoustic conditions in environments in which education takes place should provide the highest

CONTACT Douglas MacCutcheon Douglas.MacCutcheon@hig.se

quality acoustics as speech perception is such an important aspect of education. However, the value of the acoustic environment for learning is underestimated in Sweden, as a consideration of the Board of Director's regulations, BBR (Boverkets Byggregler), for classroom environments indicate. As the BBR states that sound classroom environments must be in accordance with Swedish standards (SS25268: 2007), RT in a Swedish classroom must be in the range of 0.5–0.8 s as per specifications for buildings in sound Class C. It is unclear what motivates the BBR's choice of RT for Class C sound environments, but how speech perception is negatively affected as a function of increased RT has been considered in a number of studies (Hurtig et al., 2016; Klatte et al., 2010; Klatte, Bergstroem, & Lachmann, 2013; Ljung, 2010; Nábělek & Donahue, 1984; Sörqvist et al., 2014).

A sound source in a room is perceived as a combination of its direct sound, which is the sound travelling straight from the source to the listener, and so-called "early" and "late" reflections from nearby reflective surfaces such as the floor, walls and ceiling (Kuttruff, 2016). Depending on listener position and room geometry, "early" reflections are those occurring within a certain period of time after the direct sound (e.g. 80 ms) and are perceptually reinforcing the direct sound by making it appear louder while increasing speech intelligibility (Kuttruff, 2016). "Late" reflections can be seen as a dense accumulation of reflections that evolve stochastically after the early reflections (Kuttruff, 2016). Commonly, RT is used to describe the reverberant nature of a room and is defined as the time (in seconds) it takes for the energy decay curve to drop by 60 dB after a sound event occurred. RT depends on the volume of the room, its surface area, and material properties, the latter defining the amount of absorption and scattering of reflections (Vorländer, 2007). Larger room volumes result in longer RTs, which in turn increase the perceptual smearing of the speech signal due to the manner in which late reflections interfere with the direct sound (Everest & Pohlman, 2015; Kuttruff, 2016). Therefore, acoustical design that keeps RT in an optimal range (i.e. lower) potentially leads to more successful verbal interaction among talkers and listeners.

The combined effects of temporal masking, spectral masking and RT could theoretically explain problems with listening comprehension under increasingly reverberant listening conditions. However, all vowels fall below or within the frequency range of consonants (Scheuerle, 2000), and spectral masking of speech occurs when lower frequency components of the speech signal (e.g. vowels) mask higher frequency counterparts (e.g. consonants), referred to as "upward spread of masking" (Moore, 2012). This poses a problem for speech perception because the informational component of speech is far more dependent on consonant intelligibility than vowel intelligibility (Ljung, 2010). Additionally, vowels are voiced and subsequently louder than unvoiced consonants. Temporal masking of speech occurs with increasing RT as the duration of louder sounds (i.e. vowels) increases, mixing with late reverberations from preceding phonemes and adding to noise masking the target speech (Everest & Pohlman, 2015; Klatte et al., 2010; Klatte et al., 2013; Moore, 2012). Therefore, a longer RT can degrade the speech signal temporally and spectrally, affecting the perception of the segmentation of speech by making word onsets unclear, impairing perception of prosodic cues and smearing the boundaries between words (Lecumberri, Cooke, & Cutler, 2010).

The role of cognitive factors second language listening comprehension under degraded conditions

Cognitive factors mediate the central processing of speech in reverberant sound environments in various ways, by helping listeners to track and remember what is being said or to use contextual and other cues (e.g. word frequency, lexical neighbourhood) to inform guesses and thereby fill in the perceptual gaps when uncertain (Lecumberri et al., 2010; Wingfield, Alexander, & Cavigelli, 1994). When the signal is degraded, for example, by excessive reverberation or noise, listening comprehension could be compromised if the listener incorrectly fills in the missing information with likely alternatives based on incorrect lexical guesses or expectations. As L2 listeners' cognitive resources are primarily taken up with bottom-up perceptual decoding integration when the signal is degraded, referred to as "bottom-up dependency", less resources remain for semantic integration of the words into larger units of meaning (Field, 2004). Therefore, contextual cues would be less useful in directing lexical expectations when the signal is degraded, leading to the perpetuation of a circle in which misheard words build incorrect contexts which misinform further incorrect lexical predictions (Field, 2003). Studies on adults have found that the listener's language background further mediates L2 listening comprehension at the cognitive level. Stæhr (2009) indicated that the degree of L2 vocabulary knowledge is predictive of L2 listening comprehension generally, and Bradlow and Bent (2002) state this to be a result of a mismatch between first language (L1) and lexically similar L2 items. Studies investigating L2 speech perception in a number of different noise types indicated the benefits of general proficiency in L2 for L2 speech perception in noise (Kilman, Zekveld, Hällgren, & Rönnberg, 2014; Peng & Wang, 2016). However, when the signal is degraded by RT, a cumulative effect of signal distorting factors in conjunction with L2 levels of ability has been noted (Nábělek & Donahue, 1984; Sörqvist et al., 2014; Takata & Nábělek, 1990). Nábělek and Donahue (1984) showed that, even when L2 proficiency was high, there was a 10% drop in L2 listeners' perception of consonants relative to L1 listeners when signal distortion due to increase in RT was introduced. A study by Takata and Nábělek (1990) indicated a drop in L2 speech perception from 97% to 73% when reverberation was introduced (RT = 1.2 s). The drop in performance was 8% worse than the corresponding drop in L1 listeners. Meanwhile, Sörqvist et al. (2014) indicated that L2 proficiency (measured by a L2 reading comprehension test) could help mitigate the negative effects of a longer RT.

