ISSN: 2083-389X
eISSN: 2084-3127

J Hear Sci, 2024; 14(2): 33–38
DOI: 10.17430/jhs/191377

SPEECH PERCEPTION IN NOISE IN MALAYALAM-SPEAKING YOUNG ADULTS WITH NORMAL HEARING

Contributions:
A Study design/planning
B Data collection/entry
C Data analysis/statistics
D Data interpretation
E Preparation of manuscript
F Literature analysis/search
G Funds collection

Vipin Ghosh[1ADE], Darshan Devananda[2ACEF], Harisanker S B[3BF], Harshan Kumar[2E]

[1] Department of Audiology, JSS Institute of Speech and Hearing, Mysore, India
[2] Department of Speech-Language Pathology, JSS Institute of Speech and Hearing, Dharwad, India
[3] JSS Institute of Speech and Hearing, Mysore, India

Corresponding author: Darshan Devananda, Department of Speech-Language Pathology, JSS Institute of Speech and Hearing, Kelageri, 570001, Dharwad, India; email: darshandevaslp@gmail.com

Abstract

Introduction: Various types of noise have been used with speech material to assess speech perception in noise (SPIN) abilities. The literature suggests that speech identification varies with different types of background noise, and it has been reported that the target native language and the language of the babbling influence performance. Such efforts in an Indian context have not yet been reported. The aim of the study is to evaluate the speech perception in noise performance of Malayalam-speaking young adults with normal hearing using three different background noises.

Material and methods: A repeated measure research design were adopted with a random sampling method. 30 native Malayalam speakers with normal hearing between the ages of 18 and 25 participated in the study. A standardized sentence list in Malayalam was used as the speech stimulus. Nine lists were chosen and randomly divided so that there were three lists to each background noise. Noises were speech spectrum-shaped noise, non-native language multi-talker babble (Kannada), and native language multi-talker babble (Malayalam). Each successfully repeated keyword received a '1' and each incorrectly repeated word received a '0.' Because each sentence had four important words, each collection of 10 sentences scored a maximum of 40. The percentage of correct answers was determined and further analyzed.

Results: Scores were significantly different in all three different background noises across different SNRs. The highest scores were obtained at +5 dB SNR and the poorest scores at –5 dB SNR. Among the three different background noises, native multi-talker babble (Malayalam) yielded better scores than non-native multi-talker babble (Kannada), followed last by speech spectrum-shaped noise.

Conclusions: The findings of the current study may be attributed to the increased efficacy of speech spectrum noise due to its energetic masking characteristics and the similarity between the two languages in terms of its origin and acoustic-phonetic properties.

Keywords: speech recognition • speech identification in noise • speech perception in noise • SPIN • non-native speech

PERCEPCJA MOWY W SZUMIE U MŁODYCH DOROSŁYCH Z PRAWIDŁOWYM SŁUCHEM MÓWIĄCYCH W JĘZYKU MALAJALAM

Streszczenie

Wstęp: Do oceny percepcji mowy w hałasie (SPIN) wykorzystuje się różne rodzaje szumu tła. Wyniki badań opisane w literaturze przedmiotu sugerują, że identyfikacja mowy w szumie różni się w zależności od rodzaju szumu, w tym od tego, czy jest to szum mowy ojczystej czy obcojęzycznej. W Indiach do tej pory nie były prowadzone tego typu badania. Celem pracy jest ocena percepcji mowy w hałasie u młodych dorosłych ze słuchem w normie, mówiących w języku malajalam, z wykorzystaniem trzech różnych szumów tła.

Materiał i metody: W badaniu zastosowano metodę powtarzanego pomiaru z losowym doborem próby. W badaniu wzięło udział 30 rodzimych użytkowników języka malajalam ze słuchem w normie, w wieku od 18 do 25 lat. Bodźcem mowy były opracowane listy zdań w języku malajalam. Wybrano dziewięć list i podzielono je losowo, tak aby na każdy szum tła przypadały trzy listy. Jako szumy tła wybrano: szum dopasowany do widma mowy, szum mowy obcojęzycznej (kannada) i szum mowy w języku ojczystym (malajalam). Każde prawidłowo powtórzone słowo z listy zdań otrzymywało „1", a każde niepoprawnie powtórzone – „0". Ponieważ każde zdanie zawierało cztery istotne słowa, za każdy zbiór 10 zdań można było uzyskać maksymalnie 40 punktów. Określono procent poprawnych odpowiedzi i poddano dalszej analizie.

