Research Article

Perception of Temporal Fine Structure in Individuals with Normal Hearing Sensitivity: A Comparison of Different Measures

JAIISH(2019)
Vol 38 pp. 47-57

Pavan Mahadevaiah[a], Merin Mathews[b], Ramiz Malik M.[c],
Indira C. Pazhayapisharath[b], Arivudai Nambi[d] & Rajalakshmi Krishna[b]

Affiliations

[a]Resonance Speech and Hearing Clinic, Mysuru, Karnataka, India
[b]All India Institute of Speech and Hearing, Department of Audiology, University of Mysore, Mysuru, Karnataka, India
[c]Department of Audiology, Mar Thoma College of Special Education, Kasargod, Kerala, India
[d]Department of Audiology and Speech Language Pathology, Kasturba Medical College, Mangalore, Karnataka, India

Corresponding Author

Pavan Mahadevaiah
Resonance Speech and Hearing Clinic
Mysuru
Karnataka
India.
pavandev1987@gmail.com

Key Words

TFS speech
Temporal fine structure
Recovered envelope speech

Abstract

Speech is a complex signal fundamentally decomposed by the auditory filters into corresponding narrowband signals, each of which can be considered as a slowly varying temporal envelope (ENV) superimposed upon a rapidly oscillating temporal fine structure (TFS). Both ENV and TFS are coded in the auditory nervous system in terms of time related changes in the neural firing. Encoding these cues is considered vital for speech perception, especially in the presence of background noise. The study explored the relationship between different measures of sensitivity to TFS (TFS-speech, Recovered envelope speech (RENV), and sensitivity to TFS using complex tones) and also explored how these different measures of TFS were related to performance on speech perception in noise (SPIN) testing using sentence stimuli on twenty young adults with normal hearing. The sentences were degraded using five schemes and TFS perception of complex tones were assessed using two schemes. The findings of the study showed no significant correlation of the different measures considered in the study, namely, the perception of TFS-speech, the perception of RENV-speech, and the perception TFS in complex tones with SPIN scores in the participants of the study. The results of the study show that, in a normal hearing young adult, speech perception in the presence of continuous noise is not related to their sensitivity to different measures of TFS perception.

INTRODUCTION

The perception of speech involves the interpretation of complex acoustical patterns and perceiving them as linguistic units (McRoberts, 2008). It is a complex task because a single acoustic pattern may not always represent the same speech segments. Instead, the patterns may vary depending upon the preceding or following segment as well as the auditory environment. Speech perception becomes even more difficult in the presence of background noise (Moore, 2003)

The interpretation of speech depends on how well the auditory system decodes the temporal cues present in it (Moon & Hong, 2014; Shamma & Lorenzi, 2013). The encoding of temporal cues such as the temporal fine structure (TFS) and temporal envelope (ENV), is considered crucial for speech understanding in the presence of background noise (Lorenzi et al., 2009). The TFS, sometimes called the carrier, is characterised by rapid oscillations in the signal with a rate close to the centre frequency of the frequency band of the signal. The ENV, on the other hand, corresponds to the slow, amplitude modulation of the carrier (or the TFS) over time (Lorenzi et al., 2009; Moon & Hong, 2014). Both TFS and ENV are coded in the auditory nervous system in terms of time related changes in neural firing. More specifically, TFS is coded as the neural phase locking to the phase of the carrier signal (the synchronisation of nerve firing with a particular phase of the stimulus), and ENV is encoded as fluctuations in the short-term rate of firing of neurons (Buss et al., 2004; Moon & Hong, 2014).

It is well documented that the ENV cues are sufficient to have good speech perception in quiet, but are inadequate in the presence of background noise (Fu et al., 1998; Moore, 2019; Shannon et al., 1995; Wilson et al., 1991). This may be because the ENV cues alone fail to provide perceptual segregation in a complex listening environment (Moore, 2008), due to smearing of the envelope fluctuations by the noise – the noise fills the dips across the waveform, rendering the slow modulations incomprehensible (Shetty, 2016). TFS information is reported to be useful when listening to speech in the presence of fluctuating noise (Hopkins & Moore, 2009) as well as steady-state noise (Moore, Glasberg, Flanagan, & Adams, 2006). Individuals with normal hearing benefit from 'dip listening' in noise, especially the fluctuating type (gathering snippets of the signal when it is audible over the noise). TFS cues could be relatively robust to the effects of noise because the auditory system can detect the presence of speech signals in the dips (or vallies) from neural phase locking to the TFS (Lorenzi et al., 2006a). Different signals have specific TFS cues and this will help in speech perception from the snippets of information obtained in the presence of noise (Yellamsetty, 2016). Also, pitch is reported to be an important cue to differentiate two signals (speech and noise in this case), and TFS contributes to the perception of pitch significantly (Moore, 2019). At the level of the auditory nerve, acoustic TFS is coded not just as "true TFS", but it also contributes to the recovery of envelope information (Swaminathan & Heinz, 2012). It is hypothesized that, at the neural level, TFS cues can encode the acoustic spectrogram (Shamma & Lorenzi, 2013).

The ability to effectively use TFS information in the presence of noise is considerably impaired in individuals with cochlear hearing loss, and it has been attributed to changes in their cochlear mechanics (Henry & Heinz, 2013; Hopkins & Moore, 2007; Lorenzi et al., 2006b). Assessment of speech perception abilities in noise, therefore, indicates one's abilities to utilise TFS information.

However, clear information about the different aspects of TFS perception is needed to understand the actual contribution of TFS cues to the perception of speech. In view of this, studies that explore sensitivity to TFS information and the relative contribution of TFS and ENV components to speech perception are essential.

