



Review

Temporal cues and the effect of their enhancement on speech perception in older adults – A scoping review

Hemanth Narayan Shetty

All India Institute of Speech and Hearing, Mysuru, India

Received 17 July 2016; revised 18 August 2016; accepted 22 August 2016

Abstract

Temporal envelope is a low frequency amplitude modulation conveying segmental and suprasegmental information during speech perception. Unfortunately, we seldom find ourselves in completely quiet listening environments and noise, commonly found in the surrounding environment, obscures both the fine structure cues and partly the temporal envelope cues in speech. Available temporal content of speech emanating from noise is often enough to convey required information in normal hearing individuals. However, the case is different in older adults (with and without hearing loss) who lack such capabilities due to the impairment in temporal processing. This calls attention of a researcher to delineate the importance of temporal enhancement of speech in improving speech perception. There are many temporal envelope strategies available in the literature, but each one has its own lacunae. An envelope enhancement by a deep band modulation (DBM) is found to be beneficial for those individuals who have had a temporal processing impairment. The reason could be attributed to the 15 dB enhancement in the temporal envelope bandwidth between 3 and 30 Hz, extracted from each channel, which significantly increases the modulation depth such that masking of a consonant by a vowel is minimized. Additionally, output of deep band modulated speech is rescaled such that its duration increases and it provides relatively easy access to the word of the lexicon. Thus, in the near future, with more experiments related to DBM algorithm, it can be utilized in the rehabilitative devices to lessen the impact of the temporal processing impairment.

Copyright © 2016, PLA General Hospital Department of Otolaryngology Head and Neck Surgery. Production and hosting by Elsevier (Singapore) Pte Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Temporal; Envelope; Modulation; Older adults

Contents

1. Introduction	96
2. Importance of temporal envelope cues	96
2.1. Modulation depth and bandwidth	96
2.2. Modulation detection	96
2.3. Masking release	97
3. Speech perception in noise	97
3.1. Older adults with normal hearing	97
3.2. Speech perception in older adults with hearing loss	98
4. Temporal enhancement strategies	98
5. Conflicts of interest	100
Acknowledgement	100
References	100

E-mail address: hemanthn.shetty@gmail.com.

Peer review under responsibility of PLA General Hospital Department of Otolaryngology Head and Neck Surgery.

<http://dx.doi.org/10.1016/j.joto.2016.08.001>

1672-2930/Copyright © 2016, PLA General Hospital Department of Otolaryngology Head and Neck Surgery. Production and hosting by Elsevier (Singapore) Pte Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The deterioration in hearing ability occurring with advanced age is known as presbycusis (Mills et al., 2006). The severity of hearing loss in them varies from mild to severe with sloping configuration (Mills et al., 2006). If it is left untreated, it may have a significant effect on their communication skills. Individuals with normal hearing understand speech in the presence of noise through temporal cues such as listening in dips (Stuart and Phillips, 1996), modulation detection (Grose et al., 2009) and release from masking (Hopkins and Moore, 2011). However, older adults with hearing loss are unable to process temporal cues due to asynchronous neural firing at higher auditory centres (Pichora-Fuller and Cheesman, 1997). Due to this reason, an older adult with hearing loss finds it difficult to follow speech in adverse listening conditions. Thus, in older adults with hearing loss, merely alleviating the audibility factor by an amplification device may not solve the problem. Enhancement of temporal cues has shown improvement in speech perception in the presence of noise, nevertheless each strategy has its own critics. From this viewpoint, the review has been focused on the following objectives: a) importance of temporal envelope cues in speech perception b) speech perception in noise in older adults with and without hearing loss and c) effect of temporal enhancement strategies on speech perception.

Temporal structure cues of speech are classified into three categories based on frequency (Rosen, 1992). They are envelope cue, periodicity cue and temporal fine structure cue. Envelope cue contains a frequency range from 2 to 50 Hz, which transmits voicing and stress information. Periodicity cue (50–500 Hz) conveys information on voicing, manner and intonation. Whereas, temporal fine structure cue (600–10 KHz) passes information of consonant place and vowel quality. In order to recognise speech, these cues should be processed by the auditory system. However, in naturalistic situations, these envelopes are concealed from existing noise in the environment.