Wyniki: Wyniki były znacząco różne dla wszystkich trzech szumów tła przy różnych wartościach SNR. Najwyższe wyniki uzyskano przy SNR +5 dB, a najniższe przy SNR –5 dB. Spośród trzech różnych szumów tła, wyniki dla szumu mowy w języku rodzimym (malajalam) były lepsze niż wyniki dla szumu mowy obcojęzycznej (kannada), a te z kolei były lepsze niż dla szumu dopasowanego do widma mowy.

Wnioski: Zaobserwowano podobieństwo między dwoma językami uwzględnionymi w badaniu (malajalam i kannada), prawdopodobnie ze względu na podobne właściwości akustyczno-fonetyczne obu języków i znajomość języka nierodzimego.

Słowa kluczowe: rozpoznawanie mowy • identyfikacja mowy w szumie • percepcja mowy w szumie • SPIN • mowa obcojęzyczna

Key for abbreviations

ANSI	American National Standards Institute
FFT	fast Fourier transform
LTSSN	long-term speech-shaped noise
MTB	multi-talker babble
NMTB	native multi-talker babble
NNMTB	non-native multi-talker babble
PTA	pure tone averages
SCAP-A	Screening Checklist for Auditory Processing in Adults
SIN	speech-in-noise
SIS	speech identification scores
SNR	signal-to-noise ratio
SPIN	speech perception in noise
SRT	speech recognition thresholds
SSN	speech spectrum noise

Introduction

Speech perception in noise (SPIN) testing can reveal important information about a patient's auditory system. It has the potential to be used in diagnosing and evaluating the hearing system's functional capability, providing clinicians with extremely useful information while needing very little clinical time. A SPIN test can provide important information concerning real-world complaints for these patients. The measurement of speech perception provides useful information in assessing communication difficulties experienced by listeners. The scope of speech perception tests extends to the assessment and monitoring of communication difficulties experienced by listeners [1]. A variety of test materials such as nonsense syllables, monosyllables, bisyllables, and sentences are used to assess the speech perception abilities of individuals [2].

Over the years, different forms of sentence tests have been developed, keeping in mind the perceptual difficulties of those with hearing loss and the language of the individual [3]. Studies reported that the mother tongue of an individual affects his or her perception of speech and that participants consistently have better discrimination scores in their mother tongue compared to other languages [3]. Numerous tests can be used to determine whether or not someone understands speech in a noisy environment. Any listener, particularly those with hearing problems, faces a significant barrier in understanding speech in background noise. Because of the difficulty this exercise presents to listeners, it can provide valuable insight into an individual's capacity to cope with regular everyday listening conditions, which are frequently noisy [4]. Researchers have found that people with hearing impairment need a higher SNR (10–15 dB) than people with normal hearing. With the increase in SNR, a hearing-impaired person's ability to recognize speech increases by about 3% [5]. Therefore, it has been found that the addition of noise to the SIN perception test increases the sensitivity and specificity of the test; by adding multiple noises, the difficulty of the perception increases and the possibility of differentiating people with normal hearing from people with hearing impairment improves [6]. Hence, it is important to have speech material in the mother tongue of an individual.

There have been some attempts in the literature to compare speech perception in the presence of noise across different types of background noises. However, a comprehensive evaluation of differences in speech identification across native language multi-talker babble, non-native language multi-talker babble, and speech spectrum-shaped noise are few in the Indian context.

Various types of noise have been used with speech material to assess speech perception in noise abilities. Earlier findings in a study of English phrase recognition in quiet and two types of maskers, multi-talker babble (MTB) and long-term speech-shaped noise (LTSSN), with varying signal-to-noise ratios for English-, Chinese-, and Korean-native listeners. The test results showed that background noise affected non-native listeners' sentence recognition more than native listeners [5]. Sentence recognition in native- and foreign-language multi-talker background noise has been studied by some researchers [7]. This study looked at speech-in-noise detection when the background noise language was the same as or different from the target speech language. Regardless of the language of the babbling, native English listeners had a harder time understanding English sentences in six-talker babble than in two-talker babble. Furthermore, their results showed that native English listeners were more negatively influenced by English babble than Mandarin Chinese babbling in two-talker babble [7]. These findings show "linguistic interference" as a sort of informational masking on sentence-in-noise detection. Thus, the evidence suggests that speech identification varies with different types of background noise. However, it is reported that the native language and the language of babbling influence the performance. Researchers have also noted the interaction effect of gender and ear laterality in quiet and at different SNRs are not significant [8]. Nevertheless,

