Research Article

Perception of Temporal Fine Structure Speech and Recovered Envelope Speech in Younger and Older Adults with Normal Hearing Sensitivity

JAIISH(2019)
Vol 38 pp. 58-66

Pavan Mahadevaiah[a], Indira C. Pazhayapisharath[b], Arivudai Nambi[c] &
Rajalakshmi Krishna[b]

Affiliations

[a]Resonance Speech and Hearing Clinic, Mysuru, Karnataka, India
[b]All India Institute of Speech and Hearing, University of Mysore, Manasangangothri, Mysore, 570006, Karnataka, India
[c]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
RENV speech
Young normal hearing
Old normal hearing

Abstract

Perception of temporal fine structure (TFS) speech and recovered envelope (RENV) speech were compared between seven younger adults and seven older adults with normal hearing sensitivity. To create TFS speech, Kannada sentences were processed to remove temporal envelope cues while retaining TFS cues in 2, 4 and 8 frequency bands spanning the range of 80 to 8020 Hz. To create RENV speech, envelope cues were recovered from TFS speech extracted from 2 frequency bands, by passing it through 40 band-pass filters. RENV speech was also generated with simulated widened auditory filters, 2 and 4 times the normal auditory filter bandwidth. The findings show a general trend of reduction in scores with increase in number of frequency bands in TFS speech. The scores were significantly different across conditions when the age groups are combined, whereas significant difference between age groups was seen only for TFS speech extracted from 4 bands. Similarly, the scores reduced with simulation of cochlear hearing loss for perception of RENV speech. The scores were significantly different across the RENV speech conditions in both age groups. No significant difference was seen between the two age groups in any of the RENV speech conditions. The results of the study indicate that the ability to perceive TFS cues for sentence identification would degrade significantly with increase in the severity of cochlear pathology, without any significant age effect.

INTRODUCTION

Speech is a complex signal and its perception by the human brain depends on the peripheral auditory system's ability to decode the acoustic cues present in it. The acoustic cues of speech may be primarily divided into spectral and temporal cues (Moon & Hong, 2014). The temporal cues consist of temporal envelope (ENV) and temporal fine structure (TFS). ENV is also called the 'modulator', and refers to the slow amplitude variations of the speech signal over time obtained at the output of cochlear frequency bands.(Ardoint, Sheft, Fleuriot, Garnier, & Lorenzi, 2010; Moon & Hong, 2014; Swaminathan et al., 2016). The TFS, also called the 'carrier', involves rapid oscillations with a rate close to the center frequency of the frequency band of the signal. Both ENV and TFS are known to be critical for speech understanding, especially in the presence of background noise (Ardoint et al., 2010).

Researchers have tried to understand the independent roles of TFS and ENV in speech perception (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Smith, Delgutte, & Oxenham, 2002; Swaminathan & Heinz, 2012) by extracting them separately (techniques like the Hilbert transform can be used to extract the ENV and TFS components), and assessing the resultant speech perception. These findings indicate that ENV cues, when presented alone, are sufficient to understand speech in quiet (Shannon et al., 1995); but are not in the presence of noise (Loizou, Dorman, & Fitzke, 2000). This was true for fluctuating as well as steady state noise (Loizou et al., 2000; Moore, Glasberg, Flanagan, & Adamas, 2006). In the presence of background noise, the availability of TFS information is found to aid speech perception. Further, speech can be understood to a certain extent even when only TFS speech is presented (Moore, 2019). However, TFS speech intelligibility is good when TFS is extracted from wide frequency bands (Drullman, Festen, & Plomp, 1994a, b; Drullman, 1995; Smith et al., 2002), but reduces drastically when extracted from narrow frequency bands. Nevertheless, the type and extent of contribution of TFS during sentence perception remains unclear.

Speech identification using extracted TFS information cannot be considered as an evidence of the contribution of TFS information itself, since it has been shown that temporal envelope is reconstructed at the level of the auditory filters even when only TFS information is presented (Apoux, Youngdahl, Yoho, & Healy, 2013; Apoux, Millman, Viemeister, Brown, & Bacon, 2011).

