



The Influence of the Type of Background Noise on Perceptual Learning of Speech in Noise

Liping Zhang^{1,2*}, Friederike Schlaghecken³, James Harte^{2,4} and Katherine L. Roberts^{3,5}

¹ Department of Otolaryngology-Head and Neck Surgery, Shandong Provincial ENT Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China, ² Warwick Manufacturing Group, University of Warwick, Coventry, United Kingdom,

³ Department of Psychology, University of Warwick, Coventry, United Kingdom, ⁴ Interacoustics Research Unit, Technical University of Denmark, Lyngby, Denmark, ⁵ Department of Psychology, Nottingham Trent University, Nottingham, United Kingdom

Objectives: Auditory perceptual learning studies tend to focus on the nature of the target stimuli. However, features of the background noise can also have a significant impact on the amount of benefit that participants obtain from training. This study explores whether perceptual learning of speech in background babble noise generalizes to other, real-life environmental background noises (car and rain), and if the benefits are sustained over time.

Design: Normal-hearing native English speakers were randomly assigned to a training ($n = 12$) or control group ($n = 12$). Both groups completed a pre- and post-test session in which they identified Bamford-Kowal-Bench (BKB) target words in babble, car, or rain noise. The training group completed speech-in-babble noise training on three consecutive days between the pre- and post-tests. A follow up session was conducted between 8 and 18 weeks after the post-test session (training group: $n = 9$; control group: $n = 7$).

Results: Participants who received training had significantly higher post-test word identification accuracy than control participants for all three types of noise, although benefits were greatest for the babble noise condition and weaker for the car- and rain-noise conditions. Both training and control groups maintained their pre- to post-test improvement over a period of several weeks for speech in babble noise, but returned to pre-test accuracy for speech in car and rain noise.

Conclusion: The findings show that training benefits can show some generalization from speech-in-babble noise to speech in other types of environmental noise. Both groups sustained their learning over a period of several weeks for speech-in-babble noise. As the control group received equal exposure to all three noise types, the sustained learning with babble noise, but not other noises, implies that a structural feature of babble noise was conducive to the sustained improvement. These findings emphasize the importance of considering the background noise as well as the target stimuli in auditory perceptual learning studies.

Keywords: perceptual learning, auditory training, speech in noise, generalization, babble noise

OPEN ACCESS

Edited by:

Benedikt Zoefel,
UMR 5549 Centre de Recherche
Cerveau et Cognition (CerCo), France

Reviewed by:

Vasiliki Maria Iliadou,
Aristotle University of Thessaloniki,
Greece

Michel Hoen,
Oticon Medical, Sweden

*Correspondence:

Liping Zhang
Liping_zhang0814@163.com

Specialty section:

This article was submitted to
Perception Science,
a section of the journal
Frontiers in Neuroscience

Received: 25 December 2020

Accepted: 06 April 2021

Published: 03 May 2021

Citation:

Zhang L, Schlaghecken F, Harte J
and Roberts KL (2021) The Influence
of the Type of Background Noise on
Perceptual Learning of Speech
in Noise. *Front. Neurosci.* 15:646137.
doi: 10.3389/fnins.2021.646137

INTRODUCTION

Understanding speech in noise can present a considerable challenge, even for listeners with good hearing. This challenge is exacerbated for listeners with hearing impairment, and particularly for older adults who are affected by age-related declines in both hearing and cognition (Roberts and Allen, 2016). For listeners who find it particularly difficult to understand speech amid background noise, such as those with hearing impairment or auditory processing disorder (APD) (Iliadou et al., 2017), auditory perceptual training has the potential to improve speech-in-noise comprehension (Sweetow and Palmer, 2005; Anderson et al., 2013; Weihing et al., 2015). However, auditory training does not always result in robust benefits, or generalize to untrained tasks or stimuli (Henshaw and Ferguson, 2013; Karawani et al., 2016), resulting in a need to optimize auditory training paradigms. In this study, we ask whether the type of background noise used in training affects the extent to which any improvements in speech-in-noise identification generalize to other background noises, and are sustained over time.

Improvements in auditory training paradigms could prove highly beneficial to both normal-hearing and hearing-impaired listeners. Normal-hearing listeners can experience difficulty hearing in noisy environments, and while speech recognition by people with hearing aids and cochlear implants has improved significantly over the past several years due to technological improvements, the ability of most hearing-impaired people to understand speech in noisy environments is still quite poor (Dorman and Wilson, 2004; Ricketts and Hornsby, 2005). Due to the plasticity of the auditory system (Irvine, 2018), auditory perceptual training has the potential to improve the listening performance of all listeners, and to help hearing-aid and cochlear-implant users make better use of their prosthetic device (Sweetow and Sabes, 2006; Boothroyd, 2007; Moore and Shannon, 2009). The present study investigates the effect of auditory perceptual training on normal-hearing listeners, to enable an evaluation of the effect of background noise on training outcomes in the absence of potentially confounding variables related to hearing impairment.

Auditory perceptual learning is defined as an improvement in the ability to detect, discriminate, or group sounds and speech information (Goldstone, 1998; Halliday et al., 2008). While many studies show substantially improved auditory perception following training, practical benefits to listeners will arise only if that training benefit generalizes to untrained tasks and stimuli, and is sustained over time. Generalization occurs when training on one auditory task leads to improvement on a novel auditory task (Grimault et al., 2003), and when training with one set of stimuli leads to improvements in perception of untrained, novel stimuli (Tremblay and Kraus, 2002; Davis et al., 2005; Wong and Perrachione, 2007).

The extent to which training generalizes is likely to depend on whether the untrained stimuli or task include the same internal perceptual and cognitive noise (Amitay et al., 2014) and/or decision-making processes (Jones et al., 2015; Karawani et al., 2016) as those in the training task. Many studies find that generalization of learning to untrained tasks and/or stimuli is not

robust (Henshaw and Ferguson, 2013). For example, Karawani et al. (2016) found that both normal-hearing and hearing-impaired listeners benefited from speech-in-noise training, but that there was little generalization to untrained speech and non-speech sounds. An improved understanding of how and when auditory perceptual training generalizes in normal-hearing participants is needed to help devise better training programs for people with hearing impairment (Loebach et al., 2009).

Auditory perceptual learning studies frequently study identification of speech or other sounds amid background noise (e.g., babble or speech-shaped noise), because the ability to detect speech signals in a noisy environment is critical in people's daily communication. When evaluating whether benefits of perceptual training generalize to other stimuli, these studies tend to focus on generalization to novel target speech or other sounds (Clarke and Garrett, 2004; Hervais-Adelman et al., 2011). These studies have demonstrated that there is greater generalization when training and test materials are similar (Hirata, 2004), and that high-variability training with a number of talkers leads to increased generalization to novel talkers compared with single-talker training (Clopper and Pisoni, 2004; Stacey and Summerfield, 2007; Casserly and Barney, 2017) (see Samuel and Kraljic, 2009 for a review). However, the background noise is also constantly changing in the real world, and so it is equally important that training generalizes to other types of background noise, particularly real-world environmental noise. The novelty of this study is that we aim to determine whether changing the background noise affects the amount of benefit that is maintained from auditory perceptual training.

Background noise can interfere with speech understanding through energetic masking, where the background noise has energy in the same frequency region as the speech signal, thus preventing the speech signal from being perceived. When the background noise fluctuates, as is likely with real-world environmental sounds and competing speech, the listener is afforded opportunities to "listen in the dips," or "glimpse" the speech signal (Howard-Jones and Rosen, 1993). Alternatively, background noise can produce informational masking that results from difficulties with auditory scene analysis (Bregman, 1990), particularly when the listener has difficulty segregating target speech from the background masker (Brungart, 2001) due to failures of object formation or selection (Shinn-Cunningham, 2008).