2. Importance of temporal envelope cues

2.1. Modulation depth and bandwidth

Temporal envelope is a slow modulation signal, which occurs between 5 and 50 Hz and conveys segmental and suprasegmental information. Unfortunately, noise tends to obscure slow modulation of speech by filling dips across the waveform. Nevertheless, overall amplitude of competing signals varies due to which some amount of available dips in desired signal enable a listener to hear out segments of the target signal. The interfering effects of competing signals on understanding the desired speech depends on factors such as the number of competing signals (Duquesnoy and Plomp, 1980), spectrum of competing signal and desired speech (Sommer and Gehr, 1998), informational masking (Ezzatian et al., 2011) and correlated and uncorrelated masking noise with respect to desired signal (Veloso et al., 1990). If a subject

is able to comprehend a spoken message in the noise, then it may partly indicate that his/her temporal processing ability has played an instrumental role in inferring the information. To support this hypothesis, Turner et al. (1994) utilized unprocessed and processed nonsense speech syllables to assess the importance of temporal content of speech on identification of syllables. The target test signals were subjected to many steps of mathematical operations to obtain the processed signals and they were: a) broadband noise modulated by an envelope of broadband speech signal b) low pass noise modulated by a low pass speech signal c) high pass noise modulated by a high pass speech signal and d) combined two channel signal which comprised of low and high modulated signals. In order to make the study participants (both normal hearing individuals and individuals with hearing impairment) rely only on temporal cues, nonsense syllables were subjected to certain processing strategies. These stimuli were presented at a comfortable level in quiet; and in modulated and steady state background noise conditions. For both unprocessed and processed conditions, individuals with hearing impairment performed poorer than normal hearing listeners, in noisy conditions. This inferred that hearing impaired listeners were unable to utilize temporal dips in the background noise. The results are inconsistent with other research reports, which believed that elderly listeners relied on temporal dips in modulated noise than steady state noise for recognition of speech (Stuart and Phillips, 1996; Shannon et al., 1995). Interestingly, it is well established that recognition of speech improves with increase in the temporal envelope bandwidth. On this line of research, Shannon et al. (1995) conducted a study where spectral information was greatly reduced and envelope bandwidth was emphasized to make use of only temporal cues. In their method, temporal envelope of speech was extracted using Hilbert transform from a set of filters. The extracted envelope was used to modulate noise of the same bandwidth. The recognition of consonants, vowels, and words in simple sentences improved distinctly as the bandwidth of the bands enlarged to the restricted range. Thus, dynamic temporal cues restricted to a few broad frequency regions are sufficient for the identification of speech.

2.2. Modulation detection

It is reported that older adults find it difficult to follow speech at higher rates of masker modulation. Grose et al. (2009) conducted a study by using low and high predictive sentences embedded in modulated noise of two modulation rates (16 Hz and 32 Hz). It was observed that recognition scores were lesser for low predictive sentences than high predictive sentences at both modulation rates in younger and older adult groups. Also, older listeners did not show an exacerbated deficit in the recognition of high predictive sentences at lower rate of modulation noise compared to higher rate of modulation. This infers that aging affects recognition ability if the noise modulation rate is high. In yet another study, Fullbrage et al. (2003) investigated the effect of cochlear damage on the recognition of complex temporal

envelopes using first- and second-order amplitude modulation (AM) detection thresholds. They used 2 kHz pure tone as the carrier. In the first-order, thresholds were calculated for amplitude modulation rates ranging from 4 to 87 Hz, whereas, in the second-order, thresholds were calculated for amplitude modulation rates ranging from 4 to 23 Hz. The results revealed that second-order amplitude modulation detection thresholds in hearing impaired listeners were similar to that of normal hearing listeners at all modulation rates. However, it was not true for the first-order amplitude modulated thresholds at the rate of 87 Hz. From this finding, the authors concluded that older adults with hearing loss failed to recognize speech, especially in increased rate of temporal envelope modulation.