Ghitza (2001) based his suggestion on two theories- the first one states that Hilbert instantaneous amplitude and phase are related (Voelcker, 1996), and the second one states that the output of a band-pass filter is an envelope related to the phase of Hilbert fine structure passed through the band-pass filter (Rice, 1973). Therefore, if we consider the cochlear filters as bandpass filters, and TFS information from speech is passed through these filters, the output could be the envelope of the corresponding TFS. Using psychophysical experiments, Apoux et al. (2011) showed that individuals with normal hearing could perceive amplitude modulations when only Hilbert fine structure of amplitude modulated signals were presented to them. However, they attributed the recovery of envelope to the retention of some envelope information. Later, Apoux et al. (2013) showed that in individuals with normal hearing sensitivity as well as in individuals with hearing loss, good speech recognition was achieved from recovered envelopes for up to 15 analysis bands. They state that the possible role of envelope recovery cannot be excluded in understanding TFS speech. Therefore, understanding the contribution of TFS to speech perception is a topic of intrigue.

Simulated cochlear frequency bands can be used to recover temporal envelope from TFS cues. This speech, recovered from TFS is termed 'recovered envelope speech' (RENV speech). The maximum number of frequency bands that results in envelope recovery with good speech intelligibility varies from 8 (Gilbert & Lorenzi, 2006) to 20 (Chen, Tsao, & Lai, 2016). Chen et al. (2016), in their study, synthesized TFS stimuli and used a computation measure to predict the intelligibility of RENV speech. These computed speech intelligibility scores were compared with speech intelligibility in the presence of noise from individuals with normal hearing. They reported good speech intelligibility for the processed speech for up to 20 frequency bands used for TFS extraction. However, this number also depends upon the rate of amplitude modulation within the bands.

Alternatively, studies have shown that TFS cue by itself, contributes to speech perception (Hopkins, Moore, & Stone, 2010; Moore, 2019; Sheft, Ardoint, & Lorenzi, 2008). Sheft et al. (2008) used three stimulus processing paradigms (termed the Phase modulation, Frequency modulation and Envelope only conditions) and examined the fidelity of envelope reconstruction, and compared it to the perception of TFS speech. They found that the contribution of TFS to perception of speech segments was not only due to envelope reconstruction, since their manipulations that varied envelope cues did not affect the perception of TFS cues. Their study showed that TFS cues were used more for place and manner perception compared to ENV cues. Other studies that have explored TFS perception also show that TFS information does contribute to speech intelligibility (Hopkins & Moore, 2010; Moore, 2019; Sheft et al., 2008). While Hopkins and Moore (2010) observed that TFS cues contribute to speech perception across a wide frequency range, Sheft et al. (2008) observed that TFS cues contribute to perception of place cues in the identification of consonants.

The perception of TFS is influenced by advanced age, even in the absence of hearing loss (Moore, 2019). In a study of TFS perception in different age groups, Hopkins and Moore (2011) compared the sensitivity to TFS and frequency selectivity in younger (20-35 years) and older (63-66 years) adults with normal hearing sensitivity. They used the Temporal Fine Structure-1 (TFS1; Moore & Sek, 2009) and Temporal Fine Structure-Low Frequencies (TFS-LF; Hopkins & Moore, 2011) (fast methods to measure sensitivity to TFS information at high and low frequencies, respectively) tests to check for TFS sensitivity. The frequency selectivity of the participants of the younger and older groups was comparable. But, the older group had significantly poorer performances on the two tests assessing sensitivity to TFS. In a follow-up study with the same tests, with participants in the age group of 22 to 61 years. Moore et al. (2012) observed good negative correlation between age and sensitivity to TFS.

Studies have also assessed sensiti,Moore, & Stone, 2015vity to TFS information in different age groups and compared it to their speech perception abilities (Füllgrabe et al., 2015; Strelcyk & Dau, 2009). Füllgrabe et al. (2015) observed differences in sensitivity to TFS information in younger (18–27 years) and older adults (60–79 years) and state that it is a good predictor of speech perception in noise.

The ability to perceive TFS and use TFS cues effectively is impaired in individuals with cochlear hearing loss (Lorenzi et al., 2006; Ardoint et al., 2010). Lorenzi et al. (2006) observed that individuals with hearing loss faced difficulty in understanding TFS speech (but not unprocessed speech or ENV vocoded speech). Their performance in TFS speech perception also reflected on their ability to perceive speech in noise. In addition, their difficulty to perceive TFS cues may be speculated to adversely influence their speech perception. One may expect this difficulty in processing TFS information to adversely influence day-to-day communication as well as their performance during clinical testing.