Working memory capacity (WMC), a limited capacity cognitive system for the temporary storage and rehearsal of sensory information (Engle, 2002), has been considered to play a further mediating role in listening comprehension in degraded conditions. Speech perception impaired as a result of hearing loss, versus signal distortion due to environmental factors (e.g. noise, RT), is similar to the extent that they are both theorised to impair perception by placing additional demands on cognition. In listeners with hearing impairment, the benefits for speech perception of improvements in the speech processor in hearing aids have been linked to WMC, specifically the freeing up of cognitive resources (described as "cognitive spare capacity") for higher order language processing (Rudner & Lunner, 2014). Rönnberg et al. (2013) Ease of Language Understanding (ELU) model of speech comprehension describes this process in reverse, theorising that perceptual deficits such as hearing loss divert cognitive resources away from top-down language processing, forcing the listener to devote more cognitive processing to assist with perceptual processes. However, environmental degradation due to RT in those with ordinary hearing is also thought to place similar demands on cognitive processes. For example, Kjellberg (2004) theorised that speech degraded by RT increases the number of lexical possibilities, thereby increasing the amount of processing required to match the degraded signal with representations stored in long term memory, turning an almost automatic bottom-up process into a cognitively demanding top-down one. Field (2003, 2004) has advanced the view that L2 listeners are already particularly dependent on bottom-up perceptual processes than L1 listeners. This provides grounding for the hypothesis that the effect of RT on cognitive processes could be even further compromised and that the benefits of reducing RT would be especially pronounced in L1 listeners. Sörqvist et al. (2014) reported correlations between WMC and L2 listening comprehension under three RT conditions. It has been shown that executive functions and WMC underlie a wide range of complex cognitive processes, including reading and language comprehension (Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Engle, Cantor, & Carollo, 1992; Engle, Nations, & Cantor, 1990; Turner & Engle, 1989), reasoning (Daneman & Carpenter, 1980), word-problem solving (Wiley & Jarosz, 2012), learning to spell (Ormrod & Cochran, 1988) and learning a new vocabulary (Daneman & Green, 1986). According to Miyake et al. (2000), executive functions as a subprocess of working memory are based on three key elements, shifting, inhibition and updating, which together help people stay on task by suppressing task-irrelevant information (Sörqvist, Halin, & Hygge, 2010). Shifting is the ability to switch attention back and forth between mental sets or operations and inhibition is the ability to inhibit unwanted information (Miyake et al., 2000). Updating is intrinsically related to WMC (Carretti, Cornoldi, & Pelegrina, 2007) as it is theorised to monitor incoming information and revise items held in working memory by replacing old information with new, relevant information (Miyake et al., 2000). It is therefore theorised to play a significant role in the accurate recall of sequences of words, and indeed, poorer updating has been associated with lower recall and a higher incidence of misremembered items (Carretti, Cornoldi, De Beni, & Romanò, 2005; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001). Muijselaar and de Jong (2015) have considered the function of updating in reading comprehension. They theorised that updating could be related to how a listener builds up a so-called "situation model" of a text, which is effectively a mental model of the situation being narrated, built by successively updating events in memory as the narrative unfolds in time. However, they found no effect of updating on reading comprehension in their study when considered at such a broad level. It should be noted that updating could be applicable to language comprehension in quite different contexts depending on the task and the way in which updating is assessed.

The present study

This research investigated L2 listening comprehension in a virtual classroom context. All stimuli were simulated in virtual acoustic environments based on actual acoustic measurements taken at two participating schools in Sweden: one with poor acoustics, specifically a longer RT of 0.9 s, and one with good acoustics, with a shorter RT of 0.3 s. This study builds on Sörqvist et al. (2014) by evaluating the role of L2 proficiency using a more exact measure of L2 proficiency, namely L2 vocabulary size, in helping bilingual listeners cope with L2 speech in poor acoustics under more plausible RT conditions. We therefore aimed to investigate whether those falling into higher vocabulary groups are less susceptible to the negative effects of a longer RT on L2 listening comprehension. As a secondary question, we asked whether updating, as a subcomponent of working memory relevant to speech perception, provides additional advantages in helping individuals to cope with poor acoustics during L2 listening comprehension tasks. We therefore hypothesised that higher L2 proficiency and stronger memory updating skills in bilinguals could provide advantages in L2 listening comprehension under optimal listening conditions, aiding the release from reverberation (i.e. release from spectral masking and temporal masking), and leading to higher listening comprehension scores in groups split according to L2 vocabulary and updating ability.