Due to the differing effects of energetic and informational masking, the amount and type of benefit that participants receive from perceptual training may differ depending on the type of background noise. Steady-state noise, such as speech-shaped noise, is likely to provide consistent energetic masking but little informational masking. On the other hand, the temporal variation in babble-noise will afford more opportunities for glimpsing, but increased informational masking if words are partially audible. Correspondingly, training strategies that improve glimpsing or segregation may be more useful for speech presented in babble than for speech presented in steady-state noise. Van Engen (2012) trained participants on English sentence recognition in three different background noise conditions: speech-shaped noise, Mandarin babble, and English babble. She

found that English sentence recognition was much better with babble background noise (both English and Mandarin) than with speech-shaped noise. The results suggest that to improve people's speech perception in speech-in-speech environments, it is better to train them with background noise that is structured in a similar way to speech than with noise that has relatively consistent amplitude over time. Similarly, Green et al. (2019) found that training with speech-in-babble-noise improved cochlear-implant users' perception of sentences in babble noise, but did not result in improved perception of phonemes in speech-shaped noise. These studies suggest that speech-like noise may enable listeners to develop strategies that allow them to "listen in the dips," where energetic masking is reduced. This benefit of dip-listening appears to offset any costs associated with increased informational masking for babble noise relative to steady-state noise.

Other features of the background noise can also change the amount of perceptual learning that is obtained. Felty et al. (2009) demonstrated that listeners showed greater improvement in word recognition performance when the same sample of background babble-noise was presented on each trial, compared with when different noise samples were presented on each trial. Similar results were found in a visual texture segmentation task in which a background mask was either consistent from trial-to-trial or varied on each trial (Schubö et al., 2001). In contrast, a previous perceptual learning study with vowel-consonant-vowel (VCV) stimuli found that consonant identification improved more when stimuli were presented against a random-noise background than against a fixed-noise background (Zhang et al., 2014), indicating that the effects of background noise on perceptual learning differ with different types of target stimuli (very short VCV targets contrasting with longer word stimuli).

During everyday listening, background noise is likely to include environmental sounds (e.g., washing machine, traffic) as well as speech sounds (e.g., television, other people's conversations). To date, though, studies that have looked at perceptual training with environmental stimuli have included environmental sounds as the target rather than background (e.g., footsteps, slamming door, air conditioner, dishwasher; see Burkholder, 2005; Reed and Delhorne, 2005; Kidd et al., 2007). A study by Loebach and Pisoni (2008) trained normal-hearing participants with a simulation of a cochlear implant. Each group was trained with one type of auditory stimulus (from a choice of words, sentences, and environmental sounds) and then later tested on all types of auditory stimuli. The results showed that all groups obtained significant improvement in the specific stimuli they were trained on. However, while perceptual learning did not transfer from training on speech to the recognition of environmental sounds, it did transfer from training on environmental sounds to both untrained environmental sounds and speech sounds. This finding suggests that there are differences in how well perceptual learning transfers from speech to environmental stimuli, and vice versa.

The present study will investigate the perceptual learning benefit obtained when participants are trained to identify Bamford-Kowal-Bench (BKB) sentences presented in background noise. Because auditory perceptual learning

studies have demonstrated that word-identification training outcomes are better with word and sentence stimuli than with nonsense syllables phonemes (Stacey and Summerfield, 2008), BKB sentences were used as the target stimuli. The study has two aims: one is to investigate whether perceptual learning generalizes from the background noise used in training (babble noise) to other real-life environmental background noises (the sound of a car passing by, and rain noise), the other is to explore whether any generalization of perceptual learning is sustained over a period of several weeks. The environmental sounds were selected as examples that commonly occur in everyday life, typically persist for long enough to mask a short sentence, and do not include speech (e.g., not background television noise or music with lyrics).

TEST METHODS

Participants

Twenty-four normal-hearing native English speakers (8 males and 16 females, aged 18–33 years) participated in this study. None had previously participated in psychoacoustic experiments. Pure-tone audiometry thresholds were less than 20 dB HL at all frequencies between 250 and 8,000 Hz (BSA, 2011). Participants were not tested for auditory processing disorder (APD; Iliadou et al., 2017), but all confirmed that they were not aware of any problems with their hearing and that they had not been exposed to loud noises in the past 24 h. Participants in the training and control groups did not differ in terms of age [mean age 23.42 in the control group and 23.00 in the experimental group, $t(22) = -0.24, p = 0.81$]. Participants were all volunteers recruited from the student and staff population of the University of Warwick. All gave written informed consent before participating. Ethics approval for this study was given by the Biomedical and Scientific Research Ethics Committee (BSREC) of the University of Warwick.

Test Stimuli

Bamford-Kowal-Bench (BKB; Bench et al., 1979) sentences recorded by a female British speaker were used as speech material. The speech material consisted of 21 different lists, with each list containing 16 sentences and a total of 50 target words. The sentences were centrally embedded in 2 s of background noise. The babble noise comprised 2 s of 8-speaker babble noise with four female and four male British-English speakers (Verschuur et al., 2013). Recordings of rain and a car passing by were downloaded from <https://www.soundsnap.com>. The signal-to-noise ratio for each sentence was determined by comparing the root mean square average amplitude of the signal file with the background noise file (the portion that actually overlapped with the sentence). The root mean square intensity was normalized to the same fixed value for all background noise.

All sounds were presented through Sennheiser HD 580 headphones. Calibration was carried out prior to the main test. An IEC 711 acoustic coupler and a precision microphone were used to calibrate the output. Then the maximum sound pressure levels from the PC were controlled to ensure that output from

the software (MATLAB) was within the exposure action value (65 dB SPL for the speech plus noise stimuli). Sampling rate for all signals was 44.1 kHz. Signal-to-noise ratios (SNRs) were fixed for a given noise type, but varied for each noise condition (babble noise -20 dB, car noise -12 dB, rain noise -15 dB). These SNR levels were based on a pilot study ($n = 8$), in which they gave approximately 50% correct target-word identification with each of the background noises. **Figure 1** shows the waveforms for an example sentence “The clown had a funny face” with three kinds of background noise.

Experiment Procedure

All tests were carried out in a sound-attenuating room. Participants were randomly assigned to either a control ($n = 12$) or training ($n = 12$) group (see **Figure 2**). Before the test, a pure tone audiogram was carried out to confirm that the person qualified to participate. Participants then received written instructions and were presented with one example sentence without background noise to familiarize them with the stimuli. The participant was asked to repeat the sentence. Next, babble, car, and rain background noise samples were presented separately.

Participants were informed that during the experimental trials, the speech sounds would be softer than the background noise. Both the training and the control group completed a pre-training test session (“pre-test”) and a post-training test session (“post-test”) lasting approximately 5 min each. The training group attended three consecutive daily 30 min training sessions with BKB sentences presented amid babble noise between the pre- and post-test sessions. The control group only attended for the pre-test and post-test sessions. There was a 1 day fixed time interval between pre- and post-test session for the control group. Participants were asked to repeat each sentence and were encouraged to guess even if the sentences they repeated would result in a nonsense or incomplete sentence. All the test sessions were conducted over three consecutive days (Day 1: pre-test and training session one; Day 2: training session two; Day 3: training session three and post-test). The pre- and post-test sessions included one 16-item BKB sentence list with babble noise, one with car noise, and one with rain noise. The order of the three noise conditions was randomized in the pre- and post-tests but the BKB sentence list was the same across participants. Different BKB sentence lists were used for the pre-test (lists 1–3), training (lists 4–15), and the post-test (lists 16–18) sessions. Sessions were kept intentionally short to ensure that participant motivation remained high. The number of sessions was based on a longer study with vowel-consonant-vowel stimuli (Zhang et al., 2014) in which participants reached an asymptote after three training sessions. No feedback was given in any session.

A final follow-up test session was carried out to investigate whether training effects and generalization to other background noises could be maintained over time. Participants were recalled between 8 and 18 weeks after the post-test session [with no significant difference in the time gap between the training and control groups, $t(14) < 1$]. The procedure was the same as the pre- and post-test sessions, though a new set of BKB sentence lists were used (lists 19–21). As some of the participants had already

left the university, not all the listeners attended the follow-up study. Nine training group and seven control group participants attended the follow-up test session.