The role of TFS in speech understanding lacks consensus in terms of the extent of contribution and its importance in different auditory environments. Factors like advanced age and cochlear hearing loss influence the perception of TFS information. Even though studies have compared sensitivity to TFS information in different age groups and explored their implications on speech perception in noise, TFS speech and RENV speech have not been employed in this accord, nor have they used sentences as stimuli for the purpose. It is also seen that the parameters used for TFS extraction and envelope recovery for speech stimuli influence their intelligibility, but the impact of the same in younger and older adults has not been studied. Further, studying these factors using speech units like sentences will be able to provide a clearer understanding of the abilities to use TFS cues during real-world communication.

Perception of TFS speech is reported to be influenced by envelope recovery (Apoux et al., 2013; Ghitza, 2001). TFS perception, as seen earlier, is influenced by advanced age as well as by hearing loss. Therefore, differences in TFS perception abilities as a factor of age may be expected to reflect the abilities to recover envelope. Further, widening of auditory filters characteristic of cochlear hearing loss, influences the recovery of envelope and contributes to inferior TFS perception (Heinz & Swaminathan, 2009). This change in recovery of envelope seen in cochlear hearing loss can be simulated in the RENV speech, by widening the recovery bands by a certain factor. The resultant RENV speech is likely to reflect the perception of TFS speech in individuals with cochlear hearing loss (Léger, Desloge, Braida, & Swaminathan, 2015; Swaminathan & Heinz, 2012). By comparing the perception of RENV speech using simulated widened auditory filters between younger and older adults, we can derive the perceptual weightage of TFS cues in speech perception. Findings from these comparisons will add to our understanding of perception of TFS cues recovered from widened auditory filters in different age groups.

Therefore, objectives of the present study were 1) to compare the perception of TFS speech using sentence stimuli in younger and older adults with normal hearing sensitivity, and 2) to compare the perception of RENV speech in younger and older adults with, and without simulation of widened auditory filters.

METHODS

Speech material

Recorded sentence lists from the standardized ‘Sentence identification test in Kannada’ developed by Geetha, Kumar, Manjula, and Pavan (2014) were used to prepare the stimulus for the study. The corpus consisted of 24 lists with 10 sentences each and each sentence had 4 key words to be scored. There were 14 to 16 syllables in each sentence. The language Kannada was used considering the geographical location the study was conducted in.

Participants

This study reports the findings from a pilot data collected to answer the research questions. An a-priori power analysis indicated 17 participants to reach 80% power with medium effect size and 0.05 alpha level. However, data from 14 normal hearing individuals with air conduction audiometric thresholds within 15 dB hearing level (HL) between 250 and 8000 Hz are presented in this study. The participants belonged to two groups- the Young normal hearing (YNH, age range from 27 to 33 years, average hearing threshold- 8.6 dB HL) group and the Old Normal hearing (ONH, age range from 57 to 63 years, average hearing threshold- 14.3 dB HL) group. Each of the two groups included seven participants. All the participants’ hearing thresholds were tested in a sound treated room using a calibrated audiometer (Maico MA-52 Diagnostic audiometer), under TDH 39 headphones. They had A-type tympanogram along with the presence of acoustic reflexes, indicating normal functioning of middle ear as tested using a tympanometer (Maico easyTymp), and all of them passed OAE screening (Bio-logic Natus AuDX PRO), indicating functional outer hair cells. All the tests for inclusion were carried out bilaterally. None of the participants reported any history of hearing problems, difficulty in comprehension, or memory loss. They were native speakers of Kannada (a language spoken in the south Indian state of Karnataka) and had good comprehension of the spoken language as well as the written script in Kannada. The study abided by the ethical guidelines for bio-behavioural research in human subjects (Venkatesan, 2009) and an informed consent was signed by all the participants before their participation in the study. Additionally, 2 unprocessed, randomly selected sentence lists (from the corpus selected for the study) were presented to the participants at their most comfortable level for listening. Only those participants who obtained scores higher than 95% from these lists took part in further testing.

Stimulus processing

The original sentence lists were used to generate stimuli to assess perception of TFS speech, RENV speech and RENV speech simulating widened auditory filters. The methods used for the processing are similar to Swaminathan, Reed, Desloge, Braida, & Delhorne, (2014), and are detailed below.

TFS speech: Each sentence was band-pass filtered using third order elliptical filter into 2, 4 and 8 frequency bands within 80 to 8020 Hz following logarithmic spacing within the bandwidth. The signals were forward and backward filtered to avoid phase delays. Hilbert transform was applied to the signal in each frequency band and the signal was separated into the component envelope and TFS.

