Review Article

Central Representation of Speech-in-Noise Perception: A Narrative Review

Seyede Faranak Emami[1,2], Elnaz Shariatpanahi[1,3*]

[1.] *Hearing Disorder Research Center, Hamadan University of Medical Sciences, Hamadan, Iran*
[2.] *Department of Audiology, School of Rehabilitation, Hamadan University of Medical Sciences, Hamadan, Iran*
[3.] *Department of Otorhinolaryngology, School of Medicine, Hamadan University of Medical Sciences, Hamadan, Iran*

Citation: Emami SF, Shariatpanahi E. Central Representation of Speech-in-Noise Perception: A Narrative Review. Aud Vestib Res. 2023;32(3):166-73.

https://doi.org/10.18502/avr.v32i3.12932

Highlights

- Speech-in-Noise (SIN) perception is not the same as in quiet
- Speech-in-quiet perception tests cannot assess actual perceptual abilities of people
- SIN perception tests should be used routinely as a confirmatory audiological test

Article info:
Received: 29 Nov 2022
Revised: 15 Dec 2022
Accepted: 01 Jan 2023

*** Corresponding Author:**
Hearing Disorder Research Center, Hamadan University of Medical Sciences, Hamadan, Iran.
eshariatpanahi1@gmail.com

ABSTRACT

Background and Aim: Speech-in-Noise (SIN) perception is one of the most important issues processed by human listeners. The purpose of speech tests is to determine the perceptual abilities of people in real life conditions; therefore, can speech tests in quiet be a valuable measure of this ability? Is the central representation of speech-in-quiet and SIN perception the same? This review study aimed to investigate the central representation of SIN perception in healthy individuals aged 14 to 60 years.

Recent Findings: Central representation of SIN perception is influenced by various peripheral factors and includes several neural processes. All auditory nerve fibers are stimulated by speech and noise. Low-frequency sounds play a much more important role than high-frequencies. The auditory nerve fibers that are stimulated by speech, respond only to fundamental frequencies (F0). The degree of neural synchronization that increases by noise and causes the simultaneous activity of these fibers, develops auditory processing. Large areas of the auditory cortex and its external parts (parietal, premotor, and mirror neurons) are stimulated. Larger groups of cortical nerve fibers are used for speech signals of the same family with significant ecological importance.

Conclusion: Central representation of SIN perception is not the same as in quiet. Speech perception tests in quiet cannot assess real-life perceptual abilities of people. SIN tests should be used routinely as a practical confirmatory test in audiology clinics. It is very necessary that the list of words and sentences required for SIN perception tests be prepared for different languages.

Keywords: Speech; noise; perception; auditory cortex; brainstem

Introduction

Speech is a complex function; Speech-in-Noise (SIN) perception requires the listener to process the speaker's message so that he/she can quickly code it and adapt to the changes in speaker's expression and the conditions that affect the meaning of the message [1]. SIN perception is the result of the activity of the auditory cortex which has a role in speech perception and production [2-4], and facilitates the perception of meaning for the listener, and creates certain methods according to which the listener has a better understanding of the message, e.g. reading aloud or saying something out loud for better perception [3]. Speech perception and production are in direct interaction with each other; we first understand the meaning of the sound. Then, to produce that sound, our speech organs are activated and make sounds, and again we convert what was produced and expressed by the speech organs into meaningful codes for the perception of the produced sound. This two-way circle reduces the neural activities for vocalization and pronunciation in the central nervous system [2, 3]. For example, small children say the vowel parts of the word or express the sound of a word, which is the same as pitch changes or the expression of speech vowels, and they cannot pronounce consonants properly [3-5]. As another example, in melodic languages, adults need changes in the pitch or tone of speech to understand speech, because each of these changes can alter the meaning of the words [4, 6, 7].