Table 1. Demographic details of the participants

Number of participants	30
Age range	18–25 years
Mean age	22.4 years
Male: female distribution	15:15
Education	undergraduate university students

such attempts in the Indian context have not yet been reported in the literature.

Hence, the present study was proposed to evaluate the SPIN performance in Malayalam-speaking young adults with normal hearing abilities using three different background noises: speech spectrum-shaped noise, non-native language (Kannada) multi-talker babble, and native language (Malayalam) multi-talker babble.

The aim of the study is to evaluate the speech perception in noise performance in Malayalam-speaking young adults with normal hearing abilities using three different background noises. To compare the SPIN performance in Malayalam-speaking young adults with normal hearing abilities using different background noises such as speech spectrum-shaped noise, non-native language multi-talker babble, and native language multi-talker babble.

Material and methods

Approval to conduct the study was obtained from the Research Advisory Committee of the JSS Institute of Speech and Hearing in January 2022. The study was carried out in four phases as follows.

Phase 1: Participant selection

A repeated measure research design was adopted with a random sampling method. A total of 30 participants between the ages of 18 to 25 years were chosen. Prior written and oral consent was obtained from all the participants and their capacity to distinguish speech from background noise was tested. The demographic details of the participants are provided below in **Table 1**. All the participants had bilateral normal hearing sensitivity (PTA < 15 dB HL; SRT +10 dB of PTA; SIS > 90%; ANSI, 1996) and normal auditory processing abilities, assessed using Screening Checklist for Auditory Processing in Adults (SCAP-A). All participants were native Malayalam speakers. Participants with any history of middle ear pathology or neurological, psychological, visual, or behavioural problems were excluded.

Phase 2: Preparation of different background noises

PRAAT software was used to record a 2-minute 4-talker speech jumble. The method for recording speech babble was adapted from another study [2]. Native language multi-talker babble was recorded in a classroom setting, with four native Malayalam speakers seated in a circle configuration in the centre of the room using an Omni directional microphone. The distance between the microphone and each speaker's mouth was approximately 30 cm. Speakers were instructed to read a variety of Malayalam newspaper articles at the same time. They were told to keep their speech loud and rate at standard conversational levels. The recorded speech babble was saved in the WAV file on a PC. The level of the recorded voice babble was later standardized to 70 dB SPL using PRAAT.

The same procedure was adapted to record non-native language multi-talker babble. Speakers were instructed to read a variety of Kannada newspaper articles at the same time. For recording speech spectrum-shaped noise, all of the selected audio samples of sentences were concatenated in random order, and a fast Fourier transform (FFT) was conducted separately for each language on these concatenated sentences. To generate back auditory speech noise signal, a reverse FFT with random phase was generated using the obtained spectral values. As a result, the noise generated had a frequency spectrum that was similar to the selected words' long-term average spectrum. The recorded background noises were added to the speech stimuli. The reasoning behind this was that a matched noise would mimic the actual form of noise that would disguise speech in a real-life circumstance [3]. The RMS level of the generated noise was matched to the same level as the sentences.

The Malayalam sentences for the speech stimuli were taken from the sentence list in Malayalam and Telugu [9]. The test consists of 16 lists with 10 sentences each. The phoneme frequency in each list correlated with the overall phoneme frequency from another study [10]. This maintained the phonemic balance throughout all created sentence lists. The same procedure was adopted for adding the three different background noises to the speech stimuli in the present study. A total of nine lists were selected and it was randomly divided so that three lists were added to native talker babble non-native talker babble and speech spectrum-shaped noise respectively. Matlab software (v. R2017a) was used to add three different types of noise. Within the divided list, background noises were added at three different SNR levels (i.e. at +5, 0, and –5 dB SNR) for each list.