In order to compensate for the amplitude loss of TFS when separated from the envelope, the TFS was multiplied with the RMS power of the band-pass filtered signal. The amplitude-corrected TFS was summed across frequency bands to create the final TFS speech. This processing resulted in three different conditions to test for TFS speech perception - TFS speech extracted from 2 frequency bands (TFS2nb), 4 frequency bands (TFS4nb), and 8 frequency bands (TFS8nb).

RENV speech: The TFS speech from TFS2nb condition was passed through a bank of 40 band-pass filters (1 ERB wide), with center frequencies varying from 80 to 8020 Hz. Hilbert transform was applied to the output from each frequency band to extract the envelope. The extracted envelope was low-pass filtered using second order Butterworth filter. Backward and forward filtering were used here also to avoid phase shiftd. The resultant envelope was used to modulate sinewave with frequency of the center frequency of the corresponding filter band (but with random starting phase). The output from each band was then combined to create RENV stimulus. In order to understand the influence of widening of auditory filters in envelope recovery, the recovery of the envelope from TFS2nb condition was carried out using the same procedure, but by implementing a widening factor of 2 and 4 to the number of recovery bands. This resulted in RENV speech with widening factors of 2 and 4 termed as ‘RENV2wf’ and ‘RENV4wf’ resulted in 3 different test conditions, namely RENV, RENV2wf and RENV4wf.

Experimental Procedure

Each participant was tested in six stimulus conditions-three under TFS speech (TFS2nb, TFS4nb and TFS8nb), and three under RENV speech (RENV, RENV2wf and RENV4wf). During an initial test on 3 individuals with normal hearing, perception of TFS speech extracted from bands more than 8 showed floor effect. Therefore, testing was not continued with TFS speech extracted from higher bands. The participants were seated comfortably in a sound treated room. They were instructed to carefully listen to the speech stimuli presented and repeat verbatim. Before the actual test session, the participants were presented 2 unprocessed lists of sentences to familiarize them with the task of repeating the stimulus verbatim. Following this, 2 sentence lists were randomly selected and presented in each stimulus condition, from a Lenovo laptop computer (Lenovo ThinkPad X1 Carbon, third Gen with interl core i7). The stimuli were delivered to the participants’ ears using HDA200 headphones at 70 dB SPL. The overall levels were calibrated using a Bruel & Kjaer (2250) sound level meter and ear simulator, complying with IEC 60318-1. Each participant, therefore, responded to 14 sentence lists, presented across six test conditions and familiarization. The test conditions were randomized across participants. The testing was completed in a single sitting and breaks were given to the participants whenever necessary.

The responses were recorded using a custom program written to record the speech output using MATLAB software version 2019 (Mathworks Inc., Natick, MA, USA). The recorded responses were analyzed by one examiner (a native speaker of Kannada) and the key words were scored in each sentence. Each correctly repeated key word was given a score of 1 (maximum achievable score in each condition, from 2 sentence lists was 80) and errors or skipped words were given a score of 0. Scoring method followed was as per the instructions given in the test material employed. The speech perception testing and response analysis were carried out by different investigators, to eliminate tester bias.

RESULTS

The word identification scores were the dependent variables, whereas age groups and the different testing conditions were the independent variables. As the statistical power is less due to small sample size, individual data trend is analyzed followed by planned pairwise comparisons. We also present the results of mixed ANOVA which explore the interaction effect between age and perception of RENV speech in different conditions (RENV tests) and age and use of TFS information (TFS tests). Mixed ANOVA is chosen, as there is no non-parametric equivalent which explores the interaction effect.

The data was normally distributed (based on Shapiro-Wilk test of normality) and homogeneity of variances was observed on Levene’s test of normality ($p>0.05$). The individual data, mean data, and standard deviation (SD*)* as well as median data and interquartile range of speech perception scores obtained during test conditions using TFS speech (TFS2nb, TFS4nb and TFS8nb) and RENV speech (RENV, RENV2wf, RENV4wf) are given in Figure 1 and Figure 2 respectively. The data shows a general trend of reduction in scores with increase in number of frequency bands in TFS speech. When data for the two age groups are compared, we can see that YNH group shows higher speech identification scores compared to the ONH group in each of the three test conditions. Upon visual inspection, the individual data also look separated between the two groups, especially for TFS2nb and TFS4nb conditions.