Methods

Participants

A total of 57 participants (23 male) with a mean age of 25.26 years (SD = 5.712) and Swedish as a first language and English as a second language took part in the experiment and received a cinema ticket for their participation. An in-house questionnaire was used to determine participants' language backgrounds and to screen for diagnosed hearing, language and/or cognitive deficits.

Cognitive assessments

Boston Naming Test

The Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 2001) is a picture naming test in English that was used to determine baseline second-language naming ability for all participants in this experiment. Participants were required to match 60 images with the correct word presented in a multiple choice format. For example, the first picture item was bed and the four multiple choice word options were bell, sleep, bed and pillow. Participants were first given two practice pictures developed in-house with additional response feedback to ensure that they understood their task. There was a 5-s pause between items. The entire test had a mean running time of 7 min (range: 5.4–10.1 min). Results were scored based on the total number of correct responses for each participant.

Number Updating Test

A number updating task was adopted from Carretti et al. (2007) for this test. The test material consisted of 12 lists each containing 10 different two-digit numbers and 2 training lists. The numbers were presented in the centre of a computer screen using 72-point Helvetica at a screen resolution of 1366×768 pixels. Prior to each number list, the symbol ## was displayed to indicate the position of the subsequent ten numbers. Thereafter, all numbers of the list were displayed consecutively for 2 s. After the last number, a blank screen was displayed for 1 s before the response screen was shown. The participants were instructed to recall the three lowest numbers of the presented list in their order of appearance. After typing the response into three different edit boxes the subjects pressed a "Next" button to continue with the subsequent number list. Before the actual test, the subjects were familiarised with the procedure by presenting two training lists with additional feedback. The numbers of the lists were distributed pseudo-randomly between 17 and 94 and showed an arithmetic distance of 2 and 7 between sorted adjacent numbers. All lists exhibited an arithmetic range between 30 and 44. The lists varied in difficulty as one half required two updates, i.e. how often the subjects must discard one number of a previously memorised number triplet, whereas the other half required five updates. Per list, all presented numbers only occurred once. All lists were presented in the same order to all participants. For each correctly recalled number at the correct serial position, one point was given, with a maximum score of 36 if three items were recalled correctly for all 12 lists.

Listening comprehension test

Stimulus material

The monaural audio files News Items 1 and 2 from section "Focus: Listening" (Nilsson, 2016), which form part of the National Assessment Project (Erickson, 2009), were used for the listening comprehension test. The anechoically recorded stimuli of the first and second news item set consisted of 7 and 12 news items, respectively, and included a signature sound with a length of 1.9 s. Each news item was presented by nine different female and male talkers. To allow for balancing test conditions, we removed the last news item of the second set, resulting in 18 different news items; 9 news items for each RT condition. Participants were split into two groups, where the first group with odd subject IDs received odd news items simulated in condition RTlow and even news items in condition RThigh, and vice versa for the group with even subject IDs. While the news items order was maintained, the RT simulation was alternated, starting with RTlow and RThigh for subjects with odd and even IDs, respectively. To avoid fatigue, the listening comprehension test was split into two parts, news item set 1 and 2, with a break of 1 min between parts and a countdown of 3 s, displayed on screen, before each news item set. Each news item was only presented once. The participants were prompted with a multiple choice question after each news item without receiving any feedback if their answer was correct or not. All participants were instructed by the moderator to move on to the next question when hearing a signature sound. The signature sound was played 20 s before and after each news item to indicate the timeframe within which to focus on and answer each question. The results were evaluated under the following criteria: one point was given for every correct answer with a maximum of 18 points, resulting in 7 and 11 points, respectively, for the first and second set of news items.

Simulation of the virtual acoustic environment

The listening comprehension task was conducted under simulated room acoustic conditions with low and high RTs, reflecting two physically existing classrooms in two different Swedish schools, which were used for teaching purposes during the time the study was conducted. Room acoustic measurements in both classrooms were taken according to ISO (2008). Classroom 1 (Eslov) was treated acoustically, including additional low-frequency absorber elements, which resulted in linear RTs T_{30} (i.e. the time the normalised energy decay curve exhibits to drop by 60 dB, calculated by extrapolation based on the time the energy decay curve takes to decay from 5 to 35 dB with respect to the initial value) across all measured octave band centre frequencies, exhibiting a mean RT of 0.3 s, as shown in Figure 1. Representing a room with high RTs, Classroom 2 (Höllviken) was selected and shows a distinct increase in RTs towards lower frequencies with a mean RT T_{30} of 0.9 s, see Figure 1. Based on these two real-world classrooms, a three-dimensional computer-aided design (CAD) classroom model ($L \times W \times H = 11.8 \times 7.6 \times 3\ m^3$, $V \approx 262\ m^3$) was created in SketchUp (Trimble Navigation Ltd., Sunnyvale, CA). The simulated room was equipped with typical furniture while assigning absorption and scattering coefficients to the room's surface materials (Vorländer, 2007). The simulation of binaural room impulse responses (BRIRs) was carried out using Room Acoustics for Virtual Environments (RAVEN), a real-time framework for the auralization of