SIN perception disorders are usually caused by a decrease in temporal resolution or frequency resolution, or an inability to separate signals that are received bilaterally, depending on the individual and their specific hearing needs [5, 8]. One of the most critical aspects of everyday speech communication in the real world is that people interact with different groups in different areas and noisy situations [6]. Alviandi et al. reported that speech perception test and speech discrimination test in quiet [9] and in noise reflect different areas of auditory function [10, 11]. Wilson and Watts stated that SIN tests provide valuable information about the patient's hearing problems [10]. Kujawa and Liberman proposed the hidden hearing loss, probably caused by the occurrence of several abnormal functions, and suggested that SIN perception tests are very important for its diagnosis [11]. In clinical practice, the patient's speech discrimination is evaluated in quiet [9], and SIN perception tests are rarely used [12]. The findings of studies on SIN perception are related to the methods and materials used for assessment [13-16]. Many studies have indicated the importance of SIN perception tests for evaluating the ability of understanding speech in noise, language skills, communication and learning [10-12, 17]. This review study aimed to investigate the central representation of SIN perception in healthy individuals aged 14 to 60 years.

Methods

In this review study, to find the related studies published in English from 1985 to 2022, a search was conducted in Medline, Springer, Google Scholar, Web of Science, Scopus, PubMed, and Oxford and Lippincott databases. The search yielded 106 articles, of which 64 were selected based on the inclusion criteria. The inclusion criteria were a study on participants aged 14–60 years with a normal hearing, published after 1985, regarding the peripheral factors affecting SIN perception. The studies that did not investigate the SIN perception were excluded.

Results

The SIN perception is affected by several peripheral factors [1, 6, 13-18]. Some of them, which were extracted from the reviewed studies, are provided in this section.

Unfavorable presentation of speech

The ability to recognize the word increases in unpleasant emotional listening states, maybe the reason is the increase of selective attention. The emotional way of speaking affects the accuracy of word recognition, especially in noisy competitive situations where the intensity of the speaker's voice is low [18]. In other words, the emotional content of speech is related to the level of perception. Since it is more difficult to understand speech in the presence of noise than in quiet [6, 9, 10, 12], it is important to consider the emotional load of the speech when creating word lists for the speech comprehension test in noise, which certainly increases the difficulty of the listener's hearing tasks [13, 18].

Presence of speakers with different characteristics

The presence of speakers with different characteristics [16, 17], such as gender [9, 13], pronunciation [6, 12], age [1, 8, 12], speech rate [3, 4], use of indigenous words [4, 6], and understanding vowel sounds [2, 3], is one of the main problems in speech perception. Monotonous speech reduces the accuracy of lip reading [19-21], phoneme identification, word comprehension, and memory [20, 22]. The rate of speech or the number of words spoken within a time by the speaker varies between different people [20, 23]. A typical speech rate is between 160 and 200 words per minute (wpm) [24]. The acceptable level of noise is influenced by the speech rate of the speaker [25]. When the speech rate is high (e.g. 288 wpm), background noise is less acceptable than when the speech is at a normal (e.g. 186 wpm) or slow rate (e.g. 130 wpm) [23, 24]. There is a significant difference in the ability to understand SIN between the listeners, which can pose significant challenges for the listeners with lower abilities [26]. People who have the ability to discover more, even in the conditions where the influencing factors of the speakers are highly diverse, obtain more phonetic and audio lexical information in competitive conditions and the presence of noise [27]. A greater difference between the speaker's fundamental frequency of voice and the background noise frequency improves the perception of speech in noise. Greater similarity between the two reduces the listener's ability to understand speech [28]. Diversity of speakers has more negative effects on people with hearing loss or cochlear implants [29], non-native listeners [30] and the elderly [21, 31].

Difference in ear maturation

The right ear matures at the age of 10 and acts just like adults' ear, while the left ear matures at the age of 13-14 years. SIN perception ability decreases significantly with ageing [31], which is associated with thinning and reduction of brain activity in the right superior temporal cortex and an increase in the level of stimulation and response to noise in the left anterior temporal lobe [8, 12].

Type and level of background noise

The type and level of background noise can interfere with the physical and spectral characteristics of the target (speaker) signal or voice [32]. The spectrum of the speaker's speech, and the challenges created by the speaker's individual characteristics, the fluency of the sentences or the text that speaker reads, and high similarity between the speaker's voice and the voices of the people who are speaking with at the same time, causes a weaker speech perception [33]. White noise or a noise with a wide frequency spectrum increases the speed of correct understanding of new words and improves the ability to understand speech in noise [34-36]. In fact, the modification and control of the dopaminergic neuronal pathway by white noise (at all frequencies) may increase the memory and the ability to retain the characteristics of sounds with different frequencies, which helps to understand speech in noise [35].