TEST RESULTS

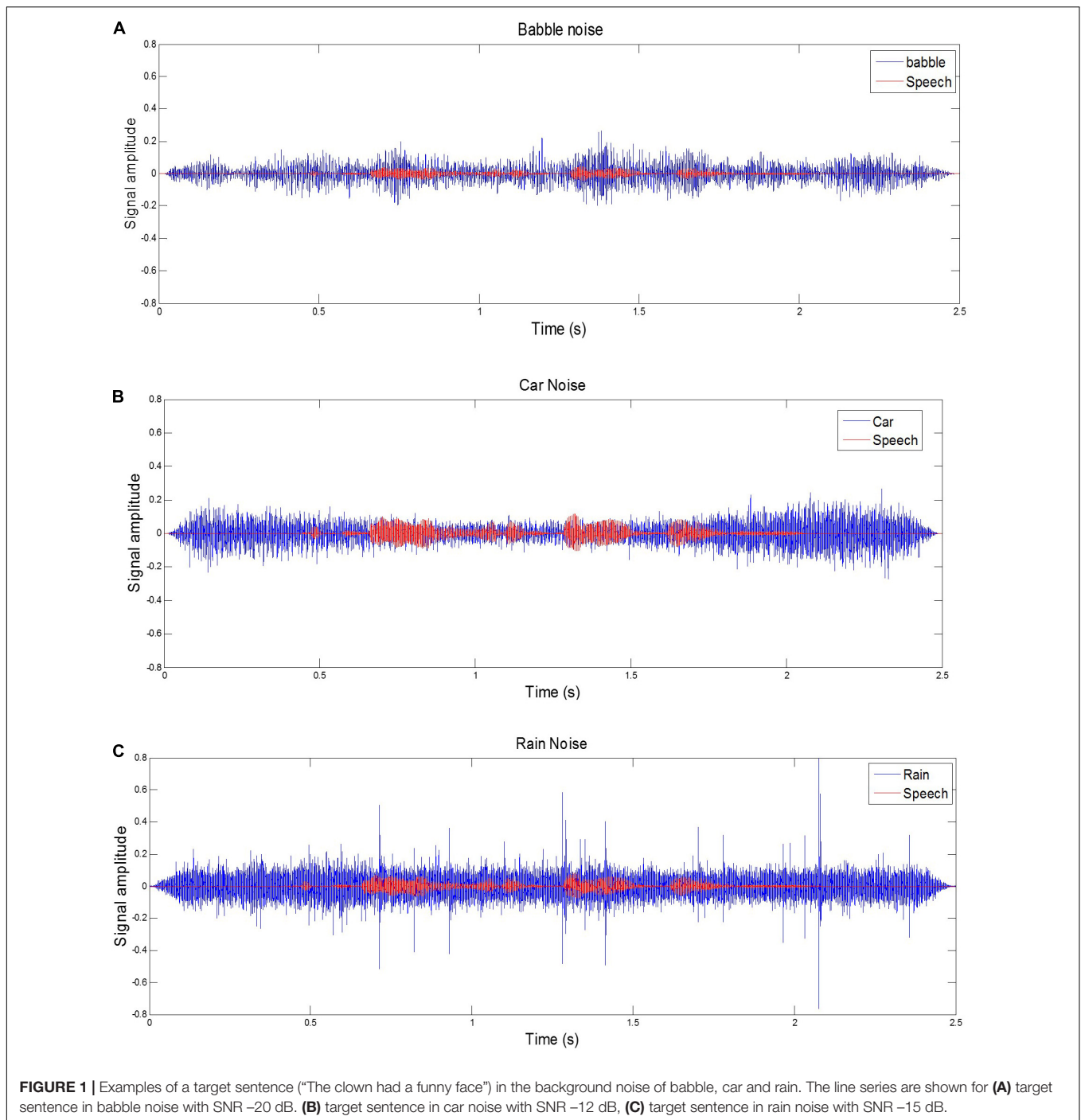
Pre- and Post-test Results

We calculated the number of BKB keywords that were correctly identified for each test, out of a maximum of 50. For the statistical analyses (but not graphs or reported values), the percentage correct was converted using a rationalized arcsine transform (Studebaker, 1985), using the formula for a small number of trials, to produce rationalized arcsine units (RAUs). This transformation corrects for deviations from normality while keeping the values numerically close to the originally percentages, for ease of interpretation. Where sphericity could not be assumed a Greenhouse-Geisser correction has been applied, and is evident from non-integer degrees of freedom.

Figure 3 shows word-identification accuracy in the different conditions. A 2 (group) \times 3 (noise condition) analysis of covariance (ANCOVA) was used to investigate whether post-test word identification was better for the trained or control group, and whether this differed across noise conditions (babble, car, rain). Pre-test word identification was used as a covariate to control for baseline differences in pre-test performance. The ANCOVA revealed that the training group had significantly higher post-test accuracy than the control group [group, $F(1, 19) = 38.56$, $p < 0.001$, $\eta^2 = 0.67$]. There was no significant difference between the three background noise conditions [noise, $F(1.55, 29.47) = 2.06$, $p = 0.15$, $\eta^2 = 0.10$], but there was a significant interaction between the group and noise conditions [noise \times group interaction, $F(1.55, 29.47) = 5.55$, $p = 0.014$, $\eta^2 = 0.23$].

Pairwise comparisons with a Bonferroni correction for multiple comparisons (critical $p = 0.0167$) showed that the training group had significantly higher post-test scores than the control group in both the trained (babble) and untrained (car and rain) conditions, when controlling for pre-test performance (training post-test accuracy – control post-test accuracy in babble noise = 15.33, $p < 0.001$; car noise = 9.83, $p < 0.001$; rain noise = 6.33, $p = 0.015$). To probe whether training led to a greater improvement in the trained background noise than other noise types, we conducted *post hoc* (group \times noise) ANCOVAs that included one pair of noise conditions at a time, still controlling for pre-test performance. A significant interaction between group and noise condition would indicate that the amount by which the training group outscored the control group was significantly greater for one type of noise than the other. The training group outscored the control group by more in the babble noise condition than the car noise [$F(1, 20) = 8.11$, $p = 0.010$, $\eta^2 = 0.29$], or rain noise conditions [$F(1, 20) = 6.73$, $p = 0.017$, $\eta^2 = 0.25$], but there was no difference in the pre- to post-test improvement between the car and rain noise conditions [$F(1, 20) = 1.02$, $p = 0.33$, $\eta^2 = 0.048$].