Sensitivity to TFS information (of complex tones) may be understood using tests like the TFS1 and TFS-LF (Hopkins & Moore, 2010a; Sek & Moore, 2012). These tests adaptively vary the TFS information, while leaving the envelope unaltered. The TFS1 test varies the frequency of the TFS component delivered to a single ear, and the TFS-LF test varies the phase of the TFS between the right and the left ears. The ability to detect the smallest change in TFS is used to understand a participant's sensitivity to TFS. For example, Hopkins and Moore (2011) showed that the TFS sensitivity is weak in the elderly with normal hearing sensitivity (63-66 years) compared to young participants with normal hearing sensitivity (20-35 years), even when their frequency discrimination abilities were comparable. In a similar study, using the same tests, Moore, Vickers and Mehta (2012) showed that age and sensitivity to TFS cues were correlated.

Various methods are used to study the relative roles of TFS and ENV cues in speech perception. One such method is 'Vocoding'. It is the extraction of TFS from a speech signal to preserve ENV cues alone or vice versa. In this method, a signal is split into different frequency bands and envelope and TFS are extracted from each band using processes like the Hilbert transform. If the envelope information from the Hilbert analysed signal is to be retained, the extracted envelope is low pass filtered, and a sine wave with a frequency equal to the centre frequency of each band is amplitude modulated with it. The output from all bands is then combined, and the final product is a signal with only envelope information (Swaminathan et al., 2014).

In order to make a TFS only signal, the envelope component is discarded following the Hilbert transformation. The TFS in each band is multiplied by a constant equal to the root-mean-square (RMS) power of the bandpass filtered signal. The 'power-weighted' TFS signals are then summed over all the frequency bands. These stimuli contain TFS information only and are termed as 'TFS-speech' (Lorenzi et al., 2006b). The cosine of instantaneous phase of the hilbert analytic signal from the output of a filterbank is used to create the TFS-speech. When the filterbank uses lesser number of filters with broad bandwidths, TFS perception is poor. When the filterbank uses more number of (narrower) bandwidth filters, TFS perception is better (Smith et al., 2002). Studies have shown that the ENV cues are reconstructed at the output of the auditory filters and that the perception of this 'recovered envelope' (Figure 1) aids in speech comprehension from TFS-speech (Ghitza, 2001).

To understand this phenomenon further, speech stimuli were made by recovering the envelope from the TFS-speech (Gilbert & Lorenzi, 2006; Léger et al., 2015). The recovered envelope is used to amplitude modulate a sine wave with a frequency equal to the centre frequency of its extracted frequency band, resulting in recovered envelope speech (RENV speech). Studies have found that these stimuli are intelligible (Lorenzi et al., 2006b; Sheft et al., 2008b), but their intelligibility depended on the number of filters used by the filterbank for recovery (Lorenzi et al., 2006b; Sheft et al., 2008b). Nevertheless, findings of these studies indicated that the auditory system could indeed re-create intelligible envelopes from TFS-speech. The recovery of envelope while trying to study the perception of TFS-speech, contaminates the findings with regard to TFS perception. In this context, TFS-speech may actually be considered to be a mere vehicle that carries the envelope information (that is promptly recovered at the level of the auditory system).

In contrast, studies that tried to eliminate the influence of RENV on speech perception have shown that TFS information does contribute to speech perception on its own; it does not function as a mere vehicle to carry ENV cues to the cochlea (Hopkins et al., 2010; Hopkins & Moore, 2010b; Sheft et al., 2008b). For example, Sheft, Ardoint and Lorenzi,

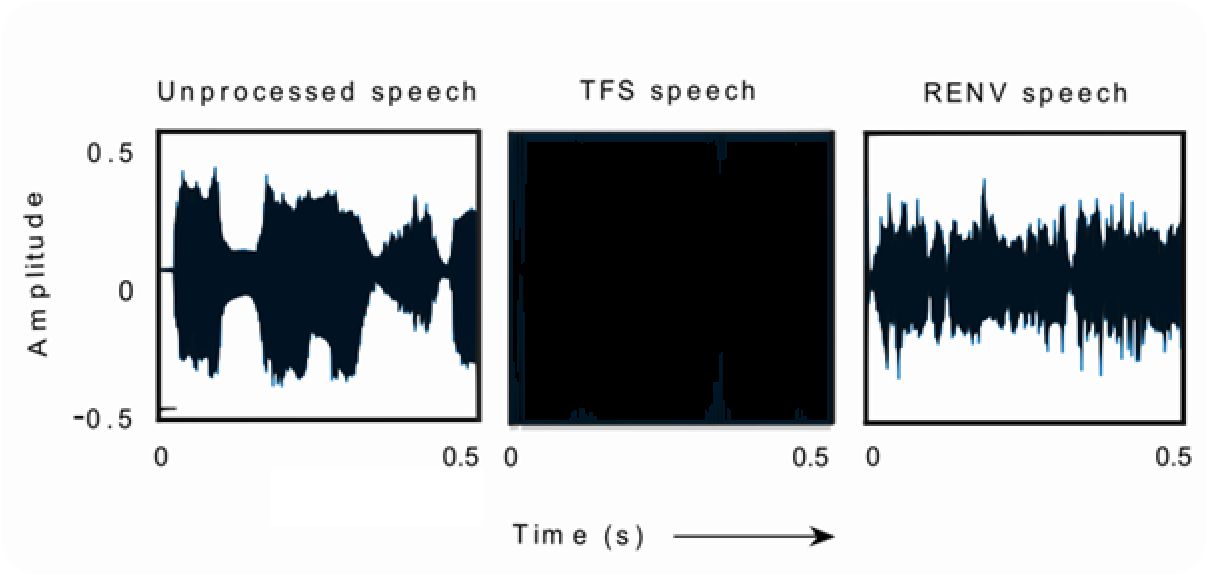

Figure 1: Example waveforms of the unprocessed speech, TFS-speech and RENV speech.

(2008) assessed the contribution of TFS information to consonant identification using different speech processing methods. They observed the consonant identification patterns under different conditions (using a Phase modulation function (PM condition), a Frequency modulation function (FM condition), and one Envelope only condition) where the ability for ENV reconstruction was restricted. They reported that TFS cues contributed more to the perception of place cues, compared to manner cues.

