ORIGINAL ARTICLE
J Audiol Otol 2020;24(4):167-173
pISSN 2384-1621 / eISSN 2384-1710
https://doi.org/10.7874/jao.2019.00472

Relationship between Speech Perception in Noise and Phonemic Restoration of Speech in Noise in Individuals with Normal Hearing

Srikar Vijayasarathy and Animesh Barman

Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, University of Mysore-Mysuru, Karnataka, India

Received December 12, 2019
Revised May 19, 2020
Accepted June 10, 2020

Address for correspondence
Srikar Vijayasarathy, MSc
All India Institute of Speech and Hearing, Manasagangotri, Naimisham Campus, University of Mysore-Mysuru, Karnataka 570006, India
Tel +91-9539732851
E-mail srkrv.y@gmail.com

Background and Objectives: Top-down restoration of distorted speech, tapped as phonemic restoration of speech in noise, maybe a useful tool to understand robustness of perception in adverse listening situations. However, the relationship between phonemic restoration and speech perception in noise is not empirically clear. **Subjects and Methods**: 20 adults (40−55 years) with normal audiometric findings were part of the study. Sentence perception in noise performance was studied with various signal-to-noise ratios (SNRs) to estimate the SNR with 50% score. Performance was also measured for sentences interrupted with silence and for those interrupted by speech noise at -10, -5, 0, and 5 dB SNRs. The performance score in the noise interruption condition was subtracted by quiet interruption condition to determine the phonemic restoration magnitude. **Results**: Fairly robust improvements in speech intelligibility was found when the sentences were interrupted with speech noise instead of silence. Improvement with increasing noise levels was non-monotonic and reached a maximum at -10 dB SNR. Significant correlation between speech perception in noise performance and phonemic restoration of sentences interrupted with -10 dB SNR speech noise was found. **Conclusions**: It is possible that perception of speech in noise is associated with top-down processing of speech, tapped as phonemic restoration of interrupted speech. More research with a larger sample size is indicated since the restoration is affected by the type of speech material and noise used, age, working memory, and linguistic proficiency, and has a large individual variability. **J Audiol Otol 2020;24(4):167-173**

KEY WORDS: Hearing · Speech perception · Speech intelligibility · Signal-to-noise ratio · Illusions.

Introduction

The human auditory system is remarkably robust in adverse listening situations. Listeners can make use of available bottom-up cues and 'restore' the sounds masked by noise by making use of higher-order cues, like context and linguistic knowledge [1,2]. One form of this top-down restoration is 'phonemic restoration of speech in noise' [1,3,4]. Phonemic restoration is typically measured as the improvement seen in intelligibility of speech interrupted by silence when the silent interruptions are filled with noise instead [5-7]. It has been demonstrated with words and non-words [3] as well as with sentences [2,6-8] using discrimination [3], continuity [9], and intelligibility paradigms [2,4]. One hypothesis used to explain this improved speech perception ability is that introduction of noise in this paradigm may help the auditory system group speech and noise into different streams, resulting in a more continuous percept of speech [10]. The noise can also mask spurious cues introduced by interrupting continuous speech with silence, leading to an increase in ambiguity and a greater activation of lexical networks facilitating restoration [10]. The magnitude of restoration, however, is characterized by large individual variability [11,12].

It has been implied by many investigators that phonemic restoration, being an aspect of top-down restoration of speech, maybe an important part of listening in adverse listening sit-

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2020 The Korean Audiological Society and Korean Otological Society

uations [6,8,13]. However, the relationship between the two is not empirically clear. While it is well-known that speech perception in noise is affected in individuals with sensorineural hearing loss and in those rehabilitated with cochlear implants, phonemic restoration has also been reported to be reduced in these two groups [6,11,14] suggesting that speech in noise performance and phonemic restoration maybe positively correlated. On the other hand, speech in noise performance is affected in older subjects [15], but phonemic restoration magnitude seems to be larger than relative to younger subjects [7,16,17]. It may well be that the two processes may not be correlated at all and may represent two independent processes. Phonemic restoration in noise is thus multidimensional and seems to be dependent on a host of factors–both bottom-up and top-down. It is an important tool since it connects envelope distortion in adverse listening situations, dip listening, and auditory scene analysis (difficulty to group silence-interrupted speech into one stream, and easier grouping in noise-interrupted speech), all in the context of speech perception. Clearly, more studies are needed to understand how speech in noise and phonemic restoration are related and how that knowledge can be used to improve perception in adverse listening situations. In this study, we measured speech perception in noise performance in the form of signal-to-noise ratio with 50% correct speech identification (SNR-50), the signal-to-noise ratio (SNR) required to achieve 50% speech intelligibility, and investigated its correlation with the magnitude of phonemic restoration of speech in noise at various SNRs in subjects with clinically normal hearing sensitivity.