Bilingualism

The ability for syllables without linguistic meaning or the ability to discover the tonality of the language is more in bilinguals than in monolinguals, while in meaningful speech and language communication, the ability of bilinguals is weaker [13, 16]. Bilingual listeners need stronger signals or sounds to understand speech in noise. Bilingual listeners with normal hearing have the ability to understand speech in a mild Signal-to-Noise Ratio (SNR; difference in the intensity of the speaker's voice and the background noise) similar to people with hearing loss. The problems of bilinguals are related to their knowledge of second, third, or fourth languages [14]; it depends on the age of learning the second or multiple languages. In fact, there is a direct relationship between the perception of speech in noise and the age of learning a second or multiple languages [24]. However, if second language learning occurs at an early age, they will perform better in speech perception [27, 29].

Musical training

Musical training significantly develops the ability of SIN perception [37-40], and improves intellectual ability [2, 3]. Playing a musical instrument improves the cognitive and mental abilities of people, especially in old age, but this correlation does not exist in musicians who do not actively play a musical instrument; i.e. listening cannot develop the mental skill of playing music [38]. In the human brain, with the onset of sound, a three-dimensional map or image is created, whose clarity can be improved by education or therapy. Brainstem responses to onset points or the onset moment of a sound stimulus do not seem to be improved with training unless the individual be at very young age. In other words, in children, the brainstem responses to the onset point of a sound stimulus are improved by practice and training [2, 3, 38, 40]. If the musical practices and training are not available, the involvement of the cerebral cortex and brainstem will not stop and will improve with the increase of age, but its speed and accuracy will be low [3, 29, 37].

The central nervous system displays the acoustic structure of speech; each sound has a multifaceted image in mind. The bottom-up process (peripheral auditory system to the cingulate cortex) requires the listener to analyze the phonological and acoustic information of the speech sounds and combine them to understand the speech clearly [37], but, when bottom-up signals cannot be detected because of the noise that masks them, listeners usually use top-down processing (cingulate cortex to peripheral auditory system) to benefit from other components of communication, including the analysis of information to discover meaning and concepts. Therefore, listeners use their lexical information and linguistic knowledge to correct the content they did not understand and fill in the blanks or misunderstood parts [3, 39, 40].

Discussion

The purpose of the present study was to review the central representation of speech in noise in healthy individuals aged 14 to 60 years. Because in real environments, listeners live in a world of competing signals and noise [41-45], common speech tests in quiet, such as speech discrimination score, lack the required validity [9, 11, 41]. The central representation of SIN perception includes several neural processes:

All auditory nerve fibers are stimulated by speech and noise: when speech materials are presented in noise, it is usually observed that the noise elicits the responses of whole neurons [2, 39], even those neurons that are stimulated by speech. Its consequence is the reduction of spatial information about the positive and negative peaks of sounds [2, 3, 40]. Reduced temporal fine structure (changes in the amplitude and frequency of a sound perceived over time) does not allow listeners to obtain information about the pitch or resonance of sounds, which are crucial in distinguishing sounds and noises [3, 45]. In a quiet environment, the envelope of signals (boundary within which the signal is contained) manage temporal fine structure [3, 4]. If the temporal fine structure information is not provided in the desired way, it will be very difficult to recognize the low frequencies of speech [42-44], and if the low frequencies are unrecognizable, it will not be possible to imagine the words spatially or create a multifaceted form or auditory image from a sound. For example, the imagination of the word "water" or its auditory image in the mind of each person is different from another; when hearing the word water, different auditory images may be created for people including the image of water in a glass, running water in a sea, river water, spring water, or a marshland water [2, 3, 40].