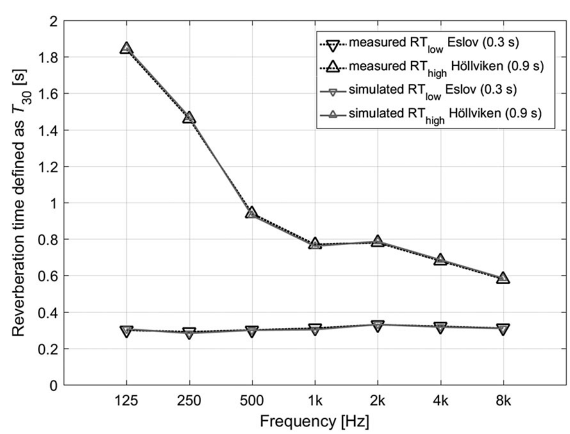

Figure 1. Simulated reverberation times T_{30} in octave bands with centre frequencies between 125 and 8000 Hz. The simulation was optimised to match the reverberation times obtained from room acoustic measurements according to ISO 3382-2 (ISO, 2008) in two different classrooms with high (Höllviken) and low (Eslov) reverberation times, RThigh versus RTlow, respectively. The values in brackets show the mean mid-frequency reverberation time (arithmetic mean of 0.5–1 kHz octave bands).

interactive virtual environments (Pelzer, Aspöck, Schröder, & Vorländer, 2014). Auralization in RAVEN is based on efficient geometrical acoustics simulation models (Kuttruff, 2016) and utilises an image source model for the generation of early reflections whereas late reverberation is simulated using a ray-tracing algorithm (Schröder, 2011). In an iterative optimisation routine, acoustical properties of the surface materials were further optimised to match the RTs T_{30} in octave bands obtained from room acoustic measurements in the two different Swedish schools, as depicted in Figure 1.

Virtual sound sources

For the generation of virtual sound sources (VSSs) a set of head-related transfer functions (HRTFs) measured from an adult head and torso simulator (Schmitz, 1995) with a spatial resolution of 1° × 1° in azimuth and zenith angles, covering the whole sphere and measured at a radius of 1.86 m, were used. Based on this data set, BRIRs with a filter length of 3 s were simulated for the two conditions RT_{low} and RT_{high}, as depicted in Figure 1, describing the acoustic transfer functions between a sound source and a listener in the simulated room. The virtual sound source was located at the typical position of a teacher in a classroom at a height of 1.6 m in front of a seated virtual listener with an ear height of 1.2 m at a distance of 6 m. The virtual listener's position and orientation was static, facing the virtual teacher. To obtain virtual sound sources, i.e. binaural signals, the relevant BRIR set was convolved with the respective stimulus selection. Verbal task instructions, verbally presented questions to each news item, as well as the signature sound were presented monaurally to ensure intelligibility and audibility, and were normalised to a common loudness level (DIN, 2010; ISO, 1975) together with the news items simulated in the two reverberant conditions.

Apparatus

The experiment was conducted in a lecture room at the University of Gävle. Each participant was seated in front of a laptop (DELL, Latitude E6430) and was provided with a pair of headphones (Sennheiser, HD 202). Monaural and binaural audio signals were played back over the headphones at a playback level of approximately 65 dB(A) using the on-board sound chip of the laptop.

Design and procedure

A within-subjects design was used. Data collection was done in 15 sessions with 1–10 participants on 3 different days. Each participant took approximately 1 h to complete the experiment. The experiment was divided into four parts: after a short introduction by the experimenter, the participants had to complete a questionnaire about personal data (gender, date of birth), language background (mother tongue, language of mother and father, home language, education language) and disabilities (hearing loss, dyslexia or other language disorders). Thereafter, the cognitive assessments (number update test and the Boston naming test) were conducted, followed by the listening comprehension evaluation. To minimise the influence of ambient noise during the listening comprehension test, all subjects were instructed to wear the headphones throughout the whole experiment even if no audio material was played back. After finishing the experiment, all participants were instructed to remain seated and quiet until all other participants had finished.