2.3. Masking release

Generally, masking release is defined as the ease with which speech is heard over the noise. If noise is modulated and its bandwidth is larger than speech, then it is easier for a listener to extract speech over noise, through listening in the dips and/or masking release. Moore and Glasberg (1987) explained masking release using phase locking capability of the auditory neurons. Phase locking pattern of neurons is robustly tuned to the signal frequency in a condition where masker envelope is at minima. Conversely, in unmodulated noise or when the envelope of the noise is at its maximum, a listener finds it hard to follow speech over noise. Masking release depends on the degree of audibility, in particular to the masker and the speech frequencies. Further, wider bandwidths of modulated noise, greater than 100 Hz, improved the signal detection. The fluctuations in modulated noise are critical and these are compared across the output of different auditory filters to detect a signal (Hall and Grose, 1991). Carlyon et al. (1989) demonstrated that signal detection was more likely in a condition where a modulated masker was at a low rate, and if it covered a wide frequency range. The probable reason could be that the modulation patterns at the output of the auditory filters, tuned to the frequencies of the target signal, are distinct from the modulation patterns of the noise. This disparity in the modulation patterns, across the different auditory filters, is sensitive to detect the signal leading to masking release (Moore and Shailer, 1991). In addition, signal amplitude depth and its amplitude above the masker are also important factors in masking release. Gustafsson and Arlinger (1994) conducted a study on the recognition of speech embedded in amplitude-modulated and unmodulated speech-spectrum shaped noise in younger and older subjects with and without hearing loss. In the amplitude-modulated noise, the modulation depth of the sinusoid varied with the frequency, ranging from 2 to 100 Hz. In addition, uneven modulation was generated by the addition of four sinusoids in random phase relation. These stimuli were presented at ± 6 dB, and ± 12 dB. The results revealed that for the normal-hearing subjects, different kinds of modulated noises facilitated some amount of masking release for speech compared to the unmodulated noise, which had a little or no masking release. It can be inferred that, for a normal-hearing individual, sinusoidal modulation provided more release of masking than the

uneven modulation. It was also noted that the release of masking over speech increased with modulation depth. Whereas, older adults with hearing loss needed relatively high modulation depth to release masking over speech due to the impaired temporal resolution. In a similar line of experiment, Bacon et al. (1998) investigated masking release in temporally complex backgrounds on three groups of participants: normal hearing group (NH), individuals with sensorineural hearing loss (HL) and normal hearing (NM) subjects with pure tone thresholds elevated to match the audibility of the SNHL group. Performance was carried out in four backgrounds where temporal envelopes were varied in a) steady-state (SS) speech-shaped noise, b) speech-shaped noise modulated by the envelope of multi-talker babble (MT), c) speech-shaped noise modulated by the envelope of single talker speech (ST) and d) speech-shaped noise modulated by a 10-Hz square wave (SQ). Results revealed that signal to noise ratio (SNR) required to just follow the speech in ST condition was less, demonstrating a higher release of masking in single talker modulated backgrounds than other types of noises. SNR thresholds were similar in steady state and multi-talker modulated noise conditions. The release of masking was larger in the normal hearing group than other groups in all the four backgrounds, where temporal envelopes were varied. The HL group consisted of 11 listeners, out of which 5 listeners exhibited similar masking release as that of the NM group. Whereas, the rest of the 6 in the HI group showed a trend of lesser masking release than the NM group. This indicates that the reduced release of masking in SNHL may be due to the temporal processing impairment. To conclude, it is supported from literature that speech is recognised in the presence of noise through listening in dips, modulation detection in speech and release from masking.

3. Speech perception in noise

3.1. Older adults with normal hearing

Older adults find it difficult to understand speech in unfavourable conditions such as noise, reverberation and rapid rate of speech. In literature, there are numerous studies on speech recognition assessed with different speech materials and in background noise. Gordon-Salant and Fitzgibbons (1993) investigated the recognition of low-predictive sentences, in which the temporal waveform was distorted by compression and reverberation. The undistorted and distorted sentences were presented to younger and older adults having normal hearing at different signal to noise ratios. It was found that older adults failed to identify sentences in reduced signal to noise ratios, but the scores deteriorated even more when the temporal envelope of the sentence was degraded by two or three combinations of distortion. In yet another study by Souza and Turner (1994), who investigated the effect of age on favourable SNRs at which maximum performance was obtained, it was observed that older adult listeners performed similar to that of younger listeners at +8 dB SNR. It infers that older adults required the minimum favourable SNR of +8 dB to recognize speech similar to that of younger adults. These

results suggest that age influences the recognition of speech in degraded conditions.