Low frequencies play a much more important role than high frequencies: vowels play the main role in word recognition and SIN perception [2, 3, 43, 44]. In noisy environments, the identity of vowels is generally preserved and facilitate the role of consonants in understanding the meaning and concepts of speech [41]. Paying attention to this phonetic feature of vowels is very important when preparing materials or word lists for speech comprehension tests in the presence of noise and can affect the validity of the results [42]. Since the sound energy level and the frequency range of speech vowels are different [43], it seems that choosing the same vowels for each independent list of phonetically similar monosyllabic words can cause the least bias in the test results [42]. Vowels transfer phonetic information and the information related to speech pitch, such as tone and stress [2, 3, 43] and play an important role in distinguishing the signal from the surrounding noise, the formation of auditory images, and the separation of different sounds [2, 3, 45, 46].

Auditory nerve fibers stimulated by speech respond only to F0: during listening in quiet, auditory nerve fibers respond to both fundamental (F0) and non-fundamental frequency sounds [28, 41, 42], while in situations where it is necessary to distinguish signal from noise, the auditory nerve fibers respond only to F0 [2, 3, 39] and send limited information to the brain for processing the meaning [2, 3, 39]. Since F0 does not convey the auditory information necessary to understand the message content [28, 39], the auditory cortex does not respond to this part of the signal [2]. Thus, an important feature of F0 perception is that native listeners are able to recognize a certain vocal message as single concept regardless of the speaker's F0, which differs between men (100 Hz), women (200 Hz), and infants (up to 400 Hz) [2, 3].

Neural synchronization improves and develops auditory processing: The phase-locking degree or the synchronization speed of auditory neurons is related to low-frequency sounds [45-47]. Brainstem neurons are temporally precise; their synchronization improves with low-frequency sounds [48]. In fact, the temporal design of the responses of the fibers in the auditory nerve and cochlear nucleus to the geniculate nucleus is periodic and their oscillation frequency is synchronized with F0. [3, 4]. Neural synchronization has an important role in the transmission of sensory messages from the thalamus to the cerebral cortex [49]. It is possible that improved neural synchronization of auditory brain and brainstem equally speed up the transmission of information to the next stations of the auditory system for auditory processing [40]. The brainstem seems to respond to pitch discrimination and synchronization stimuli with fast repetition rates (>~100 Hz) [50]. Pitch is the perceptual equivalent of F0 [4]. Pitch perception is essential for understanding music and speech, as well as for SIN perception [51]. Pitch conveys phonetic features and rhythmic aspects of speech, such as accent, emphasis (in European languages), and speech tone information (in tonal languages such as Chinese, Vietnamese, and Thai) [3, 51].

The auditory cortex and its peripheral regions are stimulated during SIN perception: speech perception is bilaterally and symmetrically organized in the superior temporal gyrus [52]. The superior temporal sulcus is also an important site for representing and/or processing phonological information [12, 53]. Responses to natural sound stimuli are significantly larger and more widespread throughout the superior temporal sulcus, particularly in the right hemisphere [54, 55]. Unrelated or unexpected emotional stimuli attract one's attention and improve speech perception through stimulation of left medial temporal gyrus, which can contribute to false word recognition [53]. The inferior temporal lobe seems involved in the retrieval of meaning and linguistic knowledge from auditory signals [54, 56]. The anterior temporal lobe seems to create a coherent and integrated form of semantic information based on the information it receives from other senses [55]. The posterior temporal sulcus seems to be a sound-selective cortex and acts as an important gateway that conveys speech information to other brain networks that have a role in semantic processes and recalling events [56, 57]. The insula in the left hemisphere is precisely stimulated when F0 sounds convey phonetic information of speech to a native speaker, while the insula of the right hemisphere is stimulated when F0 sounds convey phonetic information of non-speech sounds [58, 59]. The cingulo-opercular network is involved in top-down mechanism of attention and control [59]. Responses in the ventral premotor cortex during a vowel-identification task are improved when the SNR decreases; such responses may enhance SIN perception [60]. In the field of speech perception, it is not clear whether mirror neurons or the medial frontal cortex have a precise and essential role in SIN perception [58, 61]. Mirror-neuron brain areas play a causal role in copying the topography of body movement, and are involved in low-level processing of activities (e.g. recognizing the types of grip) but have no role in interpretation of high-level activities (e.g. inferring actors' intention) [62, 63]. The inferior frontal gyrus improves people's precision in repeating unfamiliar foreign speech sounds, and matches perceived speech to produced speech [56]. The posterior portion of the left planum temporale is activated during picture naming and exhibits length effects, frequency effects, and has a time-course of activation, that is consistent with the phonological encoding stage of naming [63]. The planum temporale, while often thought to be an auditory area, also activates in response to sensory input from other modalities. For example, silent lipreading has been shown to activate auditory cortex in the vicinity of the planum temporale [52]. Since it is more difficult to understand speech in noise than in quiet [50], less intelligible stimuli activate temporo-parietal regions related to SIN perception [53], while emotionally negative stimuli activate areas associated with emotional processing and selective attention [18].