Subjects and Methods

Participants

The study consisted of 20 native speakers (11 male and 9 female) with an age range of 40−55 years (mean=45.3 years). All the subjects had their thresholds within 15 dB HL over the frequency range (in octaves) of 250 Hz to 8,000 Hz for air conduction stimuli and between 250 Hz and 4,000 Hz for bone conduction stimuli. They had normal tympanometric findings, stapedial acoustic reflex thresholds, Transient evoked oto-acoustic emissions, and Auditory Brainstem responses. They also cleared the Screening Checklist for Auditory Processing in Adults [18]. None of them had a history of otological and/or neurological complaints. “Ethical Guidelines for bio-behavioral research involving human subjects” [19] was followed and the study was approved by the Institutional Review Board.

Speech in noise perception

Speech in noise measurement was measured using the sentence list developed by Geetha, et al. [20]. Each list consisted of 10 sentences, each with four keywords. The sentences were concatenated, the long-term power spectrum was calculated and was used to filter broad-band noise into a speech-shaped noise. The speech level was kept constant at 65 dB SPL and mixed with speech-shaped noise to produce a range of SNRs from -10 to +8 dB in 2 dB steps (10 SNR conditions in total). Each sentence list had one sentence at each of these SNRs. The order of SNRs in the lists was randomized. The sentences were routed through a personal computer and presented to the right ear using Sennheiser HDA 200 (Sennheiser, Wedemark, Germany) headphones, and subjects were instructed to listen to the stimuli and repeat what they heard in verbatim. They were encouraged to guess the words if they were not sure of what they heard. Subjects were familiarized with the procedure using a sentence list mixed at 5 dB SNR. The responses were scored online, but were also recorded for offline verification. Each keyword repeated correctly was awarded one point, and the total for each list was calculated. The SNR-50 was estimated based on the Spearman-Karber equation [21]. Two lists were used and the SNR-50 obtained in each list was averaged to calculate the final value.

Perceptual restoration of sentences in noise

Sentence lists developed by Geetha, et al. [20] were also used in this part of the experiment. The sentence lists used were different from that used for speech in noise measurement to avoid practice effect (The sentence bank cited consists of 20 different lists of similar difficulty). The sentences were interrupted in two ways: with periodic silent intervals, or with the silent intervals filled with speech-shaped noise bursts. Interruptions were applied by modulating the sentences with a periodic, 1.5-Hz square wave with duty cycles of 50% on-duration (corresponding to 333 ms), ramped with 5 ms of raised cosine. The rate of interruption and duty cycle were chosen based on what has been shown to consistently yield phonemic restoration [6]. For interruption with speech noise, sentence level was kept constant at 65 dB SPL, and noise level was varied to create lists with SNRs of 5, 0, -5, and -10 dB. The stimulus presentation and the response recording were similar to that of speech in noise measurement. The order of lists was randomized to prevent any order effect. Familiarization trials were carried out before the actual assessment using sentences interrupted at 1 Hz with silence and -2 dB SNR speech-shaped noise. Phonemic Restoration benefit was calculated as the difference in the intelligibility of sentences interrupted by silence and speech-shaped noise.

Statistical analysis

Statistical package for the Social Sciences software (version 18, SPSS Inc., Chicago, IL, USA) was used for analysis for both descriptive and inferential statistics. In the interrupted speech intelligibility task, the speech identification scores did not deviate significantly from normality in any of the conditions based on Shapiro Wilk test (p>0.05) and met the sphericity assumption (Mauchly's test of sphericity, p>0.05). Repeated measures Analysis of Variance was done with the interruption condition as the main effect and was followed up by Bonferroni post-hoc pairwise comparisons. Pearson product moment correlation was used to analyze the relationship between speech in noise perception and phonemic restoration.