Mixed ANOVA was used to compare the data within groups and across the groups in different test conditions. Mauchly’s test of sphericity indicated that assumption of sphericity was satisfied in both test conditions (TFS speech: $\chi2 = 4.618$, $p = .099$; RENV speech: $\chi2 = 1.765$, $p = .414$), and therefore, no sphericity corrections were added. There was no interaction between the test conditions and age group. Due to the small sample size, effect size and poor statistical power obtained for the test, further analysis were not done.

Similarly, the scores reduce with increase in the critical bandwidth in the cochlea (simulation of cochlear hearing loss) during the perception of RENV speech. The mean and median scores are lower in

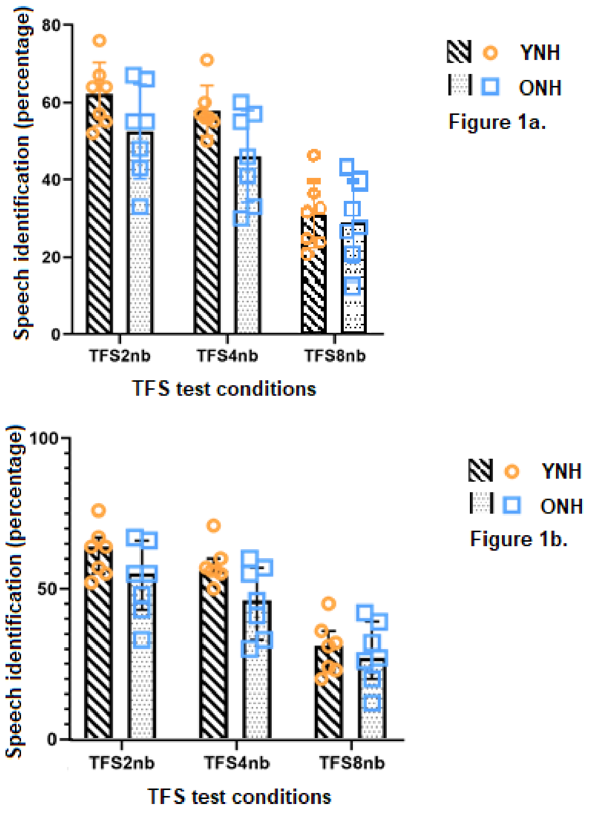

Figure 1: (a) Mean and SD, and (b) Median and inter-quartile range of speech identification scores in test conditions using TFS speech, in YNH and ONH groups, along with individual data

Table 1: Results of Wilcoxon's test for pair-wise comparisons of the different TFS and RENV speech conditions in the age groups combined.

					95% CI for Effect Size	
Condition	Test	/z/	*p* value	Effect Size	Lower	Upper
TFS2nb	TFS4nb	94.000	0.010	0.790	0.445	0.931
TFS8nb	TFS4nb	0.000	0.001	-1.000	-1.000	-1.000
TFS2nb	TFS8nb	105.000	0.001	1.000	1.000	1.000
RENV	RENV2wf	90.000	0.002	0.714	0.292	0.903
RENV2wf	RENV4wf	105.000	0.001	1.000	1.000	1.000
RENV	RENV4wf	105.000	$< .001$	1.000	1.000	1.000

Note: effect size is given by the matched rank biserial correlation.

Table 2: Results of Welch test to see the effect of age on TFS and RENV test conditions.

Condition	*t*	df	*p* value	Effect Size	Lower	Upper
TFS2nb	1.747	10.489	0.110	0.934	-0.205	2.034
TFS4nb	2.314	9.308	0.045	1.237	0.026	2.396
TFS8nb	0.362	11.597	0.724	0.193	-0.861	1.240
RENV	0.668	10.698	0.518	0.357	-0.709	1.407
RENV2wf	0.976	10.357	0.351	0.522	-0.561	1.581
RENV4wf	0.341	10.136	0.740	0.182	-0.873	1.229