Phase 3: Assessing speech perception in noise ability

The experiment was carried out in a well-lit, acoustically treated room. The sentences of each list were randomly presented to each participant through the headphones (Sennheiser HD 202) attached to the personal computer (HP Pavilion core i-3 processor). The output of the headphones was monitored using the sound level meter (B & K-2238, mediator). The stimuli were presented binaurally at 70 dB SPL loudness level (the most comfortable loudness level). The stimuli were delivered in a binaural format. The participants were advised to pay close attention to the sentences and repeat each word. The participants' responses were captured using an audio recorder for further analysis.

Phase 4: Analysis of data

Each successfully repeated keyword received a score of '1', whereas each mistakenly repeated word received a score of '0.' Any response with a glaring error was regarded as

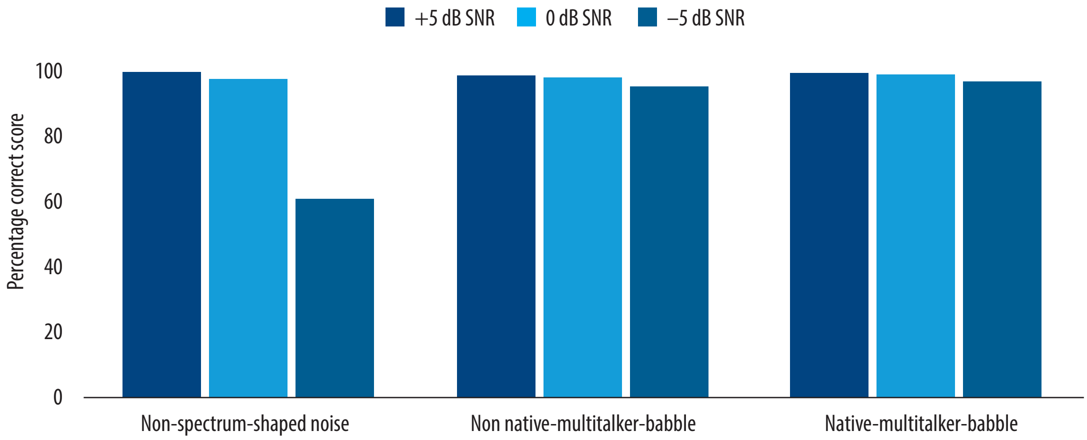

Figure 1. Mean percentage scores with different background noises

Table 2. Mean percentage and standard deviation of scores obtained at different SNRs across the background noises

SNR levels	Speech spectrum noise	Non-native multi-talker babble	Native multi-talker babble
+5 dB SNR	99.25 (*SD* = 5.20)	99.25 (*SD* = 2.84)	99.67 (*SD* = 1.34)
0 dB SNR	97.83 (*SD* = 1.63)	98.58 (*SD* = 1.34)	99.25 (*SD* = 1.08)
−5 dB SNR	61.15 (*SD* = 13.51)	95.58 (*SD* = 3.70)	97.25 (*SD* = 3.17)

Table 3. Friedman test results for comparison of speech perception in noise at different SNRs

Background noises	Friedman test values across the SNRs
Speech spectrum-shaped noise	$\chi^2 = 53.38$, $p < 0.001$
Non-native multi-talker babble	$\chi^2 = 28.22$, $p < 0.001$
Native multi-talker babble	$\chi^2 = 19.31$, $p < 0.001$

Table 4. Friedman test results for comparison of speech perception in noise at different background noises

SNR levels	Friedman test values across the background noises
+5 dB SNR	$\chi^2 = 4.85$, $p = 0.09$
0 dB SNR	$\chi^2 = 1.09$, $p = 0.58$
−5 dB SNR	$\chi^2 = 51.71$, $p < 0.001$

Table 5. Wilcoxon signed rank test values across the background noises between the SNRs

SNR comparison	Speech spectrum noise	Non-native multi-talker babble	Native multi-talker babble
0 dB and +5 dB	$Z = -2.14$, $p < 0.001$	$Z = -1.49$, $p = 0.14$	$Z = -1.25$, $p = 0.21$
+5 dB and −5 dB	$Z = -4.79$, $p < 0.001$	$Z = -4.18$, $p < 0.001$	$Z = -3.21$, $p < 0.001$
0 dB and −5 dB	$Z = -4.79$, $p < 0.001$	$Z = -3.46$, $p < 0.001$	$Z = -3.16$, $p < 0.001$

an incorrect response solely. Each sentence contained four keywords and hence, each list with 10 sentences received a maximum score of 40. Scores obtained were converted to percentage correct scores and further statistical analysis was carried out using the SPSS version 21 statistical software.