Results

Speech in noise performance

The estimated SNR-50 was used to quantify speech in noise performance. The mean SNR-50 in the sample was -4.38 dB with a standard deviation (SD) of 0.75 dB. The 95% confidence interval spread was -4.8 dB to -3.99 dB (Fig. 1).

Phonemic restoration of speech in noise

Speech identification was measured for sentences interrupted with silence or speech-shaped noise (Fig. 2). The mean score with silent interruption was 22.5 (SD=0.99) and improved when silent interruptions were filled with speech shaped noise. The mean (SD in parenthesis) for the noise interruptions at various SNRs were 25.8 (2.27), 28.9 (1.68), 28.7 (1.67), and 32.5 (1.68) at +5, 0, -5, and -10 dB SNR, respectively.

Repeated measures ANOVA indicated a main effect of interruption condition ($F_{(3,12)}$=46.1, p=0.00, nP^2=0.92). Bonferroni post-hoc pairwise comparisons revealed that all interruption conditions were significantly different from each other (p<0.05) with the exception of 0 dB vs. -5 dB SNR conditions (p>0.05). Phonemic restoration of speech in noise was calculated as the difference between the performance for sentences interrupted with silence and those interrupted by

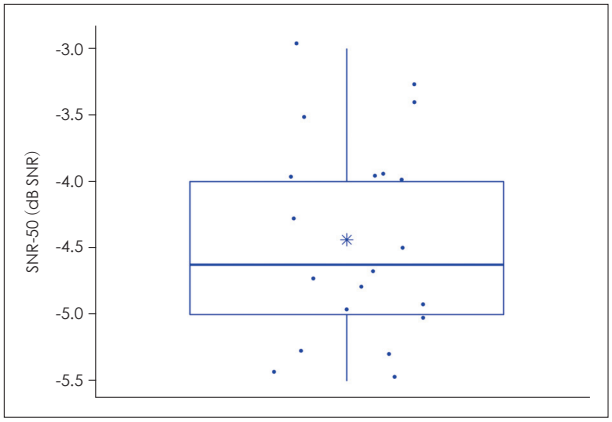

Fig. 1. Boxplot of SNR-50. The asterisk represents the mean. SNR-50 varied from as poor as -3 dB SNR to as high as -5.5 dB SNR. SNR-50: SNR with 50% correct speech identification.

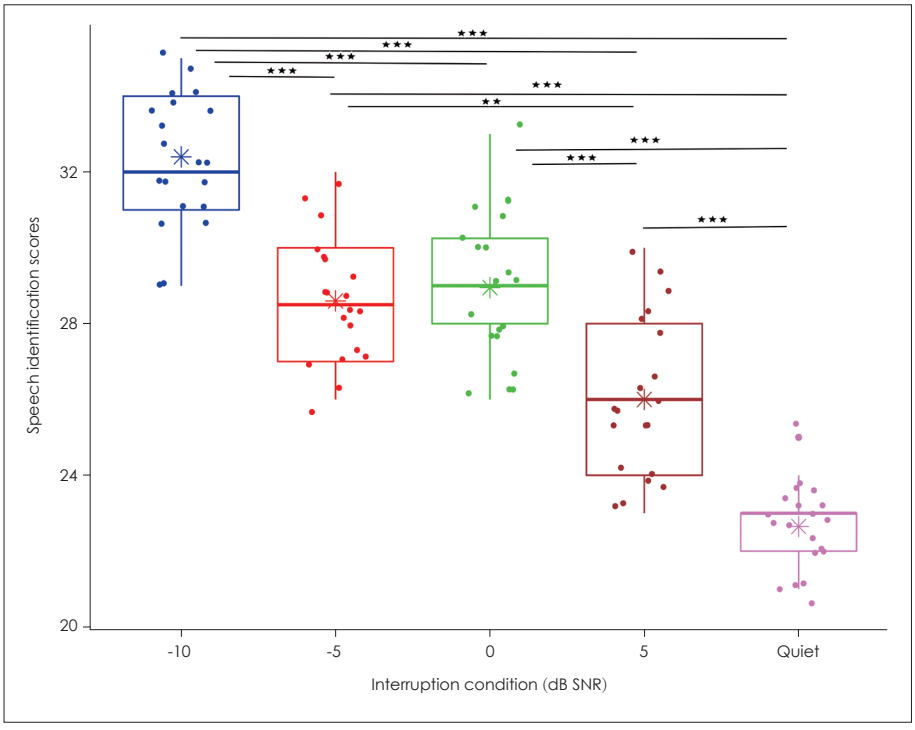

Fig. 2. Boxplot of speech identification scores for interruptions with silence and with different SNRs of speech noise (***p<0.001, **p<0.01, *p<0.05). There was a general trend for improvement of performance with increasing noise level. Note the individual variation across individuals, especially at +5 dB SNR.