The literature cited above shows that a number of measures exist to study the sensitivity to TFS cues and its relative contribution to speech perception. However, it is not known as to how these different measures compare, or if the different measures assess the same phenomenon, as there are no studies that systematically compare the different methods. It is neither known if the relationship between TFS perception and SPIN predicted in the earlier studies was due to a methodological artifact (noise effects or perception of recovered envelopes from TFS), nor how the TFS information derived from different methods predicts speech perception in noise. Therefore, the aims of this study were 1) to explore the relationship between different measures of sensitivity to TFS (TFS-speech and RENV speech using sentences, TFS1 and TFS-LF tests), and 2) to explore how the results from the different measures of TFS are related to performance on SPIN testing using sentence stimuli.

METHODS

Participants

Twenty individuals in the age range of 18 to 25 years (mean age: 20.4 years) participated in the study. Their audiometric evaluation showed thresholds better than 15 dB HL between 250 and 8000 Hz. They had no history of hearing or comprehension difficulties as reported. All the participants were native speakers of Kannada (a language spoken in the south Indian state of Karnataka) with proficiency in comprehending speech and script in the language. An informed consent form was obtained from all the participants, the method abided by the ethical guidelines for bio behavioural research involving human subjects (Venkatesan, 2009), and was approved by the ethical committee for research at the institute.

Sentence comprehension tests

Stimuli

Kannada sentence lists developed by Geetha, Kumar, Manjula, and Pavan (2014) were used to prepare stimuli for the study. There were 24 lists developed, and each list had 10 sentences, each with four keywords, resulting in 40 keywords per list. These sentences were processed in 3 different ways- initially, removing the envelope and retaining only the TFS from the sentence; secondly, by reconstructing the envelope from the extracted TFS; and thirdly, the sentences were mixed with noise to create stimuli for SPIN testing. The procedure used for processing of the stimulus in each of the two methods was similar to Swaminathan et al. (2014) and Gilbert and Lorenzi (2006).

TFS-speech: The sentences were first bandpass filtered into 2, 4, 8 and 16 bands (2nb, 4nb, 8nb and 16nb) of equal bandwidth on a log frequency scale spanning 80 to 8020 Hz (Gilbert & Lorenzi, 2006). The output from each band underwent Hilbert analysis and the TFS component within each band was extracted as the cosine of the phase of the Hilbert analytic signal. The TFS component was scaled to match the long-term average energy of the original signal in each bandpass. The resulting amplitude normalised TFS components were added to get the TFS-speech stimulus. During pilot testing, 2nb and 4nb TFS stimuli resulted in ceiling effects. Therefore, only 8nb and 16nb stimuli were used for TFS-speech testing.

RENV speech: From the TFS stimuli created with 2nb, 4nb, 8nb and 16nb conditions, the envelopes were extracted, and RENV speech was created. However, RENV speech perception was not assessed for the 8nb and 16nb stimuli, since the intelligibility of RENV speech reduced drastically after the 4nb condition, during a pilot testing. Each of the TFS-extracted sentences was first bandpass filtered into 40 frequency bands using twelfth order digital Butterworth filter. The frequency bands were of equal bandwidth on a log frequency scale between 80 to 8020 Hz, simulating a cochlear filter bank. The signal was filtered in forward and backward directions. The envelope component within each band was extracted as the magnitude of the Hilbert analytic signal and low-pass filtered at 300 Hz using a sixth order Butterworth filter. The envelope recovered from each frequency band was used to vocode a pure tone carrier with a central frequency of the corresponding frequency band and was band pass filtered. The resultant components were added to get RENV speech.

Stimuli generation for SPIN task: As mentioned above, the same sentence lists (unprocessed) were used to create stimuli for this test condition. Speech shaped noise equivalent to the spectrum of each selected list was produced and mixed at 5 dB signal to noise ratio (SNR) using custom code with MATLAB 2014 (Mathworks Inc., Natick, MA, USA).

Procedure

The stimuli for speech identification tests were created for the four different speech processing conditions- (TFSnb8, TFSnb16, RENVnb2 and RENVnb4, and for SPIN task). The participants were seated comfortably in a chair, and the testing was carried out in a sound treated room. Sentence comprehension tests, as well as the TFS tests, were carried out in random orders, to rule out order effect. All the tests were carried out using a Lenovo Laptop (running on Windows 10 OS, Intel(R) i3-2370M CPU) and the stimuli were presented through Sennheiser HDA200 headphones. Calibration was performed to set the output of the headphone to 70 dB SPL using a Bruel & Kjaer (2250) sound level meter and ear simulator, complying with IEC 60318-1. This presentation level was chosen for the speech perception tests as a number of SPIN tests use presentation levels varying between 65 to 80 dB SPL (Bench et al., 1979; Cameron & Dillon, 2007; Killion et al., 2004; Nilsson et al., 1994). Also, it has been reported in the literature that perception of envelope speech and TFS-speech (for speech identification) do not vary with presentation level (Sheft et al., 2008a).

The stimuli were presented to the participants through the software Paradigm (version 2.5.0.68, Perception Research Systems Inc.). The sentence list for each condition was selected randomly from the 24 lists, and no list was presented to a participant more than once. Under each condition, one list with 40 keywords was presented (equivalency was established for all the sentence lists in the corpus). The stimuli were presented at 70 dB SPL and the participants had to repeat the sentences heard verbatim. The responses were voice recorded for the scoring of keywords. Each correctly identified keyword was given a score of 1, and wrongly identified keyword was assigned a score of 0. Therefore, the maximum achievable score in each stimulus condition was 40.

TFS perception of complex tones

The participants' sensitivity to changes in the TFS of complex tones was assessed using two tests developed by Moore and colleagues. One test assessed sensitivity to TFS within one ear for high frequency components (TFS1) (Moore & Sek, 2009) whereas the other test assessed sensitivity to TFS across two ears for low frequency stimuli (TFS-LF) (Hopkins & Moore, 2010a).