A larger population of cortical neurons is recruited: In the auditory cortex, dissimilar structures of speech signals can be used by different types of neural processing and/or groups [50, 52]. Therefore, in noisy competitive situations, a greater or different set of cortical neurons is activated for familiar speech sounds with significant ecological importance [5]. Similarly, in the human auditory cortex, there is a tonotopic map and a phonemotopic map. The tonotopic map is a linear sum of responses to the component formant stimuli, formed in the vocal tract [3]. The phonemotopic map or a spatial cortical representation of phonemes [2, 5], reflects the additional acoustic components, including the harmonic and formant structures of speech; the harmonics are produced by vocal cords [56].

Conclusion

Since everyday conversations and communication are created in the presence of background noises, and considering that the central representation of speech-in-noise perception is not the same as in quiet, speech perception tests in quiet do not assess real life perceptual abilities of people; however, unfortunately, these tests in many audiology clinics are still used in quiet. For example, the speech discrimination score is one of the routine tests, while speech in noise tests are not used as a routine and confirmatory test. Therefore, considering that Iran is one of the countries with diverse cultures and languages, it is very necessary that the list of words and sentences required for the speech perception tests in noise be prepared for different languages, and these tests be used as reliable tools in the prevention, diagnosis and treatment of hearing problems.

Ethical Considerations

Funding

There are no Funding sources for our work.

Authors' contributions

In writing and editing this article, the contribution of the authors is equal.

Conflict of interest

There is no conflict of interest.

References

[1] Chandni J, Vipin Ghosh PG, Chetak KB, Aishwarya L. Maturation of speech perception in noise abilities during adolescence. Int J Pediatr Otorhinolaryngol. 2020;139:110459. [DOI:10.1016/j.ijporl.2020.110459]

[2] Scott SK, Sinex DG. Speech. In: Rees A, Palmer A, editors. The Oxford handbook of auditory science: the auditory brain. 1st ed. New York: Oxford university press; 2010. p. 193-214.

[3] Abrams D A, Kraus N. Auditory Pathway Representations of Speech Sounds in Humans. In: Katz J, Chasin M, English K, Hood LJ, Tillery KL, editors. Hand book of Clinical Audiology. 7th ed. Philadelphia: Wolters Kluwer Health. 2015. p. 527-44.

[4] Young ED. Level and spectrum. In: Rees A, Palmer A, editors. The Oxford handbook of auditory science: the auditory brain. 1st ed. New York: Oxford university press; 2010. p. 93-124.

[5] Meyer BT, Brand T, Kollmeier B. Effect of speech-intrinsic variations on human and automatic recognition of spoken phonemes. J Acoust Soc Am. 2011;129(1):388-403. [DOI:10.1121/1.3514525]

[6] Angwin AJ, Wilson WJ, Ripollés P, Rodriguez-Fornells A, Arnott WL, Barry RJ, et al. White noise facilitates new-word learning from context. Brain Lang. 2019;199:104699. [DOI:10.1016/j.bandl.2019.104699]

[7] Shen J, Anderson MC, Arehart KH, Souza PE. Using Cognitive Screening Tests in Audiology. Am J Audiol. 2016;25(4):319-31. [DOI:10.1044/2016_AJA-16-0032]