speech-shaped noise (Fig. 3). Phonemic restoration tended to increase with the noise level and was the highest at -10 dB SNR. The improvement function was non-monotonic with a modest improvement at +5 dB SNR (with respect to quiet), larger increments at 0 and -5 dB SNR interruption conditions, followed by the largest increment at -10 dB SNR.

Correlation between speech perception in noise and perceptual restoration of speech in noise

A trend for a negative correlation with increasing noise levels was observed (Fig. 4). While correlations were not statistically significant at +5, 0, and -5 dB SNRs (lower noise levels), a moderate negative correlation (r=-0.60, p=0.017) was found between restoration magnitude at -10 dB SNR and SNR-50 indicating that those with more negative SNR-50 (better performance in noise) also tended to have a greater phonemic restoration of speech in noise.

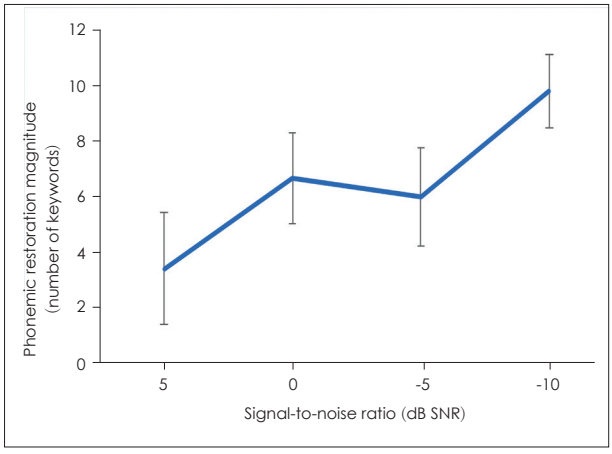

Fig. 3. Plot of magnitude of restoration across different SNR of speech-shaped noise.

Discussion

The hypothesis of the study was that there would be a correlation between speech perception in noise (SPIN) scores and phonemic restoration of speech, since those with better speech intelligibility in noise may have better top-down repair strategies to overcome the deleterious effects of noise.

Speech in noise performance

The SNR-50 score can vary widely with the material used as well as the competing signal employed [22]. Quick Speech-in-Noise test (Quick-SIN) [23] places it at +2 dB while Bamford-Kowal-Bench Speech-in-Noise test (BKB-SIN) and Hearing in Noise Test (HINT) have their 50% performances at -2.5 dB and -2.92 dB SNRs, respectively [22]. The former two tests use speech babble maskers, while HINT uses speech shaped noise instead. The SNR-50 scores obtained in this study (mean SNR-50 was -4.38 dB) are in agreement with those reported by Jain [24] who used similar speech material and noise conditions used in the present study. They reported a mean SNR-50 of -4.3 dB in their sample of 50−60 year-old subjects. The relatively large variation observed in SNR-50, even though hearing sensitivity itself was within normal limits, is in line with observations about individual variability in suprathreshold processing in normal hearing subjects [22,23].

Perceptual restoration of speech in noise

Interruption with speech-shaped noise led to improved speech intelligibility [1,2,4,6,11]. The growth of restoration

