Article17

by Author Author

Submission date: 21-Jul-2020 05:11PM (UTC+0530)
Submission ID: 1360363523
File name: 17.docx (98.94K)
Word count: 3052
Character count: 16364

Temporal Fine Structure Speech and Recovered Envelope Speech Perception in Younger and Older Individuals with Normal Hearing Sensitivity.

Background

Speech is a complex signal and its interpretation by the human brain depends on the auditory system's ability to decode the acoustical cues present in it. These acoustic cues 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) that are critical for speech understanding, especially in the presence of background noise (Ardoint, Sheft, Fleuriot, Garnier, Sheft, et al., 2010).

The sound that reaches the inner ear passes through a bank of band pass filters. ENV, also called the 'modulator' is the slow amplitude variations of the speech signal over time obtained at the output of these bands(Ardoint, Sheft, Fleuriot, Garnier, & Lorenzi, 2010; Moon & Hong, 2014; Swaminathan et al., 2016). The TFS, also called the 'carrier', involves rapid oscillations with rate close to the center frequency of the frequency band of the signal.Both ENV and TFS cues are coded as time related changes at the level of the auditory neurons:ENV cues are extracted from the amplitude variations in the neural firings or the short term rate of action potentials (Joris & Yin, 1992);TFS cues are obtained from the phase locking information, the precise timing of the action potentials (Buss et al., 2004; Heinz & Swaminathan, 2009; Joris & Yin, 1992; Moon & Hong, 2014).

Researchers have tried to understand the contribution of TFS and ENV components to speech perception (Smith et al., 2002; Swaminathan & Heinz, 2012) by separating them, presenting the extracted component, and observing the speech perception(techniques like the Hilbert transform can be used to extract the ENV and TFS components).Such studies have shown that ENV cues are sufficient to understand speech in quiet(Shannon et al., 1995). However, in the presence of noise speech perception deteriorates when only the ENV cues are provided(Loizou et al., 2000). This is true for fluctuating as well as steady state noise(Loizou et al., 2000; Moore et al., 2006). In such situations, providing TFS information improves speech perception(Eaves et al., 2011; Fogerty & Entwistle, 2015). However, the exact contribution of TFS to speech intelligibility is still unclear.

Intelligibility of speech with TFS information alone is good when the TFS is extracted from wide frequency bands(Drullman et al., 1994; Drullman, 1995; Drullman et al., 1994; Smith et al., 2002). This intelligibility decreases drastically when the TFS extraction is done from narrow, more number of frequency bands. However, speech identification using this extracted TFS information cannot be used as a proof for contribution of TFS information itself, since it has been shown that temporal envelope is reconstructed at the level of the auditory filters evenwhen only TFS information is presented (Ghitza, 2001). Therefore, understanding the contribution of TFS to speech perception is a topic of intrigue.

The number of frequency bands required for envelope recovery with good speech intelligibilityvaries from 8(Gilbert & Lorenzi, 2006) to 20 (Chen et al., 2016). Chen et al. (2016) reported that good speech intelligibility it is seen for up to 20 frequency bands used for TFS extraction. But this depends upon the rate of amplitude modulation within the bands. Such recovered speech from TFSis named ‘recovered envelope speech’. The cues from recovered envelope are used effectively. Study by Sheft et al.(2008) to explore usefulness of TFS alone showed that TFS cues were used more for place and manner perception compared to ENV cues. And this contribution was not only due to envelope reconstruction(Sheft et al., 2008). Other studies that have explored TFS perception also show that TFS information does contribute to speech intelligibility(Hopkins et al., 2010; Moore, 2019; Sheft et al., 2008).

Impaired ability to perceive TFS and use TFS cues effectively will therefore adversely affect speech intelligibility, as seen in individuals with cochlear hearing loss and in elderly individuals. It is noticeable that 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, sensitivity to TFS and frequency selectivity were compared in young (20-35 years) and older Individuals (63-66 years) with normal hearing sensitivity(Hopkins & Moore, 2011). They used the TFS1 and TFS-LF tests (Hopkins & Moore, 2010; Moore & Sek, 2009)to check for TFS sensitivity. The frequency selectivity of the participants of the younger and older groups was comparable. But, the older grouphad significantly poorer performance on the two tests assessing sensitivity to TFS. In a follow up study with the same tests, but with slight changes in the test parameters,Moore et al.(2012)observed good correlation between age and sensitivity to TFS.