In order to know the plausible cues that are utilized to recognize speech in degraded conditions, [Souza and Turner \(1994\)](#) investigated speech recognition in older adults using monosyllables as the target speech stimuli. These stimuli were embedded at different levels of speech spectrum background noise. Further, speech-spectrum noise was temporally modulated by the envelope of a multi-talker babble. Each background noise brought about a significant reduction in the speech recognition scores in older adults compared to younger adults. This was exacerbated in unmodulated noise condition, in which the temporal and spectral variations of masking noise were closer to that of the target speech, where listening through available dips was less likely. This result indicates that older adults make less use of spectral and temporal dips to recognize speech embedded in modulated noise. To know the importance of spectral and temporal processing ability in understanding speech in background noise, [Peters et al. \(1996\)](#) studied speech reception thresholds (SRT) in noise with and without spectral and temporal dips. They included younger and older individuals with normal hearing as subjects. They generated three types of noises. Each noise was modulated with respect to the envelope of the speech, steady state noise and single talker noise. The participants were asked to repeat the speech against each background noise presented at 65 dB SPL to obtain SRTn. In younger adults, the mean SRT for the speech embedded in speech modulated noise was 6.2 dB lower than that embedded in steady noise, but 1.9 dB higher than that of single talker. This indicates that the presence of temporal dips in single talker was a major cue for recognition of speech. Further, in older adults, the same trend of younger adults was observed in different noise conditions, however, their scores were significantly reduced than those of younger adults. Thus, it can be inferred that older adults appear to take slightly less advantage of the ‘dip listening’ available at different SNRs.

3.2. Speech perception in older adults with hearing loss

Individuals with cochlear hearing impairment often complain of not understanding speech, especially in background noise ([Plomp, 1994](#)). Frequency selectivity is usually impaired in individuals with cochlear hearing loss ([Hopkins and Moore, 2011](#)). Whereas, the temporal resolution is near normal or impaired depending on the degree of the hearing loss ([Schneider et al., 1994](#)). [Festen and Plomp \(1990\)](#) investigated the recognition of speech at different SNRs in individuals with cochlear hearing loss. It was observed that individuals with cochlear hearing loss required higher SNR levels to achieve identical performance as normal hearing individuals. In addition, difference in SRT varied greatly depending on the nature of the background noise for both normal and hearing-impaired groups. When the background noise used was speech-shaped noise, SRTn difference between normal and hearing-impaired individuals ranged from 2 to 5 dB. Whereas, in other background noises, such as single competing talker, time-reversed talker or an amplitude-

modulated noise, the difference in SRTn was much larger, ranging from about 7 dB up to about 15 dB ([Clarkson and Bahgat, 1991](#)). Thus, speech recognition in noise in individuals with cochlear hearing loss varies based on the type of background noise, which masks the temporal and spectral contents of the speech. Further, in case of informational masking such as single talker and four talker babble, hearing-impaired listeners fail to take benefits of ‘dips’ in competing voice. These dips may be of two types: temporal and spectral. Temporal dips are momentary fluctuations in overall signal to noise ratio, especially during brief pauses in speech or during production of low energy sounds. In these temporal dips, the signal strength is relatively higher than background noise, which allows brief ‘glimpses’ to be gained from the target speech. The spectral dips occur when the target speech spectrum differs from the background spectrum over short intervals. Although some parts of the target spectrum may be entirely masked by the background noise, there might be other parts which may be scarcely masked if at all. Thus, these parts of the spectrum of the target speech which might be barely masked may be “glimpsed” and used as cue to follow speech in the competing noise. From the literature, it is clear that speech is affected by external factors such as noise, reverberation and rate of speech, which obscure the inherent temporal and spectral cues. This altered speech further loads the auditory system of the hearing impaired individuals. In such scenario, the impaired system fails to access the barely available cues in the adversely altered speech. The possible factor in the reduction of speech recognition in noise in cochlear hearing loss subjects could be broadened auditory filters. These wider auditory filters do not mean that it removes information from speech, rather it impedes the transfer of spectral and temporal information. It can be expected that spectral peaks and valleys in the stimulus are smoothed out in these individuals. In addition, the upward spread of masking is common i.e., the high frequency components of the speech (consonants) being masked by the higher amplitude of vocalic sounds or low frequency sounds (vowel), which is found to be one of the confronting factors in sensorineural hearing loss (SNHL). It is also speculated that only a few auditory filters are available for analysis, but noise accompanied with stimulus, taxes these available filters and accumulates in the functioning filters leading to reduced recognition in lesser SNRs. To summarize, hearing-impaired individuals gained much less advantage from spectral and temporal dips to recognize speech in background noise. If the spectral and temporal content of the noise is closer to the target speech stimulus, then its effect on speech recognition is exacerbated.