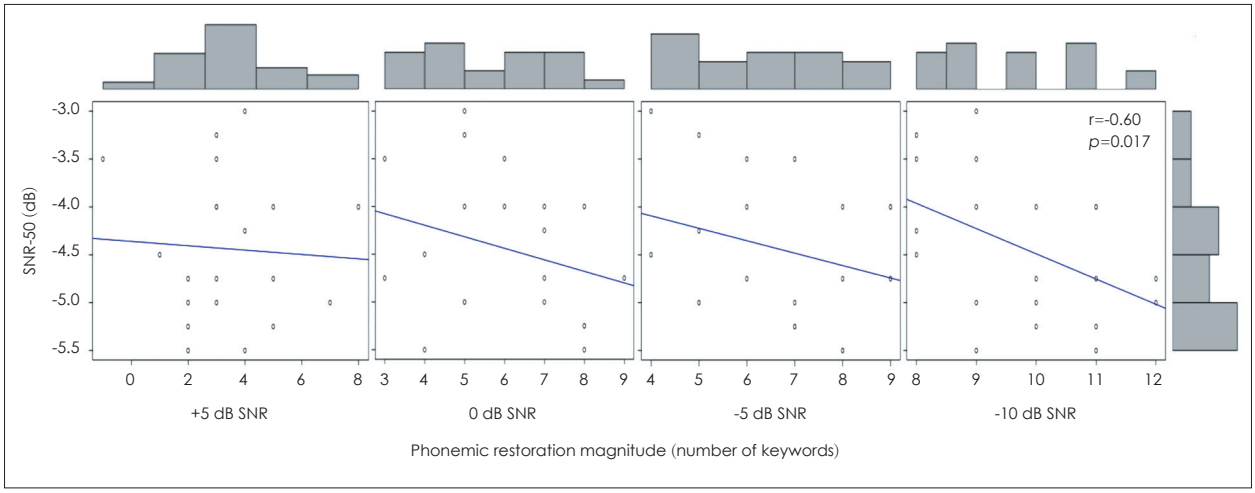

Fig. 4. Scatterplot of SNR-50 and phonemic restoration at different SNRs. The histograms of the variables are also shown on the x- and y-axis. Correlation was not statistically significant when interrupted with lower noise levels. SNR-50: SNR with 50% correct speech identification.

magnitude was non-monotonic. A similar pattern was also reported by Bhargava and colleagues [6] as well. Direct comparison of the magnitude of phonemic restoration obtained in the present study vis-a-vis other studies is difficult since we preferred to use the raw scores instead of the typical rationalized arcsine unit (RAU) transformation [6,8] or percentages [11,25]. Since the datapoints were already normally distributed without large skews, we preferred this over other transformations to preserve the actual/natural distribution of speech scores. Transforming the data scores for the sake of comparison with other studies, the average restoration in RAU was around 4.96, 9.82, 9.60, and 16.57 for +5, 0, -5, and -10 dB SNRs, respectively, which are similar to those reported by other studies [6-8], after taking into consideration the relatively easy/predictable sentence lists used as well as the age of the subjects who participated in the study. In general, the easier the list, the better the phonemic restoration [7]. Age is an important consideration since studies [7,16] suggest that phonemic restoration is in fact increased in magnitude in older individuals (without much hearing loss). They suggest that older individuals tend to have poorer access to bottom-up cues and maybe more dependent on top-down processing, resulting in a larger phonemic restoration. Variations in working memory abilities and linguistic proficiency [7,12,25] among the subjects could also contribute to the variance in the magnitude of restoration.

The reason for this intelligibility increase with noise interruption over silent interruption is likely multi-faceted. Silent interruptions may introduce spurious cues, like word-ending, stop-burst etc., thus, feeding wrong input to the top-down mechanisms. It is also possible that silent interruptions make it difficult to group the segments of speech into one stream [4]. Introduction of noise in these interruptions can alleviate this by masking off the spurious cues and may also help in integration of speech segments into one stream (noise being the other stream), resulting in activation of larger lexical networks [10]. Top-down mechanisms are thus able to make better lexical choices leading to better speech intelligibility. Increased noise levels may make speech feel more continuous [6] and may also contribute to better speech intelligibility at higher noise levels. Increased noise level can also help in better discrimination of speech and noise, which is helpful in grouping of speech and noise into different streams [7,10,13] and in turn, facilitate a better top-down repair.

Degraded perception of bottom-up cues can affect phonemic restoration by ineffective activation of top-down networks [6,7]. Reduced phonemic restoration has been reported in individuals with normal hearing when speech is distorted [6,7, 11,13]. Reduction in phonemic restoration has also been reported in those with sensory neural hearing loss [11] and in subjects who have undergone cochlear implantation [6]. Degraded bottom-up cues due to hidden hearing loss [26] resulting from cochlear synaptopathy (due to causes like ageing, noise exposure etc.) may also be a factor since cochlear synaptopathy can result in affected temporal processing [27] and poorer speech perception in noise performance [28,29]. This may explain some individual variation across suprathreshold processing tests in subjects with normal peripheral hearing sensitivity.