These studies explored the sensitivity to TFS information using complex tones. A number of studies assessing sensitivity to TFS information using speech stimulus in individuals of different ages have estimated the sensitivity in the presence of different types of noises (Füllgrabe & Moore, 2014; Peters & Moore, 1992; Strelcyk & Dau, 2009),using different speech stimuli. The contribution of TFS to the perception of speech in quiet in different age groups is not well understood. Further, the difference in the abilities of perception of TFS as a factor of age must reflect on their abilities of recovery of envelope, and might become more enhanced when the reconstruction of the envelope is carried out with widened auditory filters, simulating widened cochlear filters.

Therefore, objectives of the present study were 1) to compare the perception of TFS speech using sentence stimuli in young and older individuals with normal hearing sensitivity, and 2)to compare their perception of RENV speech with and without simulated hearing loss.

Materials and Methods

Speech material

Recorded sentence lists from the standardized ‘Sentence identification test in Kannada’developed by Geetha et al. (2014) were used to prepare 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.

Stimulus processing

The original sentencelists were subjected to threedifferent kinds of processing. Under the first kind of processing the TFS information was extracted from the sentences. Using theextracted TFS from each sentence, the envelopes of the stimuli were recovered in the second kind of processing. In the third kind of processing too, envelope recovery was carried out from the TFS, but this time it was done by simulating widened auditory filters, thereby simulating cochlear hearing loss.

TFS speech:Each sentence was band pass filtered using 3rd order elliptical filterinto 2, 4 and 8 frequency bands within 80-8020Hz 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 and the envelope was discarded. The extractedTFS loses its amplitude when the amplitude is removed. To compensate for this, the TFS was multiplied with the RMS power of the band-pass filtered signal. This amplitude-corrected TFS was summed across frequency bands to create the final TFSspeech. This processing resulted in three different conditions to test for TFS speech perception, namely TFS2nb, TFS4nb and TFS8nb.

RENV speech: The TFS speechTFS2 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 2nd order Butterworth filter. Backward and forward filtering was used here also to avoid phase shift. The resultant envelope was used to modulate sinewave with frequency of the center frequency of the corresponding filter band(but random starting phase). The outputs from each band was then combined to create 'RENV' stimulus. Separately, the recovery of the enevelope from TFS2 condition was carried out using the same procedure, but by implementing a widening factor of 2 and 4, resulting in 'RENV2' and 'RENV4' test conditions respectively.Therefore, this processing resulted in 3 different test conditions, namely RENV, RENV2wf and RENV4wf.

Participants

14 individuals with audiometric thresholds within 15 dBbetween 250 and 8000 Hz participated in the study. The participants belonged to two groups of 7 individuals each- young normal hearing (YNH, age range from 27 to 33 years)group and old normal hearing (ONH, age range from 57 to 63 years) group.All the participants' hearing thresholds were tested in a sound treated room using a calibrated audiometer (MaicoMA-52 Diagnostic audiometer).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 the ethical guidelines for bio behavioral research in human subjects(Venkatesan, 2009)and aninformed consent was signed by all the participants before their participation in the study. Additionally, only those participants who could repeat two lists of randomly selected, unprocessed sentences from the corpus selected for the study with 95% of accuracy or higher, when presented at their most comfortable level for listening were taken for further tests.

Experimental Procedure

The sentence listswere processed following the procedure mentioned above. Each participant wastested in six stimulus conditions-three under TFS speech, and three under RENV speech perception conditionsand RENV speech with simulated cochlear hearing loss. The participants were seated comfortably in a sound treated room. They were instructed to listen carefully when the speech stimuli are presented and to repeat verbatim all that they can hear, and to guess the content if they can. Before the actual test session, the participants were presented 2 unprocessed lists of sentences to familiarize them with the task. Following this, 2 sentence lists were randomly selected and presented in each stimulus condition, from a Lenovo laptop (Lenovo ThinkPad X1 Carbon, 3rd Gen with intel core i7). The stimuli were delivered to the participants' ears using HDA200 headphones, calibrated to present the stimuli at 70 dB SPL.Each participant responded to 14 sentence lists, presented across 6 test conditions and during familiarization. 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 and the key words were scored in each sentence. Each correctly repeated key word was given a score of 1 (maximum achievable score was 80 in each condition, from 2 sentence lists) and errors or skipped words were given a score of 0.