Results

In the present study, speech-in-noise scores obtained using different types of noises were compared across three different SNRs: +5 dB SNR, 0 dB SNR, and −5 dB SNR (**Figure 1**). Shapiro–Wilk test was used to assess normality and found that data were non-normally distributed. Friedman test was carried out for further analysis. Further, post hoc analysis was carried out using the Wilcoxon signed ranks test to compare between the SNRs and background considered.

The mean percentage correct scores thus obtained in the presence of speech spectrum noise (SSN), non-native multi-talker babble (NNMTB), and native multi-talker babble (NMTB) at +5 dB SNR, 0 dB SNR, and −5 dB SNR are shown in **Table 2**. A Friedman test for comparison of speech perception in noise at different SNRs (using speech spectrum-shaped noise, NNMTB, and NMTB as background noise) showed significant differences, as shown in **Table 3**.

Similarly, the speech recognition scores were then compared across different types of noises at equivalent SNRs. The comparison of SPIN across different types of background noise at +5 dB SNR and 0 dB SNR showed no significant difference. However, the comparison of SPIN across different types of background noise at –5 dB SNR showed a significant difference as shown in **Table 4**.

Further, Wilcoxon signed ranks test was carried out to test significance between the categories. The results revealed that there was a significant difference between all three categories while using speech spectrum-shaped noise but the similar trend was not seen in non-native multi-talker babble and native multi-talker babble as shown in **Table 5**.

Speech perception scores were better at +5 dB SNR than 0 dB SNR followed by –5 dB SNR using non-native multi-talker babble, native multi-talker babble and speech shaped noise. Thus, in all three noise conditions, the highest scores were obtained at +5 dB SNR and the poorest scores at –5 dB SNR.

Discussion

The present study aimed to examine the scores of Malayalam-speaking young adults with normal hearing abilities for speech perception in noise utilizing speech spectrum-shaped noise, non-native language multi-talker babble, and native language multi-talker babble. The percentage of correct answers was determined and then further analysed. The results of the present study are in agreement with the previous studies [1,12,13].

Sentence recognition in native- and foreign-language multi-talker background noise was reported in an earlier study [9]. The aim of the study was to determine whether the adverse effect of background speech is due to the linguistic content or to the acoustic characteristics of the speech masker. According to the results, in every situation, better target sentence perception was produced by greater SNRs. As the level of noise increases relative to the target, the ability to hear speech declines. The results of the present study are in agreement with the previous studies [1,12,13] in all three noise conditions, used in the present study, i.e., speech spectrum-shaped noise, non-native multi-talker babble, and native multi-talker babble, the highest scores were obtained at +5 dB SNR and the poorest scores at –5 dB SNR. The higher the signal-to-noise ratio, the more intense will be the signal. When the speech signal has a higher intensity than the background noise, it helps in better identification of speech. But as the signal intensity decreases, background noise will mask the signal, which in turn affects speech perception.

In the present study among all the three different background noises, native multi-talker babble yielded better scores than non-native multi-talker babble followed by speech spectrum-shaped noise. At higher SNRs however, there was no difference across the different background noise. But, as the SNR decreased to –5 dB SNR, speech perception scores using multi-talker babble were better than speech spectrum-shaped noise. These improved scores for multi-talker babble were reported by an earlier study [11]. Better scores with native multi-talker babble than non-native multi-talker babble is also reported in the literature [9]. This can be attributed to the similarity between the two languages included in the study as they have a common origin. The similarity between the two languages' acoustic-phonetic properties or acquaintance with the non-native tongue would have determined this. Additionally, a cognitive component would come into play, where individuals would get more easily distracted by a foreign language in the background as opposed to a familiar one.