Procedure

The software encompassing the tests (Two methods for determining TFS sensitivity, downloaded from https://www.psychol.cam.ac.uk/hearing) was installed into an HP Laptop (running on Windows 10 OS, Intel(R) i5-6200U CPU). The stimuli were presented through Sennheiser HDA200 headphones with stereo jacks. The calibration procedure adhered to the instructions by the software developers (Hopkins & Moore, 2010a), thereby ensuring appropriate levels of presentation of the test stimuli. The inbuilt sound card of the laptop was used to drive the headphones. Nevertheless, the stereo presentation of the stimuli were ensured by calibration specific to these tests.

The procedure followed for the tests TFS1 and TFS-LF were based on Moore and Sek (2009) and Hopkins and Moore (2010) respectively. Both tests used a two-interval two-alternative forced-choice method. The two intervals were separated by 500 ms, and each interval contained four tones of 400 ms duration. The tones were consecutively presented with 100 ms gap between them. All the four tones in one of the intervals had identical TFS (the standard), whereas, the other interval (the target) had TFS different from the standard for the second and fourth tones. The participants' task was to differentiate the target interval (detailed under specific sections below) from the standard. Feedback was given after each trial. The starting and variable parameters of the test stimuli were set, and the software used a 2-down, 1-up adaptive procedure to arrive at the 71% correct point on the psychometric function (Levitt,

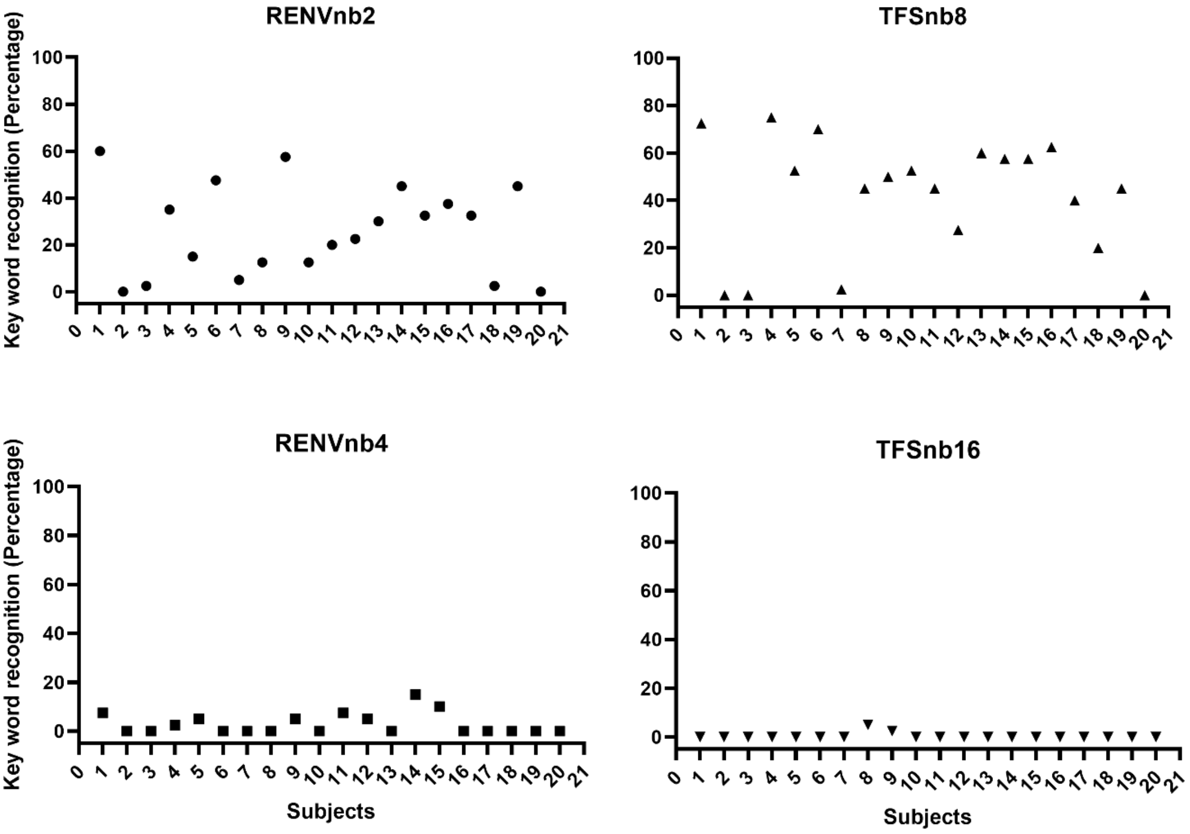

Figure 2: Individual speech perception scores from participants on RENVnb2, RENVnb4, TFSnb8 and TFSnb16 conditions.

Table 1: Mean, standard error of mean, median, and SD for scores obtained in RENV speech, TFS-speech conditions, TFS tests (TFS-LF and TFS1 - TFS1R for the right ear and TFS1L for the left ear) and SPIN.

	RENVnb2	**RENVnb4**	**TFSnb8**	**TFSnb16**	**TFS1R**	**TFS1L**	**TFS-LF**	**SPIN score at -5 dB SNR**
Mean	10.3	1.15	16.7	0.150	16.3	17.4	24.6	27.4
Std. error mean	1.73	0.386	2.24	0.109	2.05	1.50	2.88	1.27
Median	10.5	0.00	19.0	0.00	13.5	15.0	23.5	26.5
SD	7.72	1.73	10.0	0.489	9.15	6.71	12.9	5.67

Table 2: Mean, median, minimum and maximum values of normalized scores obtained in RENV speech , TFS-speech conditions, TFS tests (TFS-LF and TFS1 - TFS1R for the right ear and TFS1L for the left ear), and SPIN.