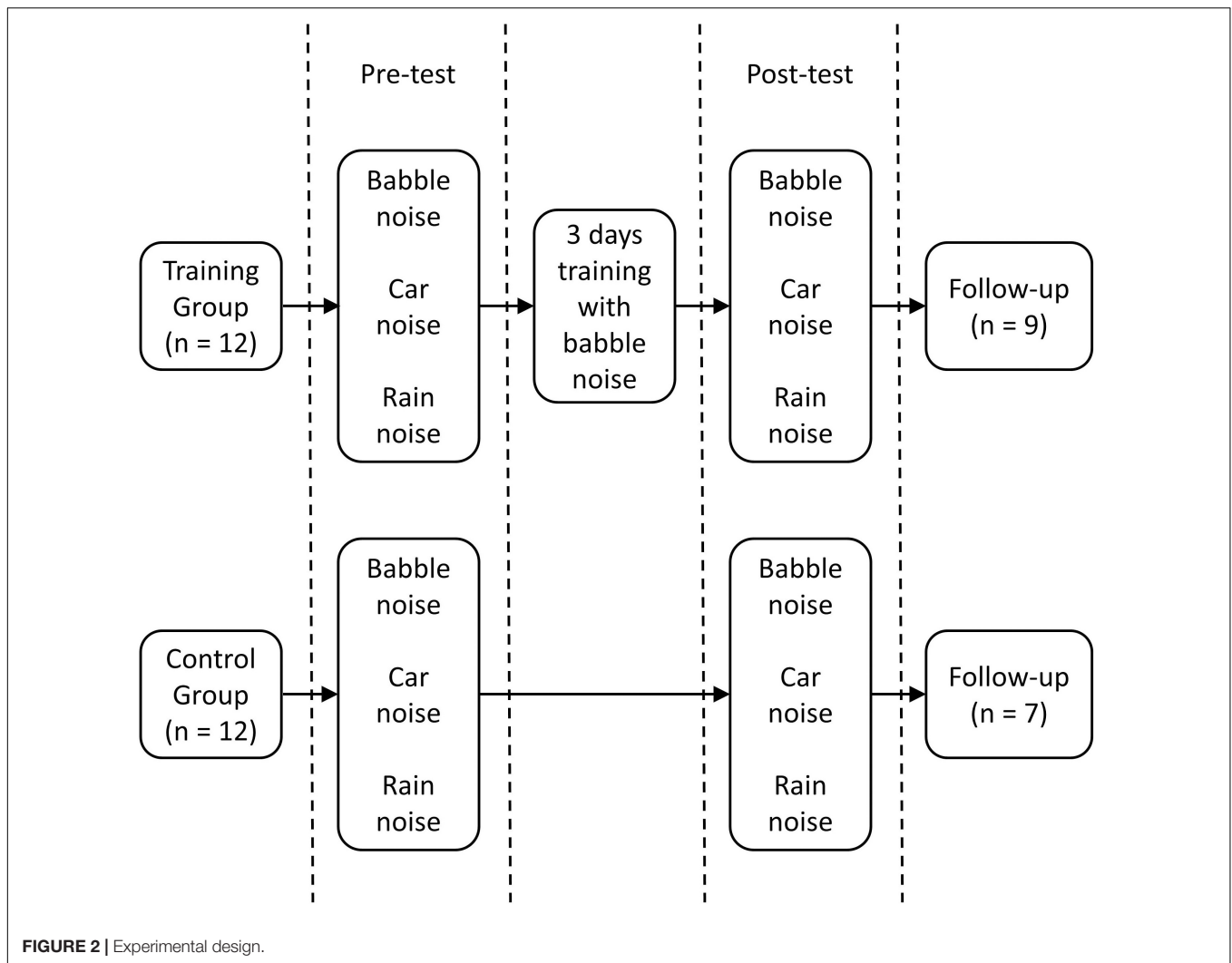
For transparency, a two-way (group \times noise) analysis of variance (ANOVA) was used to confirm that the improvement



in RAUs from pre- to post-test was greater for the training than for the control group, and differed across the noise conditions. A main effect of group [$F_{(1, 22)} = 46.17, p < 0.001, \eta^2 = 0.68$] confirmed that from pre- to post-test, the training group showed greater improvement in word identification accuracy than the control group. There was also a significant difference between the three noise conditions [noise: $F_{(2, 44)} = 42.98, p < 0.001, \eta^2 = 0.66$], and a significant interaction between group and noise condition [group \times noise: $F_{(2, 44)} = 4.21, p = 0.02, \eta^2 = 0.16$].

Summary of Pre- to Post-test Results

The training group showed significantly better post-test word identification accuracy than the control group. This was true for both the trained (babble) noise and to a lesser extent the untrained (car and rain) background noises, showing some generalization of learning to untrained background noises. The training benefit was larger for the trained background noise, with the training group outscoring the control group by a greater amount in the babble noise condition than the car or rain noise



conditions, with no difference in improvement between the car and rain noise conditions. This pattern of results was confirmed when the improvement from pre- to post-test was analyzed, with the training group showing greater improvement than the control group, with the greatest effect size for speech in babble noise.

Individual Training Performance Across Sessions

Individual performance across training sessions, for the babble-noise condition only, was analyzed to confirm that the improvement in word identification had reached an asymptote. **Figure 4** shows BKB word identification RAUs for each participant in the training group, in each test and training session. A one-way repeated-measures ANOVA showed that for the babble noise condition, there was an overall improvement across the pre-test, training, and post-test sessions [$F_{(2,4, 26.4)} = 127.83, p < 0.001, \eta^2 = 0.921$]. *Post hoc* t tests with a Bonferroni correction for multiple comparisons demonstrated an improvement from each session to the next (pre- to Training Session 1, Training Session 1 to Training Session 2, and Training

Session 2 to Training Session 3 ($t_{11} = 11.42, 5.39, \text{ and } 10.70$, respectively, all $ps < 0.001$), with the exception of the final comparison between Training Session 3 and the post-training test ($t_{11} = 1.71, p = 0.12$). In other words, no further improvement was found between the final training session and the post-training test.

Results for the Follow-Up Test

Figure 5 shows word identification RAUs for the 16 participants who took part in the follow-up test. Those who participated in the follow-up session did not have significantly different pre- or post-test scores to those who did not, and did not differ in their change in accuracy from pre- to post-test in any of the background noise conditions (all $ps > 0.1$). A mixed ANOVA was used to analyze the word identification accuracy for both groups (training and control), in the different noise conditions (babble, car and rain), across the three test sessions (pre-test, post-test and follow-up). Overall, performance was better for the training than for the control group, resulting in a significant main effect of group [$F_{(1, 14)} = 4.75, p = 0.047, \eta^2 = 0.253$], and better in the post-test than

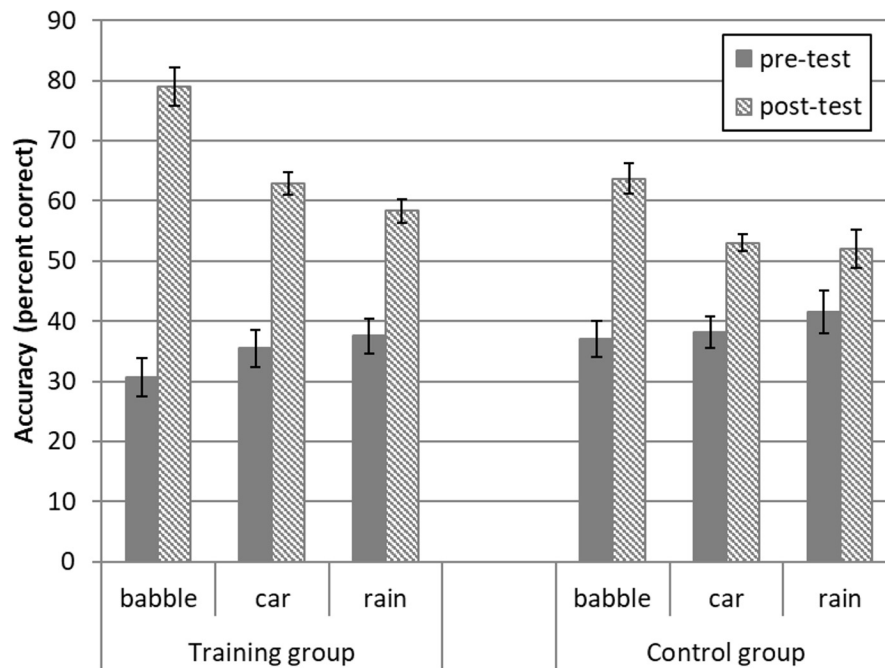


FIGURE 3 | Word identification accuracy (percent correct) for the training ($n = 12$) and control ($n = 12$) groups with three different background noises. Error bars reflect \pm one standard error.

in the pre-test or follow-up session, reflected in a significant main effect of time [$F_{(2, 28)} = 127.31, p < 0.001, \eta^2 = 0.901$]. Follow-up pairwise comparisons showed that performance improved from pre-test to post-test, then declined from post-test to follow-up, but remained better than at pre-test (all $ps < 0.001$).