4. Temporal enhancement strategies

The speech signal has prominent low-frequency amplitude modulation, which is termed as envelope of the speech signal. It conveys important cues for consonant recognition. In 1990's, a considerable amount of research on enhancement of temporal envelope of speech was conducted. This research was taken up to solve the confronting problem of difficulty in speech

understanding in adverse listening conditions. It was observed that individuals with normal hearing, cochlear hearing loss, and learning impairment showed improved speech in noise perception when the envelope cues of the speech signal were enhanced (Freyman and Nerbonne, 1996; Lorenzi et al., 1999). Envelope enhancement is an augmentation of the low frequency modulation of slow or fast temporal content of speech. In one of the experiments, Langhans and Strube (1982) enhanced the speech by non-linear multiband envelope expansion method. The test stimuli used were sentences and each sentence was sampled at 10 kHz. Fast Fourier Transform was applied to derive spectrum. A new spectrum was constructed every 10 ms and output in each band was sent through a modulation band pass filter. In addition, a gain was assigned to each band in taking ratio of a filtered envelope to an unfiltered envelope. The output of each band was recombined with the original phase to produce the enhanced signal. It was noted that sentence recognition test did not show significant improvements by non-linear envelope expansion method. The reason would be that both the vowel and consonant were assigned the same gain which more likely leads to upward spread of masking (i.e. vowels masking the weak level consonants). In a similar experiment, Kusumoto et al. (2000) enhanced the modulation depth of the signal between 2 Hz and 8 Hz. This was considered as there was a good correlation between the temporal modulation transfer function and speech intelligibility. In their study, four sensorineural hearing impaired individuals were included and asked to listen to the processed signal and original signal. A brief explanation on the signal processing needs to be stated before concluding the findings. The original signal was divided into 16 frequency bands by band pass filters with 1/3 rd octave bandwidth. From the output of each band, the envelope was extracted using a Hilbert transformer. The envelope from each band was down sampled by a factor of M and applied to several modulation filters. To convert the filtered envelope to the original signal, up-sampling was carried out with a same factor of M. To remove any artefact introduced by modulation filters, half wave rectification was applied to the filtered envelope. Finally, modulated signal from each filter was multiplied with an original band pass filter of the same filter to obtain envelope enhancement between 2 and 8 Hz. It was observed that hearing impaired listeners appreciated the perception of processed signal than the original signal. However, in the process of signal extraction, modulation filters among frequency bands were kept same. It is preferred to give the option to choose different modulation filters between frequency bands to make use of temporal content of speech, especially for those hearing impaired individuals whose dynamic range varies across frequencies. In yet another experiment, Clarkson and Bahgat (1991) used envelope expansion schema. In this schema, the target signal was filtered into several contiguous frequency bands and in each band the envelope was magnified using the non-linear expansion. These target stimuli were presented against white noise at various levels of SNR (0, -5, and -15 dB). Results showed small improvement (6%) for the envelope enhanced stimuli at 0 dB SNR, but no improvement at -5 and -15 dB SNRs. In order to improve the envelope

enhanced speech perception at lesser signal to noise ratio, Freyman and Nerbonne (1996) used simple power law function. This was used to enhance the envelope of vowel-consonant-vowel (VCV) stimulus. They used consonants as speech stimuli, which were presented at quiet and at different SNRs to normal hearing individuals. Unfortunately, the performance on this schema showed deteriorated response in quiet and at different SNRs due to reduced consonant to vowel ratio. In another approach, Lorenzi et al. (1999) applied envelope expansion nonlinearity schema to the temporal envelope of speech. Vowel-consonant-vowel syllables were presented to subjects in quiet and in the background of steady state noise at 0 dB SNR. The study comprised of four normal hearing subjects. In envelope expansion, temporal modulation of frequencies less than 500 Hz was extracted from each syllable and raised to the power of 2. The resulting envelopes were then used to modulate the white noise. The resultant output was speech enveloped noise stimuli. Their participants were instructed to recognize syllables. The results showed a deleterious effect on the identification of syllables when the envelope was expanded. There was a small improvement of 6–14%, which was consistent in performance when expanding the envelope of speech syllables with the noise. Apoux et al. (2000) extended the study by using an envelope expansion technique on perception of speech in hearing impaired subjects. It was found that improvement in recognition of speech was less. Overall, the results on temporal expansion showed no improvement in hearing-impaired listeners. The discrepancy noted from two previous studies of Freyman and Nerbonne (1996) and Lorenzi et al. (1999) could be due to the type of stimuli used and the cues on which their study participants relied for perception. In their studies, power square law technique was used to enhance the temporal envelope, which increased the vowel amplitude than the consonant amplitude leading to upward spread of masking.