	REN-Vnb2	**RENVnb4**	**TFSnb8**	**TFSnb16**	**TFS1R**	**TFS1L**	**TFS-LF**	**SPIN score at -5 dB SNR**
Mean	-7.63e−17	5.97e-17	6.66e-17	3.33e-17	-5.76e−17	-2.03e−16	-1.03e−16	1.37e-16
Median	0.0259	-0.667	0.230	-0.307	-0.306	-0.350	-0.0855	-0.168
Minimum	-1.33	-0.667	-1.67	-0.307	-1.13	-1.25	-1.29	-1.84
Maximum	1.77	2.81	1.33	3.78	2.26	2.18	1.51	1.86

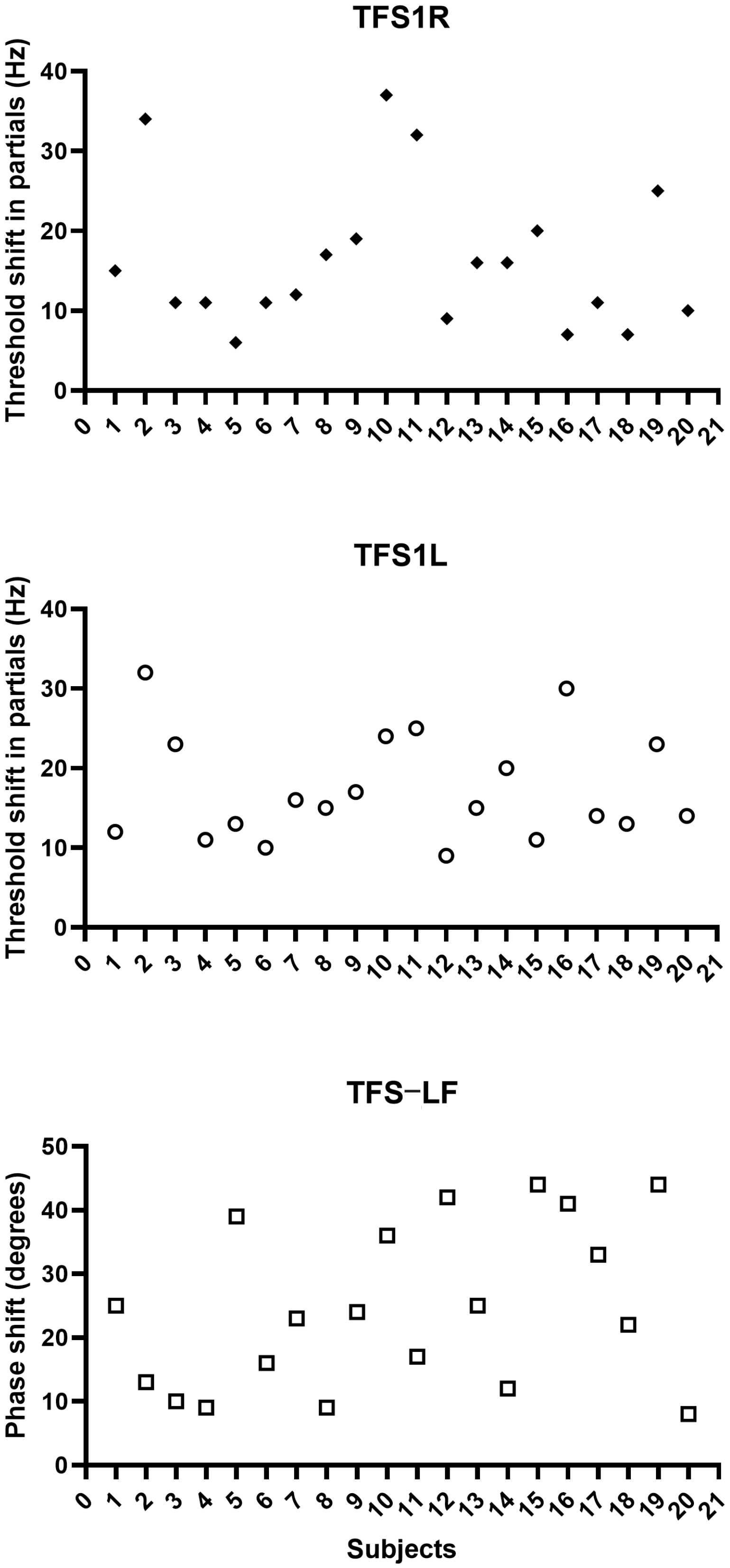

Figure 3: Individual data fromparticipants on TFS1 test (TFS1R for the Right ear and TFS1L for the Left ear) and TFS-LF test.

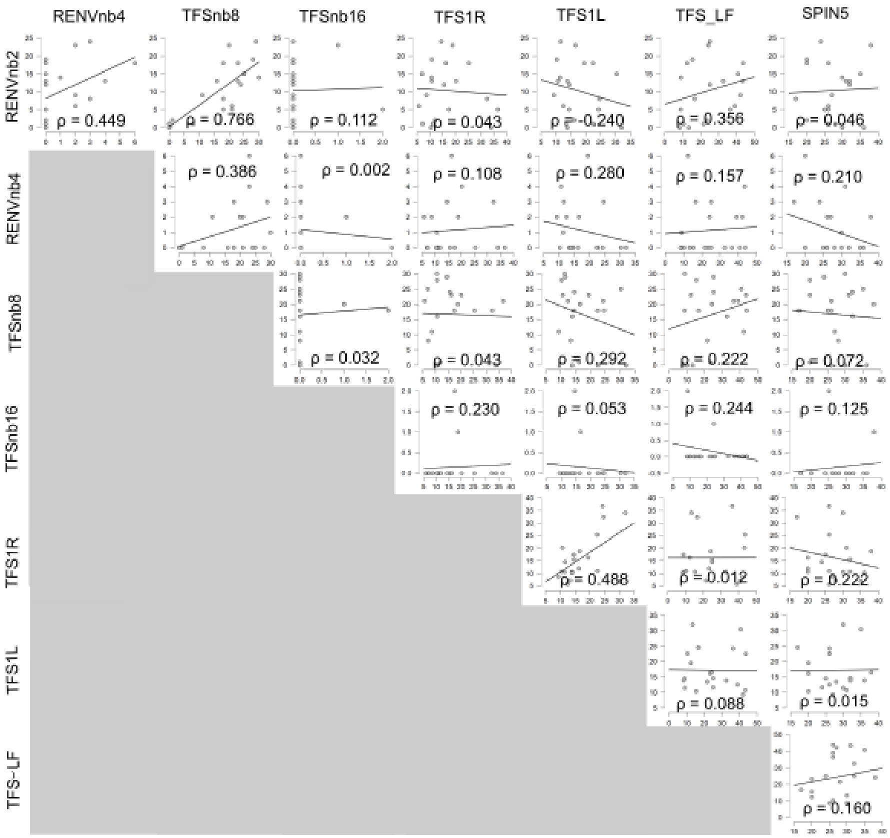

Figure 4: Correlation plots of all the parameters compared in the study.