Word identification accuracy differed significantly across the noise conditions [$F_{(1.45, 30.30)} = 8.24, p < 0.01, \eta^2 = 0.371$], and follow-up pairwise comparisons demonstrated that word identification accuracy was significantly better in the babble noise condition than in the car or rain noise conditions (both $ps < 0.05$), whereas the latter two did not differ substantially from each other ($p = 0.056$). There was also a significant interaction between time and noise condition [$F_{(3.40, 47.57)} = 29.57, p < 0.001, \eta^2 = 0.679$].

Post hoc t tests using Bonferroni correction for multiple comparisons ($p = 0.05/9$, i.e., 0.0056) showed that post-test scores were significantly higher than pre-test scores for all types of background noise (all $ps < 0.001$). For the babble-noise condition, test scores at follow-up remained significantly higher than at pre-test ($p < 0.001$), and were not significantly different from post-test ($p > 0.0056$), suggesting that the perceptual learning benefit was maintained over time. In contrast, for the car and rain noise conditions, there was a significant decrease in word identification accuracy from post-test to follow-up (both $ps < 0.001$), and scores at follow-up were not significantly different from those at pre-test ($p > 0.0056$). The interaction between test time and group was significant [$F_{(2, 28)} = 13.39, p < 0.001, \eta^2 = 0.489$], but there was no significant interaction between test time, group, and noise condition [$F_{(3.40, 47.57)} = 0.66, p = 0.621, \eta^2 = 0.045$].

Summary of Results for the Follow-Up Test

The training group showed significantly better word identification than the control group. Overall, word identification accuracy was better at post-test than at follow-up, and better at follow-up than at pre-test, but this differed across noise conditions. In the babble noise condition, participants maintained their improved performance from post-test to follow-up, but in the car and rain noise conditions performance returned to pre-test levels. The effects of time and noise condition did not interact with group, indicating that the same pattern of results was found for both the training and control groups.

DISCUSSION

The present findings demonstrate that training participants to identify BKB keywords amid babble noise can improve BKB word-identification accuracy in a range of background noises. Participants who were trained to identify BKB sentences amid background babble noise had higher post-test accuracy than the control group, not only for BKB sentences presented amid babble noise, but also, to a lesser extent, for BKB sentences presented amid car and rain noise. In other words, the benefits of training participants to understand sentences amid babble noise showed some generalization to sentences presented amid car and rain noise. Nonetheless, participants showed greater improvement in the babble-noise condition than in the car- or rain-noise conditions, as would be expected given that the training sessions

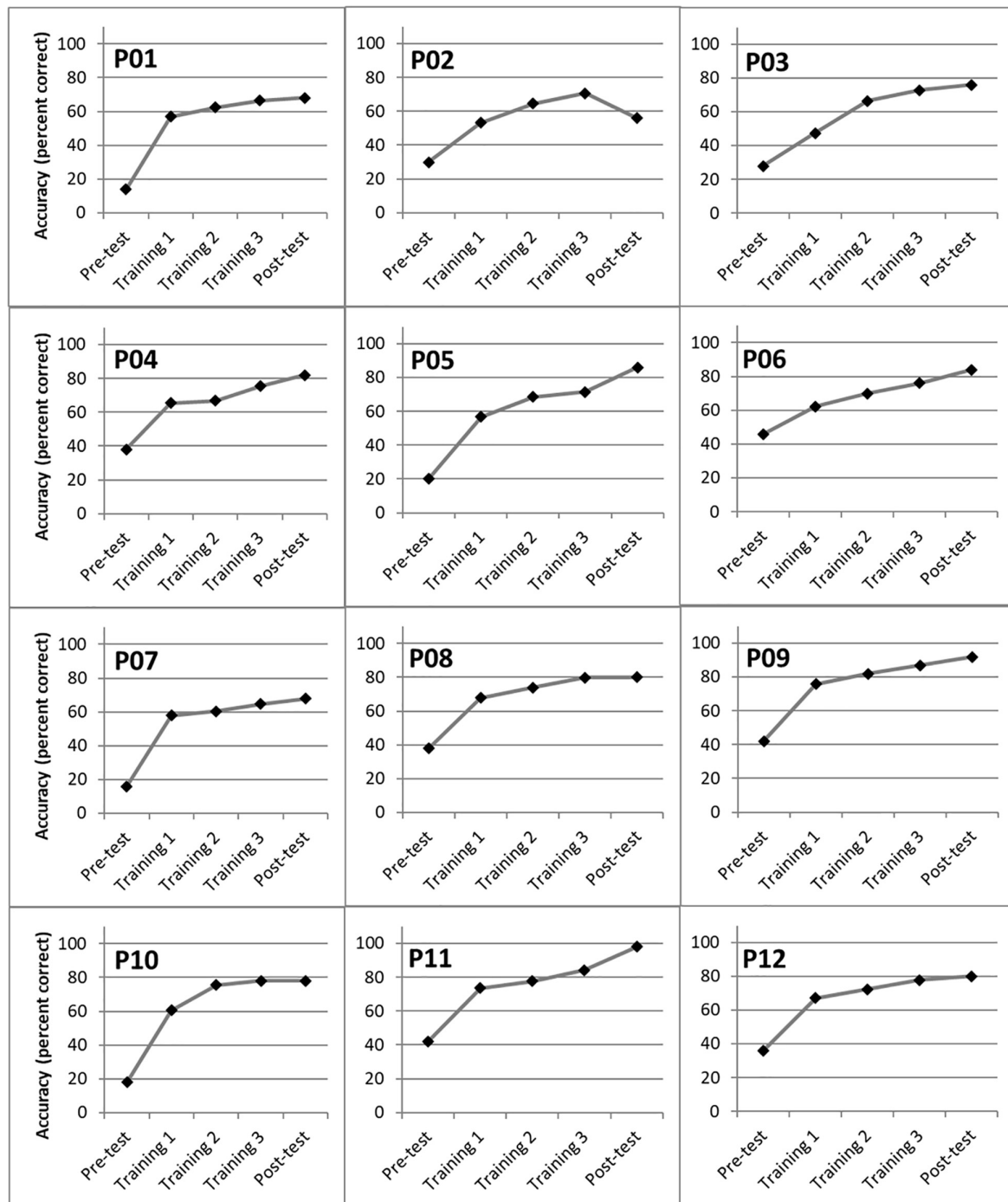


FIGURE 4 | Word identification accuracy for individual members of the training group at the pre-training test ("pre-test"), Training Session 1, Training Session 2, Training Session 3, and at the post-training test ("post-test"), for the babble-noise conditions only.

involved identical stimuli to the post-test babble noise condition (Morris et al., 1977; Roediger et al., 1989; Borrie et al., 2012).

Perceptual training can help listeners to make better lexical judgments about stimuli, process sound information to a higher cognitive order, and reduce participants' attention to lower-order acoustic features (Loebach and Pisoni, 2008). All of these skills

should generalize to other forms of background noise, resulting in the generalization of learning from babble noise to car and rain noise seen in the present study. In addition, earlier perceptual speech training studies with synthetic speech in quiet suggest that auditory perceptual training may adjust the auditory system by increasing awareness of informative phonetic cues, decreasing

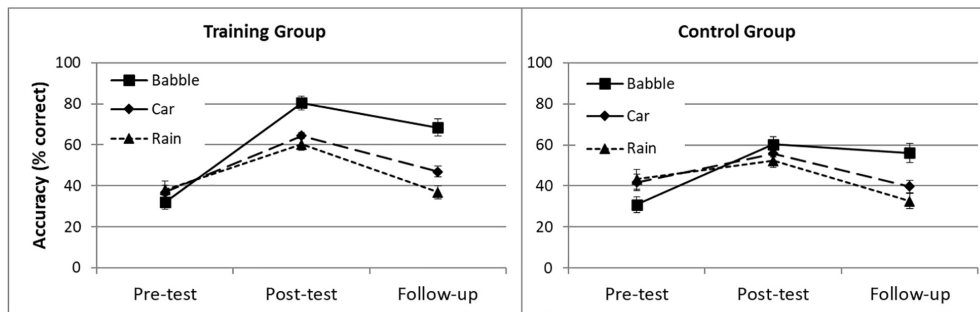


FIGURE 5 | Word identification accuracy for the training ($n = 9$) and control group ($n = 7$) from three test sessions (pre-test, post-test, and follow-up test) with three different background noises (babble, car and rain noise).

the influence of less useful stimuli, or both (Schwab et al., 1985; Francis et al., 2007; Francis and Nusbaum, 2009).