[8] Peelle JE, Wingfield A. The Neural Consequences of Age-Related Hearing Loss. Trends Neurosci. 2016;39(7):486-97. [DOI:10.1016/j.tins.2016.05.001]

[9] Alviandi W, Bashiruddin J, Bramantyo B, Rizky F. Words in noise audiometry in adult subjects with normal hearing. Oto Rhino Laryngologica Indonesiana. 2020;50(1):9-15. [DOI:10.32637/orli.v50i1.332]

[10] Wilson RH, Watts KL. The Words-in-Noise Test (WIN), list 3: a practice list. J Am Acad Audiol. 2012;23(2):92-6. [DOI:10.3766/jaaa.23.2.3]

[11] Kujawa SG, Liberman MC. Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci. 2009;29(45):14077-85. [DOI:10.1523/JNEUROSCI.2845-09.2009]

[12] Vaden KI Jr, Kuchinsky SE, Ahlstrom JB, Dubno JR, Eckert MA. Cortical activity predicts which older adults recognize speech in noise and when. J Neurosci. 2015;35(9):3929-37. [DOI:10.1523/JNEUROSCI.2908-14.2015]

[13] Chupani J, Javanbakht M, Lotfi Y. Central auditory processing in bilinguals. Aud Vestib Res. 2021;30(3):160-6. [DOI:10.18502/avr.v30i3.6529]

[14] Onoda RM, Pereira LD, Guilherme A. Temporal processing and dichotic listening in bilingual and non-bilingual descendants. Braz J Otorhinolaryngol. 2006;72(6):737-46. [DOI:10.1016/S1808-8694(15)31040-5]

[15] Sanayi R, Mohamadkhani G, Pourbakht A, Jalilvand L, Jalayi S, Shokri S. Auditory temporal processing abilities in early azari-persian bilinguals. Iran J Otorhinolaryngol. 2013;25(4):227-32.

[16] Krizman J, Bradlow AR, Lam SS, Kraus N. How bilinguals listen in noise: linguistic and non-linguistic factors. Biling (Camb Engl). 2017;20(4):834-43. [DOI:10.1017/S1366728916000444]

[17] Aarabi S, Jarollahi F, Badfar S, Hoseinabadi R, Ahadi M. Speech perception in noise mechanisms. Aud Vest Res. 2016;25(4):221-6.

[18] Olano MA, Elizalde Acevedo B, Chambeaud N, Acuña A, Marcó M, Kochen S, et al. Emotional salience enhances intelligibility in adverse acoustic conditions. Neuropsychologia. 2020;147:107580. [DOI:10.1016/j.neuropsychologia.2020.107580]

[19] Rosenblum LD, Yakel DA, Green KP. Face and mouth inversion effects on visual and audiovisual speech perception. J Exp Psychol Hum Percept Perform. 2000;26(2):806-19. [DOI:10.1037//0096-1523.26.2.806]

[20] Yakel DA, Rosenblum LD, Fortier MA. Effects of talker variability on speechreading. Percept Psychophys. 2000;62(7):1405-12. [DOI:10.3758/BF03212142]

[21] Brungart DS, Simpson BD, Ericson MA, Scott KR. Informational and energetic masking effects in the perception of multiple simultaneous talkers. J Acoust Soc Am. 2001;110(5 Pt 1):2527-38. [DOI:10.1121/1.1408946]

[22] Chang YP, Fu QJ. Effects of talker variability on vowel recognition in cochlear implants. J Speech Lang Hear Res. 2006;49(6):1331-41. [DOI:10.1044/1092-4388(2006/095)]

[23] Tamati TN, Pisoni DB. Non-native listeners' recognition of high-variability speech using PRESTO. J Am Acad Audiol. 2014;25(9):869-92. [DOI:10.3766/jaaa.25.9.9]

[24] Wingfield A, Poon LW, Lombardi L, Lowe D. Speed of processing in normal aging: effects of speech rate, linguistic structure, and processing time. J Gerontol. 1985;40(5):579-85. [DOI:10.1093/geronj/40.5.579]

[25] Tiffin S, Gordon-Hickey S. The Influence of Speech Rate on Acceptable Noise Levels. J Am Acad Audiol. 2018;29(7):596-608. [DOI:10.3766/jaaa.16159]