1971). Eight reversals were carried out with varying TFS parameters, and the values from the last six reversals were used to calculate the threshold. If the standard deviation (SD) of the last six reversals was more than 0.2, new testing was carried out. If during the adaptive procedure, the value of the variable parameter exceeded the maximum more than twice, the method of constant stimuli was used (40 trials) with the value of the parameter fixed at maximum. The thresholds were estimated once the participants were familiarised with the stimulus and task.

TFS-LF test: This test assessed the binaural sensitivity to TFS in complex tones (Hopkins & Moore, 2010a). The participants were instructed to listen carefully, and that they will hear 2 sets of 4 tones each. They were asked to point out the set with tones that seemed to shift between the ears (the target interval with the phase shift) compared to the set with tones that was heard in both ears together, or in the middle of the head (the standard interval, with identical phases between the ears). The phase shift between the ears (delta φ) was the manipulated variable, and the initial value was set at 180^0. This was carried out using 500 Hz stimulus considering good sensitivity to TFS at this frequency, compared to higher frequencies using this test. The stimuli were presented simultaneously to both ears at 50 dB SPL.

RESULTS

The participants' individual scores on each test are presented as scatterplots in figures 2 and 3 (in pages 115 and 116 respectively). The mean, median and SD of raw scores of speech comprehension tasks (RENV, TFS conditions, and SPIN) and the thresholds from tests assessing TFS perception of complex tones (TFS tests) are shown in Table 1 (in page no. 115). The Shapiro-Wilk test was done on raw scores to check the normality of distribution in each condition. Non-parametric statistics were used for further analysis, as the data was not normally distributed across conditions. Since the data were in different scales, all further analyses were done on normalized data (Table 2, in page no. 115). The correlation between scores on conditions of RENV, TFS speech perception, thresholds from TFS tests and SPIN scores were checked using Spearman's correlation coefficient (Figure 4, in page no. 117). The scores were significantly correlated between the following conditions measuring TFS sensitivity: RENVnb2 with RENVnb4 ($r = 0.449$, $p = 0.047$) and TFSnb8 ($r = .766$, $p = 0.000$); TFS1 right ear with TFS1 left ear scores ($r = 0.488$, $p = 0.029$). There was no significant correlation ($p < 0.05$) between any of the RENV speech, TFS-speech conditions, scores on TFS tests with SPIN scores at -5 dB SNR.

DISCUSSION

The goal of the present study was to examine the correlation of measures of TFS perception and sensitivity to TFS cues to the participants' ability to perceive speech in the presence of noise. The findings of the study showed no significant correlation of the different measures considered in the study, namely, the perception of TFS-speech, the perception of RENV-speech, and the perception TFS in complex tones, with SPIN scores in the participants of the study. The results of the study are in agreement with some of the literature (Neher et al., 2011; Strelcyk & Dau, 2009) and in disagreement with the others (Füllgrabe et al., 2015; Peters et al., 1998). This discrepancy may be majorly attributed to the differences in the methods used in the studies. These differences may be related to parameters such as the background noise used for SPIN, the population tested, the stimulus used for tests, etc.

For example, Peters et al. (1998) measured Speech recognition thresholds (SRT) in steady and fluctuating background noise for individuals with normal hearing and young and older individuals with hearing loss. They compared the SRTs with TFS1 and TFS-LF tests and found a good correlation between SRTs in the modulated noise and scores on TFS1 test in older individuals with normal hearing and younger and older individuals with hearing loss. However, the correlation between the measures is not considered for young participants with normal hearing. The test measure used here was SRT and not the speech identification score (SIS), and the noise used differed from the present study.

But studies have been conducted where sentences were used to measure SIS (as a measure of speech perception), and they have used different measures of TFS sensitivity (like the TFS1 and TFS-LF tests) to understand the relationship between TFS sensitivity and speech perception. Fullgrabe et al. (2015) observed a good correlation between the TFS1 and TFS-LF test scores and SPIN scores of their young normal hearing participants. However, they had used modulated noise (whereas the present study used non-fluctuating noise) for SPIN testing of sentences, and the testing was done in a sound field condition. On the other hand, with a similar testing paradigm, Neher et al. (2011) found no correlation between speech perception and scores on TFS-LF test in a group of 8 normal hearing participants during sound field testing. They had used sentences and spatially separated fluctuating background noise for stimuli. Strelcyk and Dau (2009) found no correlation between the measures in the presence of modulated noise. But reported a significant correlation between these measures in the presence of two-talker babble. Therefore, it is possible that the findings of the studies vary depending upon the stimulus used, speech perception measures considered and the noise used. The present study used a non-fluctuating noise for assessing SPIN since speech identification is possible in places where the SNR is good while listening to a fluctuating noise. This means that, in such occasions, speech is not effectively masked (Cooke, 2006). Therefore, a speech spectrum noise was used and the scores obtained using the same can be considered as that obtained from a true masker. Scores obtained on the SPIN test could also be affected by the SNR used (Shojaei et al., 2016). However, the stimuli were presented at -5 dB SNR as it is the recommended level for testing by the developers (Geetha et al., 2014).