Correlation between phonemic restoration and speech in noise

We found a significant moderate negative correlation (Fig. 4) between SNR-50 scores and phonemic restoration magnitude at -10 dB SNR. Those with larger magnitude of phonemic restoration effect tended to be more resistant to the effects of noise (had a more negative SNR-50). This correlation could indicate that there maybe an overlap of mechanisms involved in speech perception in noise and phonemic restoration. It is also possible, however, that both of them are related to something else altogether, and the two are not actually directly correlated.

Speech in noise perception is multidimensional and involves: integration of available bottom-up cues (dip listening for example, in the context of interrupted speech), using cues, like fundamental frequency, to perceive target speech as one stream and others as background, resorting to educated guesses to compensate for missed parts based on the context as well as making predictions on the go as to what maybe spoken next [30]. As such, there can be no doubt that it involves a complex interaction of bottom-up and top-down mechanisms in an effort to make sense of what is being heard. Phonemic restoration is a top-down effect based on the perception of illusory continuity of speech and grouping of speech and noise into separate streams, leading to a better access of lexical candidates [2,4,6,10,11,25]. So, it is possible that mechanisms serving phonemic restoration may well be at least partially involved in speech in noise perception as well and hence, the observed correlation. Those with better restoration abilities than their counterparts may perform better when bottom-up cues become disrupted as frequently happens in real world communication scenarios.

It is interesting that no statistically significant correlation with SNR-50 was found at better SNRs (lower noise) where the magnitude of restoration was less. This raises the possibility that a good amount of phonemic restoration is required for the mechanisms to sufficiently overlap with that of speech in noise perception, which as discussed before, also has components of top-down processing. However, the explanation

maybe simpler. Verbal working memory (like speech in noise) has been reported to be associated with phonemic restoration only when the interruptions were with speech noise, but not with other types of noise [25]. It is thus possible that top-down mechanisms are better activated when evidences of gaps are masked more efficiently (for example, with a more intense noise). Further studies are needed to verify if this is indeed the general case or if this is just a one-off finding.

The type of noise used for interruption affects phonemic restoration in a complex manner. Noise similar to speech makes it a more plausible masker [2,4] and leads to a perception of continuity of interrupted speech. However, it is also the case that noise needs to be sufficiently different from speech as well [10] in order to split into a different stream (and not become part of the same stream as speech) and thus help in better restoration. For the same reason, it is possible that a speech babble masker might lead to a lower phonemic restoration than speech-shaped noise. Studying correlation with different kinds of maskers may give us more insights on the complex interaction between opposing cues operating to serve the same phenomenon and help us better understand the relationship between phonemic restoration and speech in noise perception.

In summary, the study found a negative correlation between speech perception in noise and phonemic restoration magnitude at -10 dB SNR, suggesting that this type of top-down repair maybe operational in adverse listening situations. However, the findings have to be interpreted within the constraints of the limitations of the study. This is a correlational study and no causal relationship can be established. The findings also need to be considered provisional until they are replicated on a larger sample using different types of maskers over a wide range of SNRs. Future studies can focus on investigating phonemic restoration in the SNRs between -5 and -10 dB SNRs with smaller step sizes (e.g.: 1 dB SNR steps) since this range seems promising to gain further insight on the nature of phonemic restoration and its relationship with perception under adverse listening situations.

Acknowledgments

We thank the director of All India Institute of Speech and Hearing, Mysuru and HOD, Audiology for granting us the permission to conduct the study. We are grateful to the participants of the study for their cooperation.

Conflicts of interest

The authors have no financial conflicts of interest.

Author Contributions

Conceptualization: all authors. Data curation: Srikar Vijayasarathy. Formal analysis: all authors. Methodology: all authors. Project administration: all authors. Resources: all authors. Validation: all authors. Writing—original draft: Srikar Vijayasarathy. Writing—review & editing: all authors. Approval of final manuscript: all authors.