A subset of participants was followed up to determine if training benefits persisted over time. Improvements from pre- to post-test were sustained over a period of several weeks for words presented in babble noise, but not for words presented in car or rain noise, for which performance returned to pre-test levels. The sustained improvement for speech identification amid babble noise was present for both training and control groups, suggesting that the sustained improvement may relate more to the nature of babble noise than to the benefits of exposure or training *per se*. The follow-up study was based on only a subset of participants and therefore has lower power than the main analyses. *Post hoc* power analysis assessed the power for the main finding of a greater increase in word recognition from pre- to post-test for the training group compared to the control group, in babble noise. This indicated that the main study had power > 0.99 (one-tailed hypothesis; $\alpha = 0.05$, $n = 12$ per group), whereas the same analysis for pre-test to follow-up had power of only 0.72 ($n = 9$ in the training group and 7 in the control group). Nonetheless, the finding that improvements in speech-in-noise understanding were sustained over time for speech in babble noise, but not other types of noise, raises important questions for future research into the role of the background noise.

Why might the improvement from pre- to post-test be sustained over several weeks for words presented amid babble noise but not for words presented amid car and rain noise? Participants in the control group had identical exposure to the different background noises and yet had better word identification accuracy for words in babble noise than car or rain noise at follow-up, several weeks after the initial study. Neuroimaging studies demonstrate that speech and environmental stimuli show overlapping patterns of activation (Lewis et al., 2004; Loebach and Pisoni, 2008), and share the same auditory sound processing pathway leading to sound recognition (Kidd et al., 2007). However, it remains a controversial discussion regarding whether specific regions of the auditory cortex are selectively involved in processing speech. Overath et al. (2015) have argued there are structures in the auditory brain tuned for speech-specific spectro-temporal structure.

One simple possibility for the different levels of sustained learning is that babble noise affords more opportunities than steady-state noise for “glimpsing” (listening in the comodulated or uncomodulated dips; Rosen et al., 2013). While dips are present in the car and rain noise samples, they are less frequent and with reduced amplitude modulation (see Figure 1). Potentially, through exposure and/or training, participants learned to utilize dips more effectively, and this specific learning was sustained over time, benefiting the babble noise condition but not the car and rain noise conditions.

Understanding speech presented amid other speech sounds relies on being able to segregate the target stream from the background speech sounds (Brungart et al., 2001). Training with speech in babble noise can help listeners to “pick up” target sounds and “tune out” particular sorts of background noise (Van Engen, 2012). Listeners in the current study may make use of the speech cues (speech spectral components) in babble noise to pick up the target speech information and tune out the irrelevant sound information. Previous work has shown that auditory perceptual training can affect the distribution of attention to speech stimuli by training participants to inhibit the irrelevant sound cues, resulting in reduced processing of the unattended stimuli (Melara et al., 2002; Loebach and Pisoni, 2008; Tremblay et al., 2009; Murphy et al., 2017).

There are of course many differences in the structure over both time and frequency of the three background noises tested in this study. An active area of research in recent years (e.g., McDermott and Simoncelli, 2011) focuses on the way the brain processes and characterizes sound from the environment (e.g., rain, ocean waves, waterfall, traffic, swarms of insects, etc.). In the context of the present study we would refer to these as background noises. Sound textures are believed to be represented in the brain by statistics that summarize signal properties over space and time (McWalter and McDermott, 2018). It has been shown that when processing the underlying statistical structure of sounds, different strategies are used by the auditory system. For example, the time over which the brain integrates information on these sound statistics seems to adapt or vary for different sound textures (McWalter and McDermott, 2018). How training to ignore or suppress one sound texture might generalize to another is, to the authors’ knowledge, unknown and largely

unexplored. A potentially interesting line of research might be to determine what degree of similarity between noise/sound textures is needed to observe generalization. How one defines similarity (i.e., decides what statistical dimension to explore) for such a study is however not trivial.

Regarding the specific effects of training, it is interesting to note that participants in the training group improved in BKB word identification accuracy from pre-test to the first training session, and in each subsequent training session, but that there was no additional benefit from the third training session to the post-test. Many participants showed a sharp rise in word identification accuracy from pre-test to Training Session 1 (Figure 4), which is likely to result from increased familiarity with the test procedure and stimuli. Further gains during the training sessions are likely to result from participants learning to identify the speech sounds amid the background noise. Participants appear to have reached asymptotic performance following the three training sessions as no further benefits were found between Training Session 3 and the post-training test.

The present study found training benefits in normal-hearing young adults over a small number of short sessions. While these short sessions were sufficient to demonstrate an impact of the type of background noise on perceptual learning, longer training interventions may be needed to induce plasticity-related improvements in hearing-impaired individuals, such as hearing-aid and cochlear-implant users, or people with Auditory Processing Disorder (APD). For example, auditory training programs designed to alleviate auditory processing difficulties in APD typically involve multiple sessions per week (Weihing et al., 2015).

Participants in the control group did not attend or complete an alternative task between the pre- and post-test session. Some of the post-test benefit for the training group may therefore be due to factors unrelated to perceptual learning, such as increased concentration or familiarity with the content of BKB sentences or test environment. In future studies, an active control group could be included to identify the contribution of these factors to perceptual learning with different background noises, and to ensure equal exposure to BKB sentences across training and control groups. It is noteworthy that even without an active task, the control group showed prolonged benefits from simple exposure to speech in babble noise, but not other types of noise. One other key difference between the training and control groups is the gap prior to the post-training test session. Control participants returned after 1 day to complete the “post-training” session, whereas training-group participants completed the post-test session immediately after the Day 3 training, which potentially impacted their concentration or memory of the task. However, we are confident that these timing differences did not influence the overall results because there was no significant change for the training group between Day 3 training (1 day after the previous session) and the post-training test.

Participants gained a greater learning benefit for speech in babble noise than for speech in car or rain noise. While this is likely to be due to the increased practice with the specific speech-in-babble noise stimuli, an alternative is that the improved speech perception in the babble noise condition reflected increased

familiarity with the target stimuli (Brungart et al., 2001). This increased familiarity may have proved particularly beneficial in the babble noise condition due to the increased perceptual similarity between the target speech and background noise. One way to evaluate this possibility would be to train participants with speech in either car or rain noise and determine whether benefits are still enhanced for the speech-in-babble noise stimuli.

The present study provides evidence that the background noise should be considered in the development of speech-in-noise training paradigms, and could prove valuable in providing auditory training to participants with normal pure-tone audiometry thresholds. However, further research is needed to investigate how the background noise affects perceptual learning in participants with hearing difficulties, including hearing-aid and cochlear-implant users and those with auditory processing disorder (APD). People with APD can have particular difficulty understanding speech in noise, despite normal pure-tone thresholds in quiet (Iliadou et al., 2017), and so it may prove particularly important to understand the role of the background noise when developing auditory training paradigms for this group.

This study showed that benefits of training in babble noise generalized to car and rain background noise, in the short term at least. However, all three types of noise were broadly similar in their spectral and temporal profile. Future research could investigate whether learning benefits generalize to familiar noises with different spectro-temporal profiles (e.g., drumming), or whether generalization depends on perceptual similarity.