Results

The mean and SD of speech perception scores obtained during test conditions using TFS speech (TFS2nb, TFS4nb and TFS8nb)and RENV speech (RENV, RENV2wf, RENV4wf) speechare given in figure 1and figure 2 respectively. The data shows a general trend of reduction in scores with increase in number of frequency bands in TFS speech. Similarly, the scores reduce with simulation of cochlear hearing loss during the perception of RENV speech. The meanscore from every condition is poorer in the ONH group compared to the YNH group.

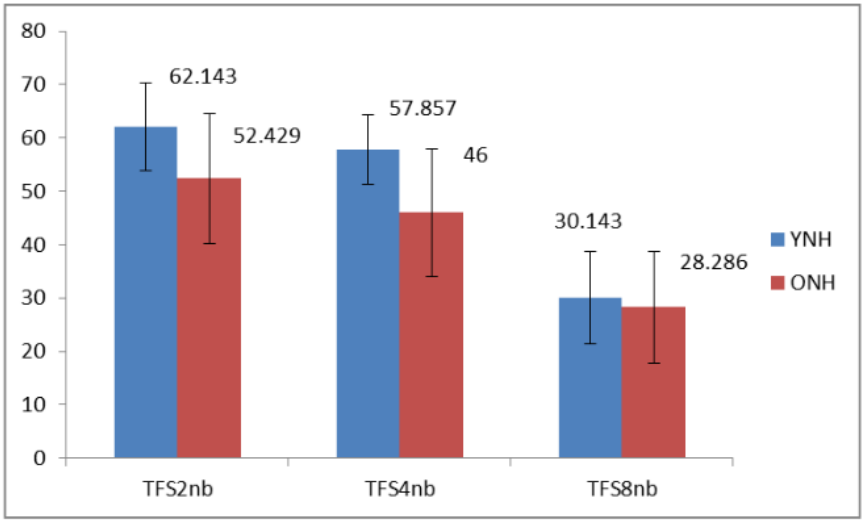

Figure 1

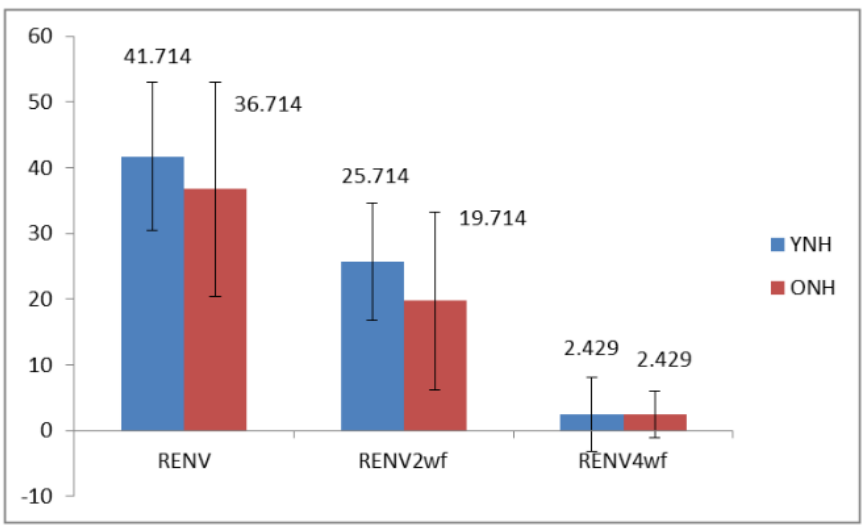

Figure 2

The scores obtained in different test conditions were the dependent variables and age group was the independent variable. The scores from YNH and ONH groups were used for between group comparisons in TFS and RENV speech perception conditions. The data were normally distributed (based on Shapiro-Wilk test of normality)and homogeneity of varianceswas observed onLevene's test of normality ($p>0.05$). Mixed ANOVA was used to compare the results within groups and across the groups in different test conditions. Muchly's test of sphericity indicated that assumption of sphericity was satisfied in both test conditions (TFS speech: $\chi 2$ = 4.618, p = .099; RENV speech: $\chi 2$ = 1.765, p = .414), and therefore, no sphericity corrections were implemented.