ORCID iDs

Srikar Vijayasarathy https://orcid.org/0000-0003-0064-3789
Animesh Barman https://orcid.org/0000-0001-8158-7584

REFERENCES

1) Warren RM. Perceptual restoration of missing speech sounds. Science 1970;167:392-3.
2) Verschuure J, Brocaar MP. Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise. Percept Psychophys 1983;33:232-40.
3) Samuel A. Phoneme restoration. Lang Cognitive Proc 1996;11:647-54.
4) Bashford JA Jr, Riener KR, Warren RM. Increasing the intelligibility of speech through multiple phonemic restorations. Percept Psychophys 1992;51:211-7.
5) Başkent D. Effect of speech degradation on top-down repair: phonemic restoration with simulations of cochlear implants and combined electric-acoustic stimulation. J Assoc Res Otolaryngol 2012;13:683-92.
6) Bhargava P, Gaudrain E, Başkent D. Top-down restoration of speech in cochlear-implant users. Hear Res 2014;309:113-23.
7) Jaekel BN, Newman RS, Goupell MJ. Age effects on perceptual restoration of degraded interrupted sentences. J Acoust Soc Am 2018; 143:84-97.
8) Başkent D, Eiler C, Edwards B. Effects of envelope discontinuities on perceptual restoration of amplitude-compressed speech. J Acoust Soc Am 2009;125:3995.
9) Bashford JA, Warren RM. Perceptual synthesis of deleted phonemes. J Acoust Soc Am 1979;65:S112.
10) Srinivasan S, Wang D. A schema-based model for phonemic restoration. Speech Commun 2005;45:63-87.
11) Başkent D. Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores. J Acoust Soc Am 2010;128:EL169-74.
12) Benard MR, Başkent D. Perceptual learning of temporally interrupted spectrally degraded speech. J Acoust Soc Am 2014;136:1344-51.
13) Clarke J, Başkent D, Gaudrain E. Pitch and spectral resolution: a systematic comparison of bottom-up cues for top-down repair of degraded speech. J Acoust Soc Am 2016;139:395-405.
14) Hopkins K, Moore BCJ. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. J Acoust Soc Am 2011;130:334-49.
15) Plomp R, Mimpen AM. Speech-reception threshold for sentences as a function of age and noise level. J Acoust Soc Am 1979;66:1333-42.
16) Saija JD, Akyürek EG, Andringa TC, Başkent D. Perceptual restoration of degraded speech is preserved with advancing age. J Assoc Res Otolaryngol 2014;15:139-48.
17) Bologna WJ, Vaden KI Jr, Ahlstrom JB, Dubno JR. Age effects on perceptual organization of speech: contributions of glimpsing, phonemic restoration, and speech segregation. J Acoust Soc Am 2018; 144:267-81.
18) Vaidyanath R, Yathiraj A. Screening checklist for auditory processing in adults (SCAP-A): development and preliminary findings. J Hear Sci 2014;4:27-37.
19) Venkateshan S. Ethical guidelines for bio behavioral research. 2nd ed. Mysuru, India: All India Institute of Speech and Hearing;2009.
20) Geetha C, Kumar KSS, Manjula P, Pavan M. Development and standardisation of the sentence identification test in the Kannada language. J Hear Sci 2014;4:18-26.
21) Finney DJ. Statistical method in biological assay. 3rd ed. London, UK: Charles Griffin & Co.;1978.
22) Wilson RH, McArdle RA, Smith SL. An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. J Speech Lang Hear Res 2007;50:

844-56.
23) Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 2004;116:2395-405.
24) Jain C. Relationship among psychophysical abilities, speech perception in noise and working memory in individuals with normal hearing sensitivity across age groups. [dissertation]. Mysuru: University of Mysore;2016. p.168.
25) Nagaraj NK, Magimairaj BM. Role of working memory and lexical knowledge in perceptual restoration of interrupted speech. J Acoust Soc Am 2017;142:3756-3766.
26) Kujawa SG, Liberman MC. Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci 2009;29:14077-85.
27) Mehraei G, Hickox AE, Bharadwaj HM, Goldberg H, Verhulst S, Liberman MC, et al. Auditory brainstem response latency in noise as a marker of cochlear synaptopathy. J Neurosci 2016;36:3755-64.
28) Kumar UA, Ameenudin S, Sangamanatha AV. Temporal and speech processing skills in normal hearing individuals exposed to occupational noise. Noise Health 2012;14:100-5.
29) Vijayasarathy S, Mohan M, Nagalakshmi P, Baraman A. Speech perception in noise, gap detection and amplitude modulation detection in suspected hidden hearing loss. Hearing, Balance and Communication. 2020 (Manuscript under review).
30) Mattys SL, Davis MH, Bradlow AR, Scott SK. Speech recognition in adverse conditions: a review. Lang Cognitive Proc 2012;27:953-78.