Thus, the present study showed that the performance of speech recognition was better when multi-talker babble was used when compared to the speech-shaped noise. Speech spectrum-shaped noise provides the same long-term average signal-to-noise ratio (SNR) in each frequency band. Hence, there is no gap within the noise generated and it provides energetic masking which effectively masks the speech signal. Moreover, when the multi-talker babbles were compared, the NNMTB was found to be a more effective masker than the NMTB. This finding may be attributed to the similarity in the origin (Dravidian) of both languages used in the current study. Moreover, the performance was found to be progressively poorer with reducing SNRs.

Conclusions

In the current study, for all three background noises speech recognition scores were better at +5 dB SNR and poorer at –5 dB SNR. Among the background noises, speech perception was found to be better when using native multi-talker babble (Malayalam) than non-native (Kannada) multi-talker babble, followed last by speech spectrum-shaped noise. Native talker-babble may provide acoustic cues which ease the perception of speech in noise compared to other background noises. The similarity between the two investigated languages – because of their common Dravidian origin – could be the cause of this difference. Alternatively, listeners may get easily distracted by a foreign language in the background as opposed to a familiar one. The present study findings warrant further research on the influence of native and non-native multi-talker babble on speech in noise perception using various native and non-native language combinations.

Limitations

The study was carried out using a small sample size. The study considered only young normal adults.

Future directions

The present study was based on a small sample size, and the study could be carried out with a larger population and with other age groups. It could also be studied using a clinical population.

References

1. Gundmi A, Himaja P, Dhamani A. Effectiveness of multi-talker babble over speech noise and its implications: a comparative study. Indian J Otol, 2018; 24(2): 88. https://doi.org/10.4103/indianjotol.indianjotol_24_18
2. 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(4): 2395–405. https://doi.org/10.1121/1.1784440
3. Prosser S, Turrini M, Arslan E. Effects of different noises on speech discrimination by the elderly. Acta Otolaryngol, 1991; 111 (Suppl 476): 136–42. https://doi.org/10.3109/00016489109127268
4. Taylor B. The acceptable noise level test as a predictor of real-world hearing aid benefit. Hear J, 2008; 61(9): 39–40, 42. https://doi.org/10.1097/01.HJ.0000339505.97796.a9
5. McArdle R, Hnath-Chisolm T. Speech audiometry. In: Katz J, Chasin M, English K, Hood LJ, Tillery KL, editors. Handbook of Clinical Audiology. 7th edition. Philadelphia: Wolters Kluwer; 2015, p. 61–76.
6. Bamiou DE, Iliadou VV, Zanchetta S, Spyridakou C. What can we learn about auditory processing from adult hearing questionnaires? J Am Acad Audiol, 2015; 26(10): 824–37. https://doi.org/10.3766/jaaa.15009
7. Jin SH, Liu Ch. English sentence recognition in speech-shaped noise and multi-talker babble for English-, Chinese-, and Korean-native listeners. J Acoust Soc Am, 2012; 132(5): EL391–7. https://doi.org/10.1121/1.4757730
8. Jain C, Kodanath S, Vimal BM, Suresh V. Influence of native and non-native multitalker babble on speech recognition in noise. Audiol Res, 2014; 4(1): 89. https://doi.org/10.4081/audiores.2014.89
9. Amiri M, Mohammadi M, Razavi SZ, Abdollahi H, Iran Pour Mobarakeh Z. Normative data for speech-in-noise perception test in young adults with normal hearing: gender and ear laterality effects. Aud Vestib Res, 2023; 32(1): 64–9. https://doi.org/10.18502/avr.v32i1.11323
10. Ramakrishna BS, Nair KK, Chiplimear VN, Atal BS, Ramachandran V, Subramanian R. Some aspects of the relative efficiencies of Indian languages. Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, 1962.
11. Van Engen KJ, Bradlow AR. Sentence recognition in native- and foreign-language multi-talker background noise. J Acoust Soc Am, 2007; 121(1): 519–26. https://doi.org/10.1121/1.2400666
12. Shojaei E, Ashayeri H, Jafar Z, Zarrin Dast MR, Kamali K. Effect of signal to noise ratio on the speech perception ability of older adults. Med J Islam Repub Iran, 2016; 30: 342.
13. Razak A, Thurston EJ, Gustainis LE, Kidd G, Swaminathan J, Perrachione TK. Talker identification in three types of background noise. J Acoustic Soc Am, 2017; 141(5): 4039. https://doi.org/10.1121/1.4989326