CONCLUSION

The present study has demonstrated that training participants to understand speech in babble noise leads to improved word identification accuracy. The benefits of training generalized, to some extent, to speech presented in car noise and rain noise, but were greatest for speech presented in babble noise. For both the training and control groups, improvements from the pre-training test to the post-training test were maintained over time for words presented in babble noise, but not for words presented in car and rain noise. As the sustained learning with babble noise, but not other noises, was observed in both groups, this implies that a structural feature of the babble noise (i.e., temporal fluctuations in envelope, residual pitch cues or some other feature) was conducive to sustained improvement in the BKB sentence task. This must be the case as the control group received equal exposure to the three noise types.

This study highlights the need to consider both the target and background sounds when creating auditory training programs. The outcomes provide important evidence for the use of background noise in perceptual auditory training programs to improve people's listening ability in challenging environments. In the future, it will be worth putting together the findings from this study with other auditory perceptual learning research (Kim et al., 2006; Wong et al., 2010; Parbery-Clark et al., 2011) to explore more methods to improve speech understanding in people who experience auditory perceptual difficulties. These

findings could be used as a baseline for further training for related auditory plasticity research in hearing impaired people, such as effects of age and hearing loss level on speech perception in noise (Dubno et al., 1984; Helfer and Wilber, 1990; Anderson et al., 2013; Henshaw and Ferguson, 2013; Alain et al., 2014; Karawani et al., 2016), or speech understanding in children with learning difficulties such as Auditory Processing Disorder (APD) (Bradlow et al., 2003; Ziegler et al., 2005, 2009).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Biomedical and Scientific Research Ethics Committee (BSREC) of the University of Warwick.

REFERENCES

- Alain, C., Zendel, B., Hutka, S., and Bidelman, G. (2014). Turning down the noise: the benefit of musical training on the aging auditory brain. *Hear. Res.* 308, 162–173. doi: 10.1016/j.heares.2013.06.008
- Amitay, S., Zhang, Y.-X., Jones, P. R., and Moore, D. R. (2014). Perceptual learning: top to bottom. *Vis. Res.* 99, 69–77. doi: 10.1016/j.visres.2013.11.006
- Anderson, S., White-Schwoch, T., Choi, H. J., and Kraus, N. (2013). Training changes processing of speech cues in older adults with hearing loss. *Front. Systems Neurosci.* 7:97. doi: 10.3389/fnsys.2013.00097
- Bench, J., Kowal, A., and Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentences lists for partially hearing children. *Br. J. Audiol.* 13, 108–112. doi: 10.3109/03005367909078884
- Boothroyd, A. (2007). Adult aural rehabilitation: what is it and does it work? *Trends Amplification* 11, 63–71. doi: 10.1177/1084713807301073
- Borrie, S. A., McAuliffe, M. J., and Liss, J. M. (2012). Perceptual learning of dysarthric speech: a review of experimental studies. *J. Speech Lang. Hear. Res.* 55, 290–305.
- Bradlow, A. R., Kraus, N., and Hayes, E. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. *J. Speech Lang. Hear. Res.* 46, 80–97.
- Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696
- Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J. Acoust. Soc. Am.* 110, 2527–2538. doi: 10.1121/1.1408946
- BSA (2011). *Recommended Procedure: Pure Tone Air and Bone Conduction Threshold Audiometry with and Without Masking and Determination of Uncomfortable Loudness Levels*. Irving, TX: The British Society of Audiology.
- Burkholder, R. A. (2005). *Perceptual Learning of Speech Processed Through an Acoustic Stimulation of a Cochlear Implant*. Ph.D. thesis, Indiana: Indiana University.
- Cassery, E. D., and Barney, E. C. (2017). Auditory training with multiple talkers and passage-based semantic cohesion. *J. Speech Lang. Hear. Res.* 60, 159–171. doi: 10.1044/2016_JSLHR-H-15-0357
- Clarke, C. M., and Garrett, M. F. (2004). Rapid adaptation to foreign accented english. *J. Acoustical Soc. Am.* 116, 3647–3658. doi: 10.1121/1.1815131

The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LZ set up the study, tested participants, and analyzed the data. All authors discussed the findings, contributed to writing the manuscript, helped to design the study, and approved the final version.

ACKNOWLEDGMENTS

We would like to thank Dr. Mark Elliott for his valuable suggestions for the experiment set up. We thank also to Dr. Dhammika Widanalage and Dr. Simon Davies for their useful suggestion about Matlab coding. We thank to WMG for providing the funding to conduct this speech experiment. Finally, we thank to all the listeners for participation in the experiment.

- Clopper, C. G., and Pisoni, D. B. (2004). Effects of talker variability on perceptual learning of dialects. *Lang. Speech* 47, 207–239. doi: 10.1177/00238309040470030101
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A. G., Taylor, K., and McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise vocoded sentences. *J. Exp. Psychol.: General* 134, 222–241. doi: 10.1037/0096-3445.134.2.222
- Dorman, M. F., and Wilson, B. S. (2004). The design and function of cochlear implants. *Am. Sci.* 92, 436–445. doi: 10.1511/2004.49.942
- Dubno, J. R., Dirks, D. D., and Morgan, D. E. (1984). Effects of age and mild hearingloss on speech recognition in noise. *J. Acoustical Soc. Am.* 76, 87–96. doi: 10.1121/1.391011
- Felty, R. A., Buchwald, A., and Pisoni, D. B. (2009). Adaptation to frozen babble in spoken word recognition. *J. Acoustical Soc. Am.* 125, EL93–EL97. doi: 10.1121/1.3073733
- Francis, A. L., Fenn, K., and Nusbaum, H. C. (2007). Effects of training on the acoustic phonetic representation of synthetic speech. *J. Speech Lang. Hear. Res.* 50, 1445–1465.
- Francis, A. L., and Nusbaum, H. C. (2009). Effects of intelligibility on working memory demand for speech perception. *Attent. Percept. Psychophys.* 71, 1360–1374. doi: 10.3758/APP.71.6.1360
- Goldstone, R. L. (1998). Perceptual learning. *Annu. Rev. Psychol.* 49, 585–612.
- Green, T., Faulkner, A., and Rosen, S. (2019). Computer-based connected-text training of speech-in-noise perception for cochlear implant users. *Trends Hear.* 23, 1–11. doi: 10.1177/2331216519843878
- Grimault, N., Micheyl, C., Carlyon, R. P., Bacon, S. P., and Collet, L. (2003). Learning in discrimination of frequency or modulation rate: generalization to fundamental frequency discrimination. *Hear. Res.* 184, 41–50. doi: 10.1016/S0378-5955(03)00214-4
- Halliday, L. F., Taylor, J. L., Edmondson-Jones, A. M., and Moore, D. R. (2008). Frequency discrimination learning in children. *J. Acoustical Soc. Am.* 123, 4393–4402. doi: 10.1121/1.2890749
- Helfer, K. S., and Wilber, L. A. (1990). Hearing-loss, aging, and speech-perception in reverberation and noise. *J. Speech Lang. Hear. Res.* 33, 149–155. doi: 10.1044/jshr.3301.149
- Henshaw, H., and Ferguson, M. A. (2013). Efficacy of individual computer-based auditory training for people with hearing loss: a systematic review of the evidence. *PLoS One* 8:e62836. doi: 10.1371/journal.pone.0062836
- Hervais-Adelman, A., Davis, M. H., Taylor, K., Johnsrude, I. S., and Carlyon, R. P. (2011). Generalization of perceptual learning of vocoded speech. *J. Exp. Psychol.: Hum. Percept. Perform.* 37, 283–295. doi: 10.1037/a0020772