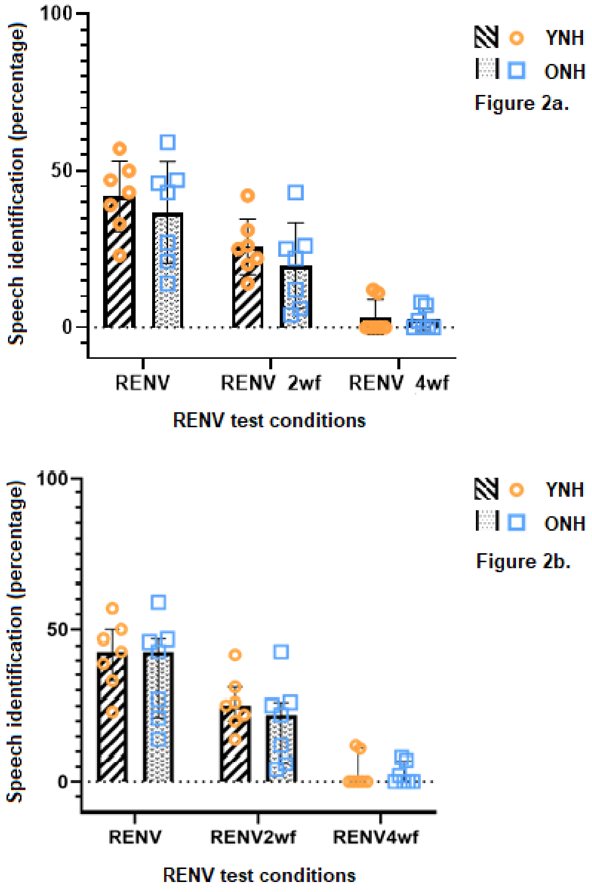

Figure 2: (a) Mean and SD, and (b) Median and inter-quartile range of speech identification scores in test conditions using RENV speech conditions, in YNH and ONH groups along with individual data

the ONH group compared to the YNH group (Figure 2). However, when we observe the individual scores, there is a considerable overlap in the scores for the two groups.

To examine the effect of test conditions and age on word identification scores, Wilcoxon, and Welch tests were used respectively. Welch test was used since it can be used for continuous variables whose variances are unequal. The scores from YNH and ONH groups were used for between group comparisons in TFS and RENV speech perception conditions. The results of test condition comparisons and age comparisons are as shown in Table 1 and Table 2 respectively.

The Wilcoxon test results show significant difference between all the test conditions in both the age groups combined. The Welch test results showed significantly lower speech identification score for ONH group compared to YNH group for TFS4nb condition.

DISCUSSION

The main aim of the study was to examine the effect of the number of frequency bands used for TFS extraction on TFS speech and envelope recovery on RENV speech perception in younger and older adults with normal hearing sensitivity. The study compared the perception of TFS speech and RENV speech in younger and older normal hearing adults, using sentence stimuli. Considering that the power observed for the study is poor, the focus was on observing the trends in individual data, even though group comparisons were done to establish statistical significance, if any. Additionally, interaction between test conditions and age was also studied. No interaction was observed between the two age groups and the test conditions. The individual trend in data showed a decrease in speech perception scores with increase in number of frequency bands used for extraction in TFS speech and with widening of recovery bands in RENV speech. The participants' speech perception scores were significantly different between the different test conditions using TFS speech as well as that of RENV speech.

An increase in the number of frequency bands used to create TFS speech results in progressive reduction of speech perception scores (Smith et al., 2002), as is observed in the study. These findings are in agreement with the literature (Drullman et al., 1994a, b; Smith et al., 2002). TFS information, when presented alone, is effectively used in speech identification when it is extracted from less number of frequency bands. In other words, TFS speech intelligibility reduces with increase in the number of extraction bands, and beyond a certain level, there is minimal speech intelligibility (Smith et al., 2002). In their study, Smith et al. (2002) examined the relative contribution of TFS and ENV information to speech perception and the influence of number of bands of extraction of the same. They used speech-noise and speech–speech chimeras (mixing TFS from one signal and ENV from another), and showed that TFS was used to comprehend speech only when the extraction of information was carried out from 2 or 4 frequency bands. The intelligibility of RENV speech also depends upon the number of frequency bands from which the TFS information is extracted (Swaminathan et al., 2014). One of the experiments carried out by Swaminathan et al. (2014) compared the perception of RENV speech using RENV bands extracted from TFS ranging from 1 to 32. The results showed that envelope recovery deteriorated with increase in the number of bands from 1 to 32. Similar findings are observed in the present study, thereby reiterating that perception of TFS speech is better when it is extracted from lesser number of frequency bands. Further, these findings were found to be true for both younger and older adults with normal hearing.

The present study also compared the perception of TFS speech in different test conditions between the two age groups. This, however, was significantly different only in the TFS4nb condition. This could be because of ceiling effect in the scores in the TFS2nb condition and floor effect in the TFS8nb condition. It is plausible that a significant difference was observed only in conditions with optimal difficulty.