The study also found correlations between some, and not all the measures of TFS sensitivity used. The different tests used in the study measured sensitivity to TFS information, but possibly the different aspects of sensitivity to TFS. Perception of TFS information in TFSnb8 and TFSnb16 conditions involved perception of the extracted TFS (or the resultant recovery of the envelope at the level of the listener's cochlea) from the sentence stimuli. The RENVnb2 and RENVnb4 conditions tested the listener's ability to perceive the simulation of extracted envelope from the TFS-speech. Significant correlation found between the two RENV conditions was possibly because the two tests measured the same construct underlying TFS perception. Good correlation was also seen between scores from RENVnb2 and TFSnb8 conditions. Even though RENV speech stimuli were derived from TFS-speech, the number of bands used for extraction of TFS from the original stimuli were different. It is possible that the recovery of the envelope from the TFS-speech at the cochlear level is correlated with simulations of recovery of the envelope. At the extreme conditions, however, the perception of speech deteriorated in the participants. It was also seen that the deterioration varied among the participants and this variability could be the reason that correlations were not observed between these conditions.

The TFS perception of complex tones were not correlated with the measures of speech perception, but TFS1 test scores correlated between the right and left ears. All the participants of the study were young individuals with normal hearing which could possibly be the reason for their comparable ability in the two ears to perceive changes in high frequency TFS information. However, these measures did not correlate with the results of TFS-LF test, indicating that the ability to perceive low frequency and high frequency TFS information were not comparable. Other studies that have reported correlation between TFS tests and speech perception in noise have compared these measures in different age groups (Füllgrabe, Moore, & Stone, 2015; Peters & Moore, 1992). A comparison of the same measures as administered in the present study with differences in procedures used in the studies (like the noise and the speech test used) could have contributed to the differences in the findings. As stated, several studies have found correlations between measures of TFS sensitivity and speech perception in the elderly (Füllgrabe et al., 2015; Hopkins

& Moore, 2010a; Peters et al., 1998) or in individuals with hearing loss (Füllgrabe et al., 2015; Hopkins & Moore, 2010a; Peters et al., 1998). The present study focussed on young individuals with normal hearing. Further studying the same measures in other populations susceptible to poorer processing of TFS cues may reveal more information regarding the relationship between different measures of TFS sensitivity and SPIN. Hearing loss also influences the perception of TFS information, as well as perception of speech in the presence of noise (Strelcyk & Dau, 2009). It is well known that pitch is important for localization, lateralization, and music perception, TFS cues contribute to pitch perception, and also perception of speech in the presence of noise. Individuals with hearing loss experience difficulty in all these domains. Studies that measure TFS sensitivity in indivudals with hearing loss employ different measures for the purpose. However, in order to make any meaningful interpretation and comparison of the findings from different studies in this domain, one needs to know how these measures are related. Therefore, better understanding of relationship between measures of TFS sensitivity will add value to the studies exploring those factors in individuals with hearing loss.

CONCLUSIONS

In the present study, an attempt was made to see if a young person's ability to understand speech in a commonly encountered adverse listening environment is related to a measure of sensitivity to TFS cues. The results of the study and the ensuing discussion show that in a normal hearing young adult speech perception in the presence of continuous noise is not related to their sensitivity to different measures of TFS perception. However, the same may not be true for speech perception in the presence of fluctuating noise, or when the tests are administered on a different population.

ACKNOWLEDGEMENTS

The authors would like to thank the Director, All India Institute of Speech and Hearing, Mysore for permission to carry out the study. The authors would also like to thank all the participants in the study for their support.

REFERENCES

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology, 13*(3), 108—112. https://doi.org/10.3109/03005367909078884

Buss, E., Hall, J. W., & Grose, J. H. (2004). Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss. *Ear and Hearing, 25*, 242–250. https://doi.org/10.1097/01.AUD.0000130796.73809.09

Cameron, S., & Dillon, H. (2007). Development of the Listening in Spatialized Noise-Sentences Test (LISN-S). *Ear and Hearing, 28*(2), 196–211. https://doi.org/10.1097/AUD.0b013e318031267f

Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America, 119*, 1562–1573.

Fu, Q. J., Shannon, R. V, & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. *The Journal of the Acoustical Society of America, 104*(6), 3586–3596. http://www.ncbi.nlm.nih.gov/pubmed/9857517

Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience, 7*(JAN), 1–25. https://doi.org/10.3389/fnagi.2014.00347

Geetha, C., Kumar, K. S. S., Manjula, P., & Pavan, M. (2014). Development and standardisation of the sentence identification test in the Kannada language. J*ournal of Hearing Science, 4*(1), 18-26.

Ghitza, O. (2001). On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception. *The Journal of the Acoustical Society of America, 110*, 1628–1640. https://doi.org/10.1121/1.1396325

Gilbert, G., & Lorenzi, C. (2006). The ability of listeners to use recovered envelope cues from speech fine structure. *The Journal of the Acoustical Society of America, 119*, 2438–2444. https://doi.org/10.1121/1.2173522

Henry, K. S., & Heinz, M. G. (2013). Effects of sensorineural hearing loss on temporal coding of narrowband andbroadband signals in the auditory periphery. In Hearing Research (Vol. 303, pp. 39–47). https://doi.org/10.1016/j.heares.2013.01.014

Hopkins, K., & Moore, B. C. J. (2007). Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. *The Journal of the Acoustical Society of America, 122*, 1055–1068. https://doi.org/10.1121/1.2749457

Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical Society of America, 125*(1), 442–446. https://doi.org/10.1121/1.3037233

Hopkins, K., & Moore, B. C. J. (2010a). Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies. *International Journal of Audiology, 49*(12), 940–946. https://doi.org/10.3109/14992027.2010.512613

Hopkins, K., & Moore, B. C. J. (2010b). The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America, 127*, 1595–1608. https://doi.org/10.1121/1.3293003

Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. T*he Journal of the Acoustical Society of America, 130*, 334–349. https://doi.org/10.1121/1.3585848

Hopkins, K., Moore, B. C. J., & Stone, M. A. (2010). The effects of the addition of low-level , low-noise noise on the intelligibility of sentences processed to remove temporal envelope information. *The Journal of the Acoustical Society of America, 128*(4), 2150-61. https://doi.org/10.1121/1.3478773

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 116*(4), 2395–2405.