Results

Individual difference measure

The mean score on the number updating test was 19.07 ($SD = 6.40$, range: 5–36) and the BNT the mean score was 47.42 ($SD = 5.91$, range: 28–59). Mean score on the listening comprehension test for the subset of news items with a maximum score of 9 presented under 0.9 s RT was 6.35 ($SD = 1.34$, range: 3–9), and mean score for the listening comprehension test with a maximum score of 9 for those presented with an RT of 0.3 s was 6.91 ($SD = 1.85$, range: 1–9), corresponding to 70.5% and 77% correct, respectively. So as to observe interaction effects in the subsequent ANOVA reported in the next section, two between-subjects factors were created by median splits of participants into two sets of groups: vocabulary and number updating. The first set of groups either achieved high or low vocabulary scores in the BNT (median = 48), named High Vocab ($n = 30$, mean = 51.60, SD 3.480, range: 48–59) and Low Vocab ($n = 27$, mean = 42.78, SD = SD 4.388, range: 28–47) groups. The second set of groups based on their scores on the number updating test (median = 19), named, High Updating ($n = 32$, mean = 23.50, $SD = 3.959$, range: 19–36) and Low Updating ($n = 25$, mean = 13.40, $SD = 3.990$, range: 5–18).

Second language listening comprehension, number updating and vocabulary

A three-way mixed analysis of variance (ANOVA) was conducted to investigate the effect of the within-subject factor RT, at two levels Shorter RT and Longer RT, on listening comprehension score. Between-subject effects for vocab groups were significant, $F(1, 53) = 20.475$, $p < .001$, $\eta_p^2 = .279$, but number updating between-subject effects were non-significant, $F(1, 53) = 2.372$, $p = .129$, $\eta_p^2 = .043$. The Low Vocab mean score on the dependent variable, listening comprehension, in Longer RT was 6.0 (SD = 1.46, range: 3–8) and 6.0 in Shorter RT (SD = 2.02, range: 1–9). High Vocab mean listening comprehension score was 6.7 in Longer RT (SD = 1.16, range: 4–9) and 7.73 in Shorter RT (SD = 1.23, range: 5–9). The Low Updating mean listening comprehension score was 6.4 in Longer RT (SD = 1.26, range: 4–8) and 6.8 in Shorter RT (SD = 1.76, range: 2–9). High Updating mean listening comprehension score was 6.3 in Longer RT (SD = 1.42, range: 3–9) and 7 in Shorter RT (SD = 1.95, range: 1–9).

Between-subjects factors were evaluated for interactions on the independent variable, RT. ANOVA preliminary assumption testing was conducted using Shapiro–Wilk and Levene's tests and no violations of normality or homoscedasticity were indicated ($p > .05$). A significant main effect of RT condition was observed, $F(1, 53) = 4.211$, $p = .045$, $\eta_p^2 = .074$. Figure 2 illustrates the significant interaction and RT condition and Vocab, $F(1, 53) = 4.211$, $p = .045$, $\eta_p^2 = .074$. A simple effects analysis was conducted in order to break down the Vocab-RT interaction effects. To assist interpretation, the advantages in listening comprehension when moving from a long to shorter RT were converted to percentages calculated on the basis of differences between conditions/groups estimated marginal means relative to the maximum value of the comprehension test (i.e. out of 9). The Low Vocab group showed a statistically non-significant difference in listening comprehension between RT conditions of close to zero (2.665e-15), corresponding to a 1% difference in estimated marginal means in listening comprehension between the two RT conditions, $F(1, 53) = 0.000$, $p = 1$, $\eta_p^2 = .000$. The High Vocab group showed a significant increase in average comprehension of 1.14 points (16%) when moving from a long to shorter RT condition, $F(1, 53) = 9.721$, $p = 0.003$, $\eta_p^2 = .155$. Differences between High and Low Vocab were significant in both the Longer RT condition, $F(1, 53) = 5.076$, $p = 0.03$, $\eta_p^2 = .087$, and Shorter RT condition, $F(1, 53) = 18.111$, $p < 0.001$, $\eta_p^2 = .255$, corresponding with listening comprehension advantages of .81 points (9%) and 1.96 points (22%). The interaction of RT condition with Updating was not significant, $F(1, 53) = 0.987$, $p = .325$, $\eta_p^2 = .018$.

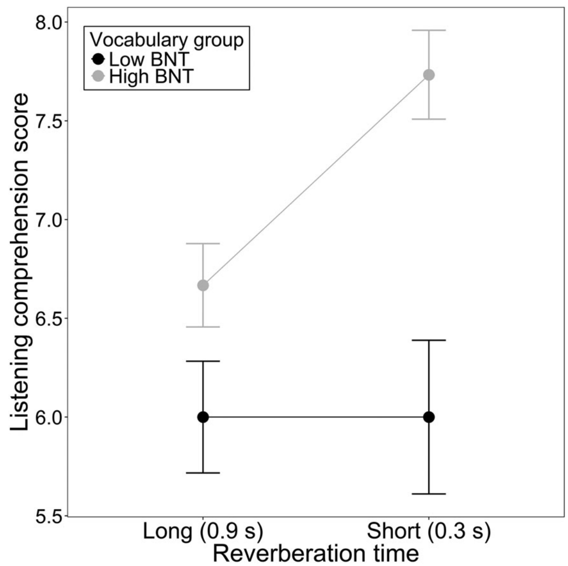

Figure 2. Interaction between vocabulary and reverberation time (RT). Results shows a 14% increase in listening comprehension score for High Vocab group moving from a long to shorter RT, and a less than 1% increase in Low Vocab group moving from a long to shorter RT. Error bars indicate standard error. A sound source in a room is perceived as a combination of its direct sound, which is the sound travelling straight from the sound source to the listener, and so-called "early" and "late" reflections from nearby reflective surfaces such as the floor, walls and ceiling.