There was no interaction between TFS speech conditions and age group ($F(2,1) = 2.031$, $p = 0.153$, $\eta^2 = 194.048$). There was significant main effect ofcondition for TFS speech($F(2) = 65.104$, $p < 0.001$, $\eta^2 = 6219.00$) but there was no main effect of age group in these test conditions ($F(1) = 3.251$, $p < .097$, $\eta^2 = 640.381$). The results of pair-wise comparisonswithin the group for condition and across group comparison for age are given in table 1. Bonferroni adjustments were made for multiple comparisons.The scores were significantly different across conditions in both age groups, whereas significant difference between age groups was seen for only TFS4nb condition.

Table 1: Results of pair-wise comparisons within the age groups for TFS2nb, TFS4nb and TFS8nb conditions and across age groups comparison for age.

TFS speech condition			Mean difference	Std error	Sig. b
TFS2	TFS4		5.357*	1.682	.024
	TFS8		28.071*	2.963	.000
TFS4	TFS2		-5.357*	1.682	.024
	TFS8		22.714*	2.977	.000
TFS8	TFS2		-28.071*	2.963	.000
	TFS4		-22.714	2.977	.000
TFS speech condition	Group				
TFS2	YNH	ONH	9.714	5.560	0.106
	ONH	YNH	-9.714	5.560	0.106
TFS4	YNH	ONH	11.857*	5.124	0.39

	ONH	YNH	-11.857*	5.124	0.39
TFS8	YNH	ONH	1.857	5.137	.724
	ONH	YNH	-1.857	5.137	.724

There was no interaction between RENV speech conditions and age group (F (2,1) = .327, p = .724, η^2 = 52.048). There was significant main effect of condition for RENV speech (F (2) = 58.336, p < 0.001, η^2 = 9279.190) but there was no main effect of age group(F (2,1) = .858, p = .373, η^2 = 164.024). The results of pair-wise comparisons within the group for condition and across group comparison for age are given in table 2. Bonferroni adjustments were made for multiple comparisons. The scores were significantly different across the three RENV speech conditions in both age groups. No significant difference was seen between the two age groups in any of the RENV speech test conditions.

Table 2: Results of pair-wise comparisons within the age groups for RENV, RENV2wf, RENV4wf conditions and across age groups comparison for age.

RENV condition			Mean difference	Std error	Sig. b
RENV	RENV2wf		16.500*	3.053	.000
	RENV4wf		36.357*	3.966	.000
RENV2wf	RENV		-16.500*	3.053	.000
	RENV4wf		19.857*	3.005	.000
RENV4wf	RENV4wf		-36.357*	3.966	.000
	RENV2wf		-19.857*	3.005	.000
RENV condition	Group				
RENV	YNH	ONH	5.000	7.481	.517
	ONH	YNH	5.000	7.481	.517
RENV2wf	YNH	ONH	6.000	6.147	.348
	ONH	YNH	-6.000	6.147	.348
RENV4wf	YNH	ONH	.857	2.513	.739
	ONH	YNH	-.857	2.513	.739

Discussion

The study aimed to compare the perception of TFS speech and RENV speech in young and older normal hearing individuals, using sentence stimuli. No interaction was observed between the two age groups and the test conditions. The participants' speech perception scores were significantly different between the TFS2nb, TFS4nb and TFS8nb conditions, and between RENV, RENV2wf and RENV4wf conditions. The increase in the number of bands in TFS speech progressively degrades the speech signal(R Drullman, 1995), resulting in reduction of speech perception scoresin the participants in these conditions. These findings are in agreement with the literature(Drullman et al., 1994a, b; R Drullman, 1995; Smith et al., 2002).

Introduction of widening factor to RENV speech essentially simulates widening of the auditory filters- this too results in degraded speech signal. The results of the study indicate that the perception of TFS in the presence of cochlear hearing loss degrades significantly with the increase in severity of the cochlear pathology or the alteration in cochlear physiology. However, the scores obtained in the two groups for perception of these stimuli were not significantly different.

The present study did not find significant difference in the speech perception scores in all the TFS and RENV speech perception conditions. Perception of TFS information is reported to be significantly impaired in older normal hearing individuals with normal pitch perception(Hopkins & Moore, 2011; Peters & Moore, 1992; Strelcyk & Dau, 2009). Differences in the findings of the present study and the literature could be due to the age differences between the participants in the older age group in the literature, and the present study; participants of the previous studies were generallyolder than the participants in the present study. Another possible contributing factor for the difference could be the number of participants in the present study. A clearer picture of the trend may be seen when the tests are administered in more participants. The study did find significant difference between the two groups at the TFS4nb condition and difference was obtained between the two groups in the other two TFS conditions. This could be because of ceiling effect in the scores in the TFS2nb condition and floor effect in the TFS8nb condition, that a significant difference is observed only in condition with optimal difficulty.