Léger, A. C., Desloge, J. G., Braida, L. D., & Swaminathan, J. (2015). The role of recovered envelope cues in the identification of temporal-fine-structure speech for hearing-impaired listeners. *The Journal of the Acoustical Society of America, 137*(1), 505. https://doi.org/10.1121/1.4904540

Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., & Moore, B. C. J. (2009). Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *The Journal of the Acoustical Society of America, 125*(1), 27–30. https://doi.org/10.1121/1.2939125

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006a). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 18866–18869. https://doi.org/10.1073/pnas.0607364103

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006b). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America, 103*(49), 18866–18869. https://doi.org/10.1073/pnas.0607364103

McRoberts, G. W. (2008). Speech PerceptionElsevier. In M. M. Haith & J. B. Benson (Eds.), E*ncyclopedia of infant and early childhood development*. (pp. 244–253). Elsevier Ltd. https://doi.org/https://doi.org/10.1016/B978-012370877-9.00154-7

Moon, I. J., & Hong, S. H. (2014). What is temporal fine structure and why is it important? *Korean Journal of Audiology, 18*(1), 1–7. https://doi.org/10.7874/kja.2014.18.1.1

Moore, B., Glasberg, B., Stoev, M., Fullgrabe, C., & Hopkins, K. (2012). The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies (L). *The Journal of the Acoustical Society of America, 131*, 1003–1006.

Moore, B.C.J. (2010). Testing for Cochlear Dead Regions: Audiometer Implementation of the TEN(HL) Test. *Hearing Review, 17*(1), 10-16,48.

Moore, B.C.J., & Ohgushi, K. (1993). Audibility of partials in inharmonic complex tones. *The Journal of the Acoustical Society of America, 93*, 452–461.

Moore, Brian C. J. (2003). An introduction to the psychology of hearing Academic Press, San Diego.

Moore, Brian C. J., Glasberg, B. R., Flanagan, H. J., & Adams, J. (2006). Frequency discrimination of complex tones; assessing the role of component resolvability and temporal fine structure. *The Journal of the Acoustical Society of America, 119*(1), 480–490. https://doi.org/10.1121/1.2139070

Moore, Brian C.J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology, 9*(4), 399–406. https://doi.org/10.1007/s10162-008-0143-x

Moore, Brian C.J. (2019). The roles of temporal envelope and fine structure information in auditory perception. *Acoustical Science and Technology, 40*(2), 61–83. https://doi.org/10.1250/ast.40.61

Moore, Brian C.J., & Sek, A. (2009). Development of a fast method for determining sensitivity to temporal fine structure. *International Journal of Audiology, 48*(4), 161–171. https://doi.org/10.1080/14992020802475235

Moore, Brian C.J., Vickers, D. A., & Mehta, A. (2012). The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. *International Journal of Audiology, 51*(10), 715–721. https://doi.org/10.3109/14992027.2012.690079

Neher, T., Laugesen, S., Søgaard Jensen, N., & Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *The Journal of the Acoustical Society of America, 130*(3), 1542–1558. https://doi.org/10.1121/1.3608122

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America, 95*(2), 1085–1099.

Peters, R. W., Moore, B. C. J., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *The Journal of the Acoustical Society of America, 103*(1), 577–587. https://doi.org/10.1121/1.421128

Plomp, R., & Mimpen, A. M. (1968). The ear as a frequency analyser II. *The Journal of the Acoustical Society of America, 43*, 764–467.

Sek, A., & Moore, B. C. J. (2012). Implementation of two tests for measuring sensitivity to temporal fine structure. *International Journal of Audiology, 51*(1), 58–63. https://doi.org/10.3109/14992027.2011.605808

Shamma, S., & Lorenzi, C. (2013). On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system. *The Journal of the Acoustical Society of America, 133*, 2818–2833. https://doi.org/10.1121/1.4795783

Shannon, R. V, Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303–304. https://doi.org/10.1126/science.270.5234.303

Sheft, S., Ardoint, M., & Lorenzi, C. (2008a). Speech identification based on temporal fine structure cues. *The Journal of the Acoustical Society of America, 124*, 562–575. https://doi.org/10.1121/1.2918540

Sheft, S., Ardoint, M., & Lorenzi, C. (2008b). Speech identification based on temporal fine structure cues. *The Journal of the Acoustical Society of America, 124*(1), 562–575. https://doi.org/10.1121/1.2918540

Shetty, H. N. (2016). Temporal cues and the effect of their enhancement on speech perception in older adults – A scoping review. *Journal of Otology, 11*(3), 95–101. https://doi.org/10.1016/j.joto.2016.08.001

Shojaei, E., Ashayeri, H., Jafari, Z., Zarrin Dast, M. R., & Kamali, K. (2016). Effect of signal to noise ratio on the speech perception ability of older adults. *Medical Journal of the Islamic Republic of Iran, 30*(1).

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature, 416*(6876), 87–90. https://doi.org/10.1038/416087a

Strelcyk, O., & Dau, T. (2009). Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *The Journal of the Acoustical Society of America, 125*(5), 3328. https://doi.org/10.1121/1.3097469

Swaminathan, J., & Heinz, M. G. (2012). Psychophysiological Analyses Demonstrate the Importance of Neural Envelope Coding for Speech Perception in Noise. *Journal of Neuroscience, 32*, 1747–1756. https://doi.org/10.1523/JNEUROSCI.4493-11.2012

Swaminathan, J., Reed, C. M., Desloge, J. G., Braida, L. D., & Delhorne, L. A. (2014). Consonant identification using temporal fine structure and recovered envelope cues. *The Journal of the Acoustical Society of America, 135*(4), 2078–2090. https://doi.org/10.1121/1.4865920

Venkatesan, S., & Basavaraj, V. (2009). Ethical guidelines for bio behavioural research. *All India Institute of Speech and Hearing.*

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., & Eddington, D K Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature, 352*(6332), 236–238.

Yellamsetty, A. (2016). Temporal fine structure frequency bands criticality in perception of the speech in the presence of noise. *Indian Journal of Otology, 22*(2), 92–99. https://doi.org/10.4103/0971-7749.182274