Discussion

This quasi-experiment evaluated the benefits of having better second language (L2) vocabulary and executive function skills when performing L2 listening comprehension tests under shorter versus longer RT conditions. The shorter and longer RTs used in this study were 0.3 and 0.9 s respectively, and the main effect of RT showed that listening comprehension was significantly better in the shorter RT condition, amounting to an increase in comprehension corresponding to 0.6 points (7%) on the listening comprehension test when moving from a long to a shorter RT. Sörqvist et al. (2014) found no significant difference in listening comprehension between their 0.3 and 0.9 s RT conditions, but found differences between 0.3 and 0.9 s and their worst condition of 1.8 s. However, this research used virtual acoustic simulations reproducing acoustical conditions encountered during actual measurements of two real classrooms in order to more closely approximate real-world listening environments. It has long been known that a longer RT negatively affects speech comprehension (Kjellberg, 2004), therefore this finding corroborates studies showing the negative effects of a longer RT on speech comprehension (Bradley, 1986; Ljung, 2010; Stansfeld et al., 2005). We further evaluated whether those falling into a higher L2 vocabulary group have better listening comprehension speech perception under difficult acoustic conditions. Unsurprisingly perhaps, the higher L2 vocabulary group had better L2 listening comprehension scores in both acoustic conditions. This corroborates the findings of Sörqvist et al. (2014) who established strong relationships between degrees of L2 proficiency (reading comprehension) in bilinguals and their L2 listening comprehension scores.

The interesting interaction between high/low vocabulary and RT conditions depicted in Figure 2 requires further explanation. Upon dividing our sample into high and low L2 vocabulary groups, we found that those falling in the lower vocabulary group performed only negligibly better in the Shorter RT condition, in the order of a 1% increase in listening comprehension (percentage calculated on the basis of differences in estimated marginal means). Meanwhile, the High Vocab group experienced a 14% advantage in listening comprehension in the Shorter RT condition, corresponding to about 1.3 items. Based on the theories already discussed, this increase might be attributable to release from reverberation, which is a release from the effects of spectral and temporal masking thought to disadvantage listening comprehension due to perceptual smearing as RT increases, as explained in the Introduction. Two theories are here proposed which might explain this interaction by looking at the results in two different ways. The first way views the data from the perspective that the High Vocab group actually did far worse than could be expected in the Longer RT condition, in relation to the Low Vocab group. The fact that there are more options in the mental lexicon available to the higher vocabulary group under conditions in which part of the word is masked could actually be detrimental to them when the signal is degraded because there would be a greater possibility that they would substitute an incorrect word as a result of having a higher number of potential alternative (and incorrect) words to draw on. The converse and more credible view is that the higher vocabulary group did far better than should be expected in the Shorter RT condition, indicating an advantage for listening comprehension in less reverberant conditions due to better L2 proficiency. This release from reverberation benefit in the High Vocab group indicates that this benefit only occurs once a certain level of L2 vocabulary has been achieved. As vocabulary is an essential indicator of general language ability, superior L2 vocabulary could also indicate superior top-down language processing in other linguistic domains, such as a greater ability to make use of contextual cues to increase the accuracy of guesses when uncertain. Therefore, benefits attributable to the resulting signal clarity due to release from spectral and temporal masking could have compounded with other aspects of L2 proficiency, such as the ability to utilise contextual cues and other language benefits not directly related to L2 vocabulary but which a higher L2 vocabulary is indicative of. Described as the hard problem in bilingual lexical access, bilinguals are thought to process double the number of associated items that are closely related to the target word/concept during word retrieval (Finkbeiner, Gollan, & Caramazza, 2006). Better L2 language proficiency is certain to make word selection more efficient, and likewise optimal acoustic conditions could compound, providing additional benefits for selecting between L1 and L2 items in the mental bilingual lexicon. It should be cautioned that this finding does not suggest that acoustic conditions do not matter to L2 learners in the early stages, but simply that the positive effects of optimised acoustic conditions are much greater for those with better L2.