- Hirata, J. (2004). Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *J. Acoustical Soc. Am.* 116, 2384–2394. doi: 10.1121/1.1783351
- Howard-Jones, P. A., and Rosen, S. (1993). Unmodulated glimpsing in 'checkerboard' noise. *J. Acoustical Soc. Am.* 93, 2915–2922.
- Iliadou, V., Ptok, M., Grech, H., Pedersen, E. R., Brechmann, A., Deggouj, N., et al. (2017). A European perspective on auditory processing disorder: current knowledge and future research focus. *Front. Neurol.* 8:622. doi: 10.3389/fneur.2017.00622
- Irvine, D. R. F. (2018). Plasticity in the auditory system. *Hear. Res.* 362, 61–73. doi: 10.1016/j.heares.2017.10.011
- Jones, P. R., Moore, D. R., Shub, D. E., and Amitay, S. (2015). The role of response bias in perceptual learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 1456–1470. doi: 10.1037/xlm0000111
- Karawani, H., Bitan, T., Attias, J., and Banai, K. (2016). Auditory perceptual learning in adults with and without age-related hearing loss. *Front. Psychol.* 6:2066. doi: 10.3389/fpsyg.2015.02066
- Kidd, G. R., Watson, C. S., and Gygi, B. (2007). Individual differences in auditory abilities. *J. Acoustical Soc. Am.* 122, 418–435. doi: 10.1121/1.2743154
- Kim, S. H., Frisina, R. D., Mapes, F. M., Hickman, E. D., and Frisina, D. R. (2006). Effect of age on binaural speech intelligibility in normal hearing adults. *Speech Commun.* 48, 591–597. doi: 10.1016/j.specom.2005.09.004
- Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., and DeYoe, E. A. (2004). Human brain regions involved in recognizing environmental stimuli. *Cereb. Cortex* 14, 1008–1021. doi: 10.1093/cercor/bhh061
- Loebach, J. L., and Pisoni, D. B. (2008). Perceptual learning of spectrally-degraded speech and environmental sounds. *J. Acoustical Soc. Am.* 123, 1126–1139. doi: 10.1121/1.2823453
- Loebach, J. L., Pisoni, D. B., and Svirsky, M. A. (2009). Transfer of auditory perceptual learning with spectrally reduced speech to speech and nonspeech tasks: implications for cochlear implants. *Ear Hear.* 30, 662–674. doi: 10.1097/AUD.0b013e3181b9c92d
- McDermott, J. H., and Simoncelli, E. P. (2011). Sound Texture Perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron* 71, 926–940. doi: 10.1016/j.neuron.2011.06.032
- McWalter, R., and McDermott, J. H. (2018). Adaptive and selective time averaging of auditory scenes. *Curr. Biol.* 28, 1405–1418. doi: 10.1016/j.cub.2018.03.049
- Melara, R. D., Rao, A., and Tong, Y. (2002). The duality of selection: excitatory and inhibitory processes in auditory selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 279–306. doi: 10.1037/0096-1523.28.2.279
- Moore, D., and Shannon, R. (2009). Beyond cochlear implants: awakening the deafened brain. *Nat. Neurosci.* 12, 686–691. doi: 10.1038/nn.2326
- Morris, C. D., Bransford, J. D., and Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *J. Verb. Learn. Verb. Behav.* 16, 519–533. doi: 10.1016/S0022-5371(77)80016-9
- Murphy, S., Spence, C., and Dalton, P. (2017). Auditory perceptual load: a review. *Hear. Res.* 352, 40–48. doi: 10.1016/j.heares.2017.02.005
- Overath, T., McDermott, J. H., Zarate, J. M., and Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by response to sound quilts. *Nat. Neurosci.* 18, 903–911. doi: 10.1038/nn.4021
- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., and Kraus, N. (2011). Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One* 6:e18082. doi: 10.1371/journal.pone.0018082
- Reed, C. M., and Delhorne, L. A. (2005). Reception of environmental sounds through cochlear implants. *Ear Hear.* 26, 48–61.
- Ricketts, T. A., and Hornsby, B. W. (2005). Sound quality measures for speech in noise through a commercial hearing aid implementing digital noise reduction. *J. Am. Acad. Audiol.* 16, 270–277. doi: 10.3766/jaaa.16.5.2
- Roberts, K. L., and Allen, H. A. (2016). Perception and cognition in the ageing brain: a brief review of the short- and long-term links between perceptual and cognitive decline. *Front. Aging Neurosci.* 8:39. doi: 10.3389/fnagi.2016.00039
- Roediger, H. L., Weldon, M. S., and Challis, B. H. (1989). "Explaining dissociations between implicit and explicit measures of retention: a processing account," in *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving*, eds H. L. Roediger III and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 3–41.
- Rosen, S., Souza, P., Ekelund, C., and Majeed, A. A. (2013). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *J. Acoustical Soc. Am.* 133, 2431–2443. doi: 10.1121/1.4794379
- Samuel, A. G., and Kraljic, T. (2009). Perceptual learning for speech. *Attent. Percept. Psychophys.* 71, 1207–1218. doi: 10.3758/APP.71.6.1207
- Schubö, A., Schlaghecken, F., and Meinecke, C. (2001). Learning to ignore the mask in texture segmentation tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 919–931. doi: 10.1037/0096-1523.27.4.919
- Schwab, E. C., Nusbaum, H. C., and Pisoni, D. B. (1985). Some effects of training on the perception of synthetic speech. *Hum. Factors* 27, 395–408. doi: 10.1177/001872088502700404
- Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends Cogn. Sci.* 12, 182–186. doi: 10.1016/j.tics.2008.02.003
- Stacey, P. C., and Summerfield, A. Q. (2007). Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *J. Acoustical Soc. Am.* 121, 2923–2935. doi: 10.1121/1.2713668
- Stacey, P. C., and Summerfield, A. Q. (2008). Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *J. Speech Lang. Hear. Res.* 51, 526–538.
- Studebaker, G. A. (1985). A 'rationalized' arcsine transform. *J. Speech Hear. Res.* 28, 455–462.
- Sweetow, R., and Palmer, C. V. (2005). Efficacy of individual auditory training in adults: a systematic review of the evidence. *J. Am. Acad. Audiol.* 16, 494–504.
- Sweetow, R. W., and Sabes, J. H. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) program. *J. Am. Acad. Audiol.* 17, 538–558. doi: 10.3766/jaaa.17.8.2
- Tremblay, K. L., and Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *J. Speech Lang. Hear. Res.* 45, 564–572.
- Tremblay, K. L., Shahin, A. J., Picton, T., and Ross, B. (2009). Auditory training alters the physiological detection of stimulus-specific cues in humans. *Clin. Neurophysiol.* 120, 128–135. doi: 10.1016/j.clinph.2008.10.005
- Van Engen, K. J. (2012). Speech-in-speech recognition: a training study. *Lang. Cogn. Processes* 27, 1089–1107. doi: 10.1080/01690965.2012.654644
- Verschuur, C., Boland, C., Frost, E., and Constable, J. (2013). The role of first formant information in simulated electro-acoustic hearing. *J. Acoustical Soc. Am.* 133, 4279–4289. doi: 10.1121/1.4803910
- Weihing, J., Chermak, G. D., and Musiek, F. E. (2015). Auditory training for central auditory processing disorder. *Sem. Hear.* 36, 199–215. doi: 10.1055/s-0035-1564458
- Wong, P., Ettlinger, M., Sheppard, J., Gunasekera, G., and Dhar, S. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear Hear.* 31, 471–479. doi: 10.1097/AUD.0b013e3181d709c2
- Wong, P. C. M., and Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Appl. Psycholinguistic* 28, 565–585. doi: 10.1017/S0142716407070312
- Zhang, L., Schlaghecken, F., Harte, J., and Jennings, P. (2014). *Learning to Ignore Background Noise in VCV Test, the Proceeding of the 2014 BSA Annual Conference, UK.* 29–30.
- Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F. X., and Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14110–14115. doi: 10.1073/pnas.0504446102
- Ziegler, J. C., Pech-Georgel, C., George, F., and Lorenzi, C. (2009). Speech perception- in-noise deficits in dyslexia. *Dev. Sci.* 12, 732–745. doi: 10.1111/j.1467-7687.2009.00817.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhang, Schlaghecken, Harte and Roberts. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.