Scores obtained for the perception of RENV speech in the present study were comparable to those obtained by Swaminathan et al. (2014). They observed similar outcomes in their experimental condition with envelope recovered from TFS extracted from one frequency band and 40 ERBN, despite the differences in the stimuli used. They had used monosyllables and the present study used sentences. The intelligibility of RENV speech also deteriorated with widening of auditory filter bands. This finding is in agreement with studies that have examined the effects of widening of cochlear auditory filters on recovery of envelope from TFS speech (Heinz & Swaminathan, 2009; Swaminathan et al., 2014). Widened auditory filters that result in reduced frequency resolution is also characteristic of cochlear hearing loss (Glasberg & Moore, 1986). Simulation studies (Heinz & Swaminathan, 2009) and studies that explored perception of recovered envelopes (Lorenzi et al., 2012; Swaminathan et al., 2014) have shown that perception of RENV speech deteriorates with the widening of RENV bands, as was the trend observed in the present study, in both younger and older participants.

It was also of interest to see if the perception of the RENV speech varied between the two groups, given that the perception of TFS varies between younger and older adults despite normal hearing sensitivity (Hopkins & Moore, 2011; Peters & Moore, 1992; Strelcyk & Dau, 2009), and that widening of recovery bands for RENV speech may reflect perception of speech in ears with hearing loss. The results, however, did not show any significant difference between the two age groups in their perception of RENV speech, with or without the simulation of widened auditory filters. Age group of the participants in the present study could be one reason why no significant differences were obtained between the groups. There is no published literature to directly compare the findings of the present study in this regard. However, studies that have shown differences in the perception of TFS cues in the elderly have used different measures of TFS perception. It should also be noted that the participants of the previous studies were older than in the present study. Additionally, floor effect in the scores obtained for RENV4wf condition and the number of participants in the study may be contributing factors. A clearer picture of the trend may be seen when the tests are administered on a larger group of participants.

Another important factor to consider in this study is the use of sentences to study the perception of TFS speech across two age groups. This is essential, since stimulus complexity does influence the intelligibility of TFS speech (Swaminathan et al., 2014). For example, Swaminathan et al. (2014) noted that speech perception scores in their subjects decreased systematically with increase in the number of bands for TFS extraction, with a prompt drop for bands beyond eight. In the present study, however, floor effects were observed for RENV speech perception even for speech recovered from 4 TFS bands (which was why it was not used for further testing in the participants).

However, findings of the present study should be compared with the existing literature with caution for two reasons- first, the stimulus used to study envelope recovery were sentences in the present study, whereas they were monosyllables in similar studies in the literature; and second, the number of bands for TFS extraction was between 2 and 8 in the present study due to floor effects in obtained scores, whereas other studies have used higher number of bands for the same (though not for sentence stimuli) (Swaminathan et al., 2014). Further, difference in language may also influence the amount of information that is used for speech understanding, as well as recovery of envelope.

CONCLUSIONS

The study compared the perception of TFS cues using processed speech stimuli in two different age groups. The perception of TFS information from TFS speech and RENV speech worsened with more degradation of the speech stimuli in the YNH and

ONH groups. Significant difference between the groups could be seen only in one condition, and it sheds light on the perception of TFS cues from a more complex unit of speech than what has been studied so far. This observation needs to be validated, on a larger population and from a wider age range to further our understanding of influence of age on TFS and RENV speech perception.

ACKNOWLEDGEMENTS

The authors would like to thank the participants of the study for their cooperation and time.

REFERENCES

Apoux, F., Millman, R. E., Viemeister, N. F., Brown, C. A., & Bacon, S. P. (2011). On the mechanisms involved in the recovery of envelope information from temporal fine structure. *The Journal of the Acoustical Society of America*, *130*(1), 273–282. https://doi.org/10.1121/1.3596463

Apoux, F., Youngdahl, C. L., Yoho, S. E., & Healy, E. W. (2013). Can envelope recovery account for speech recognition based on temporal fine structure? *The Journal of the Acoustical Society of America*, *133*(5), 3380–3380. https://doi.org/10.1121/1.4805822

Ardoint, M., Sheft, S., Fleuriot, P., Garnier, S., & Lorenzi, C. (2010). Perception of temporal fine-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss. *International Journal of Audiology*, *49*, 823–831. https://doi.org/10.3109/14992027.2010.492402

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

Chen, F., Tsao, Y., & Lai, Y. H. (2016). Modeling speech intelligibility with recovered envelope from temporal fine structure stimulus. *Speech Communication*. https://doi.org/10.1016/j.specom.2016.01.006

Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of temporal envelope smearing on speech reception. *Journal of Acoustical Society of America.*, *95*, 1053– 1064.

Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *The Journal of the Acoustical Society of America*, *97*, 585–592. https://doi.org/10.1121/1.413112

Drullman, Rob, Festen, J. M., & Plomp, R. (1994b). Effect of reducing slow temporal modulations on speech reception. *Journal of the Acoustical Society of America*, *95*(5), 2670–2680. https://doi.org/10.1121/1.409836

Eaves, J. M., Quentin Summerfield, A., & Kitterick, P. T. (2011). Benefit of temporal fine structure to speech perception in noise measured with controlled temporal envelopes. *The Journal of the Acoustical Society of America*, *130*(1), 501–507. https://doi.org/10.1121/1.3592237

Fogerty, D., & Entwistle, J. L. (2015). Level considerations for chimeric processing: Temporal envelope and fine structure contributions to speech intelligibility. *The Journal of the Acoustical Society of America*, *138*(5), EL459–EL464. https://doi.org/10.1121/1.4935079

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. *Journal of Hearing Science ®*, *4*(1), OA18–OA26. http://www.journalofhearingscience.com/abstract/index/idArt/890267

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*(3), 1628. 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

Glasberg, B. R., & Moore, B. C. J. (1986). Auditory filter shapes in sub- jects with unilateral and bilateral cochlear impairments. *Journal of Acoustical Society of America*, *79*, 1020–1033.

Heinz, M. G., & Swaminathan, J. (2009). Quantifying envelope and fine-structure coding in auditory nerve responses to chimaeric speech. *JARO - Journal of the Association for Research in Otolaryngology*. https://doi.org/10.1007/s10162-009-0169-8

Hopkins, K., & Moore, B. C. J. (2010). 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. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The 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.* https://doi.org/10.1121/1.3478773

Joris, P. X., & Yin, T. C. (1992). Responses to amplitude-modulated tones in the auditory nerve of the cat. *The Journal of the Acoustical Society of America*, *91*, 215–232. https://doi.org/10.1121/1.402757

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*(L), 505–508. https://doi.org/10.1121/1.4904540

Loizou, P. C., Dorman, M., & Fitzke, J. (2000). The effect of reduced dynamic range on speech understanding: implications for patients with cochlear implants. *Ear and Hearing*, *21*, 25–31. 10.1097/00003446-200002000-00006

Lorenzi, C., Wallaert, N., Gnansia, D., Leger, A. C., Ives, D. T., Chays, A., Garnier, S., & Cazals, Y. (2012). Temporal-envelope reconstruction for hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology : JARO*, *13*(6), 853–865. https://doi.org/10.1007/s10162-012-0350-3

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. 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, B. 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, B. 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, B. 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

Peters, R. W., & Moore, B. C. J. (1992). Auditory filters and aging: Filters when auditory thresholds are normal. In Y. Cazals, L. Demany, & K. Horner (Eds.), *Auditory Physiology and Perception* (pp. 179–185). Oxford.

Rice, S. O. (1973). Distortion produced by band limitation of an FM wave. *Bell System Technical Journal*, *52*, 605–626.

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

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

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. *Journal of Acoustical Society of America*, *125*, 3328–3345.

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., Mason, C. R., Streeter, T. M., Best, V., Roverud, E., & Kidd, G. (2016). Role of binaural temporal fine structure and envelope cues in cocktail-party listening. *Journal of Neuroscience*, *36*(31), 8250–8257. https://doi.org/10.1523/JNEUROSCI.4421-15.2016

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. (2009). *Ethical guidelines for bio-behavioral research involving human subjects* (pp. 1–23). All India Institute of Speech and Hearing. https://doi.org/10.1017/CBO9781107415324.004

Voelcker, H. B. (1996). Towards a unified theory of modulation. I. Phase- envelope relationships. *Proceedings of the IEEE*, *54*(3), 340–354.

Won, J. H., Lorenzi, C., Nie, K., Li, X., Jameyson, E. M., Drennan, W. R., & Rubinstein, J. T. (2012). The ability of cochlear implant users to use temporal envelope cues recovered from speech frequency modulation. *The Journal of the Acoustical Society of America*, *132*(2), 1113–1119. https://doi.org/10.1121/1.4726013