The relationship between updating and L2 listening comprehension was also evaluated. It was hypothesised that updating could be implicated in listening comprehension due to its role of monitoring and updating information (e.g. words) held in working memory by replacing old information with new, relevant information (Miyake et al., 2000). L2 listeners are known to be more disadvantaged than native listeners when listening to L2 speech in poorer RT conditions (Nábělek & Donahue, 1984), and this disadvantage has been linked to underlying cognitive processes, particularly those related to WMC (such as memory updating). Based on the ELU model of speech comprehension (Rönnberg et al., 2013) described in the Introduction, it has been posited that cognitive auditory information processing is taxed as a result of a longer RT, as the limited amount of mental resources available for both perceptual and cognitive processes are primarily diverted towards perception as listening conditions become more difficult (Ljung, 2010). Furthermore, there is evidence that L2 listeners place greater emphasis on decoding L2 words than forming higher level units of meaning required for listening comprehension, referred to as a "bottom-up dependency" (Field, 2003, 2004), indicating that L2 listeners' mental resources are diverted away from cognitive processing to perceptual processing. Therefore, this study hypothesised that better cognitive processing (i.e. higher memory updating ability) would provide advantages for L2 listening comprehension under difficult conditions because the mental resources left over for cognitive processing would be more efficiently utilised. Indeed, Sörqvist et al. (2014) found a significant correlation between WMC and L2 listening comprehension and deficits in updating have been linked with impaired recall and misremembering in other studies (Carretti et al., 2005; Palladino et al., 2001). At a higher level theorised by Muijselaar and de Jong (2015), updating could also be related to how a listener builds up a so-called "situation model" of a text by updating events in memory as the narrative unfolds. However, these researchers found no convincing effect of updating on reading comprehension (Muijselaar & de Jong, 2015). It should be noted that updating could be applicable to language comprehension in quite different contexts depending on the task and the way in which updating is assessed. Although the task used in the present study required updating at both the level of monitoring and revising incoming information in working memory and simultaneously updating the "situation model" of a text, our findings did not confirm any statistically significant relationship either, and therefore extend those of Muijselaar and de Jong (2015) to the context of listening comprehension.

Limitations of the study

It should be clarified that the mean RT values which are suggested by the BRR are not indicative of RT distribution in a room, which poses a danger illustrated by the classrooms simulated in this study. Figure 1 depicts simulated RTs T_{30} in octave bands. As can be seen in Höllviken, the Higher RT room, the distribution of RT in the room at different octave bands is unequal, with increasingly higher RTs in frequencies below 1 kHz. As the majority of vowels fall below 1 kHz, this could have led to disproportionately higher RT for vowel sounds, leading to greater perceptual smearing than the mean RT value would lead one to expect.

Regarding the tests used, our updating task was visually administered rather than in the auditory domain, and this crossing of domains could have had some bearing on the results. Numbers rather than words were used because they were language non-specific and we assumed this controlled for various problems when using word updating (word frequency for example). However, it was not taken into consideration that read numbers are processed phonologically on a sub-vocal level during reading, and that the sub-vocal language triggered would be the reader's native language (L1). Therefore, this was essentially a measure of L1 updating which was then contrasted with L2 listening comprehension. Additionally, as an audiogram was not taken, the role of pure tone thresholds could not be evaluated and there is a small chance that participants which reported no diagnosed hearing impairment might have been mildly impaired.

Conclusion

This study examined differences in listening comprehension under virtually simulated short and long reverberation times (RT) of 0.3 and 0.9 s respectively, conditions which aimed to reproduce poor versus good classroom environments for listening comprehension. Groups were split according to high and low second language (L2) vocabulary and updating ability. Results indicated that the High Vocab group had higher listening comprehension scores than the Low Vocab group under both acoustic conditions, but that this advantage was especially noticeable under lower RT conditions. Under both RT conditions, the Low Vocab group performed equally poorly on the listening comprehension assessment, showing no additional advantage of good acoustics. Updating was not measurably related to listening comprehension under different RT conditions. In conclusion, the results of this study demonstrate an additional benefit for listening comprehension under optimal acoustic conditions where the target speech is clear, but this benefit is in proportion with the level of L2 proficiency. Future research will be directed towards establishing why lower L2 vocabulary groups are unable to take advantage of this effect, and will particularly examine the role of L1 suppression during L2 listening tasks in good and bad acoustic conditions.

Acknowledgement

Special thanks to Sebastian Arnstrom for assisting with data collection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement FP7-607139 (iCARE).

ORCID

Janina Fels http://orcid.org/0000-0002-8694-7750

References

Bradley, J. S. (1986). Speech intelligibility studies in classrooms. *The Journal of the Acoustical Society of America*, *80*(3), 846–854.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, *112*(1), 272–284.

Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology*, *91*(1), 45–66.

Carretti, B., Cornoldi, C., & Pelegrina, S. L. (2007). Which factors influence number updating in working memory? The effects of size distance and suppression. *British Journal of Psychology*, *98*(1), 45–60.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, *25*, 1–18.

Daneman, M., & Merikle, P. M. (1996). Working memory and comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*, 422–433.

DIN. (2010). 45631/A1:2010-03 – Calculation of loudness level and loudness from the sound spectrum – Zwicker method – Amendment 1: Calculation of the loudness of time-variant sound.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Engle, R. W., Cantor, J., & Carollo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*, 972–992.

Engle, R. W., Nations, K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? *Journal of Educational Psychology*, *82*, 799–804.

Erickson, G. (2009). *National assessment of foreign languages in Sweden*. Gothenburg: University of Gothenburg.

Everest, F. A., & Pohlman, K. C. (2015). *Master handbook of acoustics*. New York: McGraw Hill.

Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal*, *57*(4), 325–334.

Field, J. (2004). An insight into listeners' problems: Too much bottom-up or too much top-down? *System*, *32*(3), 363–377.

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition*, *9*(2), 153–166.

Hurtig, A., Keus van de Poll, M., Pekkola, E. P., Hygge, S., Ljung, R., & Sörqvist, P. (2016). 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.

Hygge, S. (2003). Classroom experiments on the effects of different noise sources and sound levels on long-term recall and recognition in children. *Applied Cognitive Psychology*, *17*, 895–914.

Hygge, S., Evans, G. W., & Bullinger, M. (2002). A prospective study of some effects of aircraft noise on cognitive performance in schoolchildren. *Psychological Science*, *13*, 469–474.

ISO. (1975). 532:1975(en) Acoustics – Method for calculating loudness level.

ISO. (2008). 3382-2:2008 Acoustics – Measurement of room acoustic parameters – Part 2: Reverberation time in ordinary rooms.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston naming test*. Austin: Pro-ed.

Kilman, L., Zekveld, A., Hällgren, M., & Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology*, *5*, 651.

Kjellberg, A. (2004). Effects of reverberation time on the cognitive load in speech communication: Theoretical considerations. *Noise and Health*, *7*(25), 11.

Klatte, M., Bergstroem, K., & Lachmann, T. (2013). Does noise affect learning? A short review of noise effects on cognitive performance. *Frontiers in Psychology*, *4*, 578.

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.

Kuttruff, H. (2016). *Room acoustics*. Boca Raton: CRC Press.

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, *52*(11-12), 864–886.

Ljung, R. (2010). Room acoustics and cognitive load when listening to speech (Doctoral dissertation). Luleå tekniska universitet.

Ljung, R., & Kjellberg, A. (2009). Long reverberation time decreases recall of spoken information. *Building Acoustics*, *16*(4), 301–311.

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*(3), 257–265.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Moore, B. C. (2012). *An introduction to the psychology of hearing*. Boston: Brill.

Muijselaar, M. M., & de Jong, P. F. (2015). The effects of updating ability and knowledge of reading strategies on reading comprehension. *Learning and Individual Differences*, *43*, 111–117.

Náběelek, A. K., & Donahue, A. M. (1984). Perception of consonants in reverberation by native and non-native listeners. *The Journal of the Acoustical Society of America*, *75*(2), 632–634.

Nilsson, S. (2016). Engelska 6 – Exempel på uppgiftstype [English 6 – Examples of task types]. Göteborg University. Retrieved from http://nafs.gu.se/prov_engelska/exempel_provuppgifter/engelska_6_exempeluppg.

Ormrod, E., & Cochran, K. F. (1988). Relationship of verbal ability and working memory to spelling achievement and learning to spell. *Reading Research & Institution*, *28*(1), 33–43.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory & Cognition*, *29*(2), 344–354.

Pelzer, S., Aspöck, L., Schröder, D., & Vorländer, M. (2014). Interactive real-time simulation and auralization for modifiable rooms. *Building Acoustics*, *21*(1), 65–73.

Peng, Z. E., & Wang, L. M. (2016). Effects of noise, reverberation and foreign accent on native and non-native listeners’ performance of English speech comprehension. *The Journal of the Acoustical Society of America*, *139*(5), 2772–2783.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., … Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, 31.

Rudner, M., & Lunner, T. (2014). Cognitive spare capacity and speech communication: A narrative overview. *BioMed Research International*, *2014*, 1–10.

Scheuerle, J. (2000). Hearing and aging. *Educational Gerontology*, *26*(3), 237–247.

Schmitz, A. (1995). Ein neues digitales Kunstkopfmeßsystem [A new artificial head measuring system]. *Acta Acustica United with Acustica*, *81*(4), 416–420.

Schröder, D. (2011). *Physically based real-time auralization of interactive virtual environments* (Vol. 11). Logos Verlag Berlin GmbH.

Sörqvist, P., Halin, N., & Hygge, S. (2010). Individual differences in susceptibility to the effects of speech on reading comprehension. *Applied Cognitive Psychology*, *24*(1), 67–76.

Sörqvist, P., Hurtig, A., Ljung, R., & Rönnberg, J. (2014). High second-language proficiency protects against the effects of reverberation on listening comprehension. *Scandinavian Journal of Psychology*, *55*(2), 91–96.

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, *31*(4), 577–607.

Stansfeld, S. A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., & Haines, M. (2005). Aircraft and road traffic noise and children’s cognition and health: A cross-national study. *The Lancet*, *365*(9475), 1942–1949.

Takata, Y., & Náběelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *The Journal of the Acoustical Society of America*, *88*(2), 663–666.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Vorländer, M. (2007). *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Berlin: Springer Science & Business Media.

Wiley, J., & Jarosz, A. F. (2012). How working memory capacity affects problem solving. *Psychology of Learning and Motivation*, *56*, 185–227.

Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Language and Speech*, *37*(3), 221–235.