Conclusions

The perception of 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. This observation needs to be repeated, possibly in a more diverse population to further our understanding of TFS and RENV speech perception.

Figure Legends

Figure 1: Mean and SD of speech perception scores in TFS2nb, TFS4nb and TFS8nb conditions, in YNH and ONH groups.

Figure 2: Mean and SD of speech perception scores in RENV, RENV2wf and RENV4wf conditions, in YNH and ONH groups.

Tables

Table 1: Results of pair-wise comparisons within the age groups for TFS2nb, TFS4nb and TFS8nb conditions and across age groups comparison for age.

TFS speech condition			Mean difference	Std error	Sig. b
TFS2	TFS4		5.357*	1.682	.024
	TFS8		28.071*	2.963	.000
TFS4	TFS2		-5.357*	1.682	.024
	TFS8		22.714*	2.977	.000
TFS8	TFS2		-28.071*	2.963	.000
	TFS4		-22.714	2.977	.000
TFS speech condition	Group				
TFS2	YNH	ONH	9.714	5.560	0.106
	ONH	YNH	-9.714	5.560	0.106
TFS4	YNH	ONH	11.857*	5.124	0.39
	ONH	YNH	-11.857*	5.124	0.39
TFS8	YNH	ONH	1.857	5.137	.724
	ONH	YNH	-1.857	5.137	.724

Table 2: Results of pair-wise comparisons within the age groups for RENV, RENV2wf, RENV4wf conditions and across age groups comparison for age.

RENV condition			Mean difference	Std error	Sig. b
RENV	RENV2wf		16.500*	3.053	.000
	RENV4wf		36.357*	3.966	.000
RENV2wf	RENV		-16.500*	3.053	.000
	RENV4wf		19.857*	3.005	.000
RENV4wf	RENV4wf		-36.357*	3.966	.000
	RENV2wf		-19.857*	3.005	.000
RENV condition	Group				
RENV	YNH	ONH	5.000	7.481	.517
	ONH	YNH	5.000	7.481	.517
RENV2wf	YNH	ONH	6.000	6.147	.348
	ONH	YNH	-6.000	6.147	.348
RENV4wf	YNH	ONH	.857	2.513	.739
	ONH	YNH	-.857	2.513	.739

Article17

ORIGINALITY REPORT

6%	0%	3%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Bei Li, Limin Hou, Li Xu, Hui Wang, Guang Yang, Shankai Yin, Yanmei Feng. "Effects of steep high-frequency hearing loss on speech recognition using temporal fine structure in low-frequency region", Hearing Research, 2015 Publication	1%
2	Submitted to University of Western Sydney Student Paper	1%
3	Submitted to Prairie View A&M University Student Paper	1%
4	Submitted to All India Institute of Speech & Hearing Student Paper	1%
5	Laurianne Cabrera, Christian Lorenzi, Josiane Bertoncini. "Infants Discriminate Voicing and Place of Articulation With Reduced Spectral and Temporal Modulation Cues", Journal of Speech, Language, and Hearing Research, 2015 Publication	1%
	Submitted to Georgia Institute of Technology	

6	Main Campus Student Paper	<1%
7	Submitted to University of Liverpool Student Paper	<1%
8	Li, Bei, Limin Hou, Li Xu, Hui Wang, Guang Yang, Shankai Yin, and Yanmei Feng. "Effects of steep high-frequency hearing loss on speech recognition using temporal fine structure in low-frequency region", Hearing Research, 2015. Publication	<1%
9	www.carbon-it.co.uk Internet Source	<1%
10	Moore, . "The Influence of Hearing Loss and Age on the Binaural Processing of TFS", Auditory Processing of Temporal Fine Structure, 2014. Publication	<1%
11	Submitted to Macquarie University Student Paper	<1%
12	Brian J. Malone, Brian H. Scott, Malcolm N. Semple. "Dynamic Amplitude Coding in the Auditory Cortex of Awake Rhesus Macaques", Journal of Neurophysiology, 2007 Publication	<1%
13	Hemanth Narayan Shetty, Jijo Mathai Pottackal. "Gain adjustment at tinnitus pitch to manage	<1%

both tinnitus and speech perception in noise",
Journal of Otology, 2019
Publication

14	Submitted to La Trobe University Student Paper	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		