[26] Nabelek AK, Freyaldenhoven MC, Tampas JW, Burchfiel SB, Muenchen RA. Acceptable noise level as a predictor of hearing aid use. J Am Acad Audiol. 2006;17(9):626-39. [DOI:10.3766/jaaa.17.9.2]

[27] Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Reading and subcortical auditory function. Cereb Cortex. 2009;19(11):2699-707. [DOI:10.1093/cercor/bhp024]

[28] Wang X. The harmonic organization of auditory cortex. In: Lopez-Poveda EA, Palmer AR, Meddis R, editors. The Neurophysiological Bases of Auditory Perception. 1st ed. New York: Springer Science & Business Media; 2010. p. 211-22. [DOI:10.1007/978-1-4419-5686-6_20]

[29] Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. Philos Trans R Soc Lond B Biol Sci. 1992;336(1278):367-73. [DOI:10.1098/rstb.1992.0070]

[30] Alqattan D, Turner P. The effect of background noise on speech perception in monolingual and bilingual adults with normal hearing. Noise Health. 2021;23(110):67-74.

[31] Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. J Neurosci. 2012;32(41):14156-64. [DOI:10.1523/JNEUROSCI.2176-12.2012]

[32] Ruggles D, Shinn-Cunningham B. Spatial selective auditory attention in the presence of reverberant energy: individual differences in normal-hearing listeners. J Assoc Res Otolaryngol. 2011;12(3):395-405. [DOI:10.1007/s10162-010-0254-z]

[33] Meyer J, Dentel L, Meunier F. Speech recognition in natural background noise. PLoS One. 2013;8(11):e79279. [DOI:10.1371/journal.pone.0079279]

[34] Strait DL, Parbery-Clark A, O'Connell S, Kraus N. Biological impact of preschool music classes on processing speech in noise. Dev Cogn Neurosci. 2013;6:51-60. [DOI:10.1016/j.dcn.2013.06.003]

[35] Bradlow AR, Kraus N, Hayes E. Speaking clearly for children with learning disabilities: sentence perception in noise. J Speech Lang Hear Res. 2003;46(1):80-97. [DOI:10.1044/1092-4388(2003/007)]

[36] Temple E, Poldrack RA, Protopapas A, Nagarajan S, Salz T, Tallal P, et al. Disruption of the neural response to rapid acoustic stimuli in dyslexia: evidence from functional MRI. Proc Natl Acad Sci U S A. 2000;97(25):13907-12. [DOI:10.1073/pnas.240461697]

[37] Parbery-Clark A, Tierney A, Strait DL, Kraus N. Musicians have fine-tuned neural distinction of speech syllables. Neuroscience. 2012;219:111-9. [DOI:10.1016/j.neuroscience.2012.05.042]

[38] Hanna-Pladdy B, MacKay A. The relation between instrumental musical activity and cognitive aging. Neuropsychology. 2011;25(3):378-86. [DOI:10.1037/a0021895]

[39] Oxenham AJ. Pitch perception. J Neurosci. 2012;32(39):13335-8. [DOI:10.1523/JNEUROSCI.3815-12.2012]

[40] Hartley DEH, King AJ. Development of the auditory pathway. In: Moore DR, Fuchs PA, Rees A, Palmer AR, Plack CJ, editors. The Oxford Handbook of Auditory Science: The auditory brain. New York: Oxford University Press; 2010. p. 361-86.

[41] Emami SF. Word Recognition Score in White Noise Test in Healthy listeners. Sch J App Med Sci. 2015;3(1A):29-33.