In further efforts to eliminate the upward spread of masking, Apoux et al. (2004) investigated the effect of temporal envelope expansion on sentence identification by taking 8 normal hearing and 24 elderly cochlear hearing loss participants. In this study, envelope squaring and expansion-compression schema were used. The first method is explained earlier, whereas in method two, the depth was varied artificially based on either high or low amplitude fluctuations. Identification task was carried out against stationary and fluctuating noise, in which these noises were applied before and after their envelopes were processed by speech. In the first expansion scheme, it was observed that there were no significant improvements in identification scores in both normal hearing and hearing impaired listeners. This is because when envelope expansion was applied within a range of 0–16 Hz, the high amplitude vocalic sounds in sentences were amplified whereas low amplitude consonants were reduced, leading to reduced consonant vowel (CV) ratio. This had a negative impact on the perception of sentences. However, in expansion compression scheme, higher amplitude (vowels) was compressed and lower amplitude (consonants) was amplified. The results yielded a significant improvement when applied to speech before the addition of background noise in both normal and hearing impaired listeners.

Similar to expansion and compression scheme, Nagarajan et al. (1998) developed another temporal envelope enhancement strategy called Deep Band Modulation (DBM). This algorithm basically enhances the temporal modulation, thereby lessening the deleterious effects of noise. A DBM enhances the modulation depth of sound and increases the time scale of entire duration. In DBM, the extracted temporal envelope from each channel with bandwidth ranging from 3 to 30 Hz is enhanced by 15 dB, which significantly increases the modulation depth such that masking of a consonant by a vowel is minimized. In addition, deep band modulated output rescales the entire length of the stimulus. It was found that DBM improved speech perception score in children with learning disability who typically demonstrated temporal processing deficits. It is well established that even older adults without hearing loss have temporal resolution impairment. In this perspective, Hemanth and Akshay (2015) conducted a study with the hypothesis that DBM scheme improves phrase perception in older adults, in noisy conditions. It was found that the hypothesis was proved true and observed that the speech perception scores improved in noise using the algorithm DBM. From Nagarajan et al. (1998) and Hemanth and Akshay (2015) studies, it is understood that speech perception improved in their study participants who had temporal processing impairment. Further, DBM was used to study the speech perception in older adults having hearing loss (Sneha and Hemanth, 2015). It was found that, older adults with hearing loss performed significantly better in DBM condition than unprocessed (UP) condition at higher signal to noise ratios, such that the participants could make use of the higher amplitude of modulation depth. Another observation made from the study was that perception of phrase in DBM condition at higher signal to noise ratio from older adults with hearing loss approximated the unprocessed phrase perception scores obtained by Younger Adult Group (YAG). This study sheds a light on the fact that after correcting for the audibility, the increased modulation depth (15 dB) in speech brought about by DBM improves phrase perception in older adults with hearing loss even in noise (only higher signal to noise ratios). It was also speculated that rescaling the entire stimulus length by DBM might have got additional time to access the appropriate lexicon. Recognition of deep band modulated consonants was also studied in older adults with and without hearing loss. It was observed that at reduced SNRs, the cues from DBM facilitated the listeners to repeat the heard VCV syllables. The effect of aging and combined effect of aging and hearing loss was partly lessened by DBM by enhancing the manner feature in VCV syllables. In noisy condition, DBM helped them in listening through temporal dips, where the amplitude of the envelope was enhanced. The use of DBM strategy has potential to improve speech perception in individuals with SNHL at adverse listening conditions. Hence, there is a scope for this strategy to be utilized as a rehabilitation technique.