[42] Emami SF. Is all human hearing cochlear? ScientificWorldJournal. 2013;2013:147160. [DOI:10.1155/2013/147160]

[43] Emami SF, Pourbakht A, Daneshi A, Sheykholeslami K, Emamjome H, Kamali M. Sound sensitivity of the saccule for low frequencies in healthy adults. ISRN Otolaryngol. 2013;2013:429680. [DOI:10.1155/2013/429680]

[44] Emami SF, Pourbakht A, Sheykholeslami K, Kamali M, Behnoud F, Daneshi A. Vestibular hearing and speech processing. ISRN Otolaryngol. 2012;2012:850629. [DOI:10.5402/2012/850629]

[45] Emami SF, Daneshi A. Vestibular hearing and neural synchronization. ISRN Otolaryngol. 2012;2012:246065. [DOI:10.5402/2012/246065]

[46] Assmann PF, Summerfield Q. The contribution of waveform interactions to the perception of concurrent vowels. J Acoust Soc Am. 1994;95(1):471-84. [DOI:10.1121/1.408342]

[47] Hornickel J, Anderson S, Skoe E, Yi HG, Kraus N. Subcortical representation of speech fine structure relates to reading ability. Neuroreport. 2012;23(1):6-9. [DOI:10.1097/WNR.0b013e32834d2ffd]

[48] Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nat Rev Neurosci. 2010;11(8):599-605. [DOI:10.1038/nrn2882]

[49] Anderson S, Parbery-Clark A, Yi HG, Kraus N. A neural basis of speech-in-noise perception in older adults. Ear Hear. 2011;32(6):750-7. [DOI:10.1097/AUD.0b013e31822229d3]

[50] Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci. 2002;6(1):37-46. [DOI:10.1016/S1364-6613(00)01816-7]

[51] Martin RC. Language processing: functional organization and neuroanatomical basis. Annu Rev Psychol. 2003;54:55-89. [DOI:10.1146/annurev.psych.54.101601.145201]

[52] Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology. 2010;47(2):236-46. [DOI:10.1111/j.1469-8986.2009.00928.x]

[53] Tremblay P, Brisson V, Deschamps I. Brain aging and speech perception: Effects of background noise and talker variability. Neuroimage. 2021;227:117675. [DOI:10.1016/j.neuroimage.2020.117675]

[54] Shestakova A, Brattico E, Soloviev A, Klucharev V, Huotilainen M. Orderly cortical representation of vowel categories presented by multiple exemplars. Brain Res Cogn Brain Res. 2004;21(3):342-50. [DOI:10.1016/j.cogbrainres.2004.06.011]

[55] Obleser J, Elbert T, Lahiri A, Eulitz C. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. Brain Res Cogn Brain Res. 2003;15(3):207-13. [DOI:10.1016/S0926-6410(02)00193-3]

[56] Restle J, Murakami T, Ziemann U. Facilitation of speech repetition accuracy by theta burst stimulation of the left posterior inferior frontal gyrus. Neuropsychologia. 2012;50(8):2026-31. [DOI:10.1016/j.neuropsychologia.2012.05.001]

[57] Nuttall HE, Kennedy-Higgins D, Devlin JT, Adank P. Modulation of intra- and inter-hemispheric connectivity between primary and premotor cortex during speech perception. Brain Lang. 2018;187:74-82. [DOI:10.1016/j.bandl.2017.12.002]

[58] Heyes C, Catmur C. What Happened to Mirror Neurons? Perspect Psychol Sci. 2022;17(1):153-68. [DOI:10.1177/1745691621990638]

[59] Wong PC, Parsons LM, Martinez M, Diehl RL. The role of the insular cortex in pitch pattern perception: the effect of linguistic contexts. J Neurosci. 2004;24(41):9153-60. [DOI:10.1523/JNEUROSCI.2225-04.2004]

[60] Callan DE, Jones JA, Callan A. Multisensory and modality specific processing of visual speech in different regions of the premotor cortex. Front Psychol. 2014;5:389. [DOI:10.3389/fpsyg.2014.00389]

[61] Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature. 2000;403(6767):309-12. [DOI:10.1038/35002078]

[62] Rogers JC, Möttönen R, Boyles R, Watkins KE. Discrimination of speech and non-speech sounds following theta-burst stimulation of the motor cortex. Front Psychol. 2014;5:754. [DOI:10.3389/fpsyg.2014.00754]

[63] Hickok G, Okada K, Serences JT. Area Spt in the human planum temporale supports sensory-motor integration for speech processing. J Neurophysiol. 2009;101(5):2725-32. [DOI:10.1152/jn.91099.2008]