To conclude, a temporal enhancement strategy such as the DBM has been proven to improve the perception of phrases in older adults with and without hearing loss. DBM helps them to

access the available temporal cues by lessening the temporal asynchrony present in the older adults. In addition, upward spread of masking is also reduced since the gain provided for the consonants is more and lesser for the vowels. Further, in noisy condition, this technique has helped the older adults in listening through temporal dips, where the amplitude of envelope would be enhanced. Further, rescaling of the entire length of phrase provided additional timing to locate the words in the lexicon.

Conflicts of interest

None.

Acknowledgement

The author would like to acknowledge the Director and HOD (Audiology), All India Institute of Speech and Hearing, Mysore, India.

References

- Apoux, F., Crouzet, O., Lorenzi, C., 2000. Temporal envelope expansion of speech in noise for normal-hearing and hearing-impaired listeners. Effects on identification performance and response times. *Hear. Res.* 153, 123–131.
- Apoux, F., Tribut, N., Debrulle, X., Lorenzi, C., 2004. Identification of envelope expanded sentences in normal-hearing and hearing-impaired listeners. *Hear. Res.* 189, 13–24.
- Bacon, S.P., Opie, J.M., Montoya, D.Y., 1998. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. *J. Speech Lang. Hear. Res.* 41, 549–563.
- Carlyon, R.P., Buus, S., Florentine, M., 1989. Comodulation masking release for three types of modulator as a function of modulation rate. *Hear. Res.* 42, 37–46.
- Clarkson, P.M., Bahgat, S.F., 1991. Envelope expansion methods for speech enhancement. *J. Acoust. Soc. Am.* 89 (3), 1378–1382.
- Duquesnoy, A.J., Plomp, R., 1980. Effect of reverberation and noise on the intelligibility of sentences in cases of presbycusis. *J. Acoust. Soc. Am.* 68, 537–544.
- Ezzatian, P., Li, L., Pichora-Fuller, K., Schneider, B., 2011. The effect of priming on release from informational masking is equivalent for younger and older adults. *Ear Hear.* 32, 84–96.
- Festen, J.M., Plomp, R., 1990. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88, 1725–1736.
- Freyman, R.L., Nerbonne, G.P., 1996. Consonant confusions in amplitude-expanded speech. *J. Speech Hear. Res.* 39, 1124–1137.
- Fullbrage, C., Meyer, B., Lorenzi, C., 2003. Effect of cochlear damage on the detection of complex temporal envelopes. *Hear. Res.* 178, 35–43.
- Gordon-Salant, S., Fitzgibbons, P.J., 1993. Temporal factors and speech recognition performance in young and elderly listeners. *J. Speech Hear. Res.* 36, 1276–1285.
- Grose, J.H., Mamo, S.K., Hall, J.W., 2009. Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 30, 568–575.
- Gustafsson, H.A., Arlinger, S.D., 1994. Masking of speech by amplitude-modulated noise. *J. Acoust. Soc. Am.* 95, 518–529.
- Hall, J.W., Grose, J.H., 1991. Some effects of auditory grouping factors on modulation detection interference (MDI). *J. Acoust. Soc. Am.* 90, 3028–3036.
- Hemanth, N., Akshay, M., 2015. Deep band modulation and noise effects: perception of phrases in adult. *Hear. Balance Commun.* 13, 111–117.

- Hopkins, K., Moore, B.C., 2011. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J. Acoust. Soc. Am.* 130, 334–349.
- Kusumoto, A., Arai, T., Kitamura, T., Takahashi, M., Murahara, Y. 2000. Modulation enhancement of speech as a preprocessing for reverberant chambers with the hearing-impaired. *International Conference on Acoustics, Speech and Signal processing IEEE*. 2, 853–856.
- Langhans, T., Strube, H.W. 1982. Speech enhancement by non linear multi-band envelope expansion. *International Conference on Acoustics, Speech and Signal processing IEEE*. 89, 156–159.
- Lorenzi, C., Berthommier, F., Apoux, F., Bacri, N., 1999. Effects of envelope Expansion on speech recognition. *Hear. Res.* 136, 131–138.
- Mills, J.H., Schmidt, R.A., Dubno, J.R., 2006. Age-related hearing loss: a loss of voltage not hair cells. *Semin. Hear* 27, 228–236.
- Moore, B.C., Glasberg, B.R., 1987. Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes. *J. Acoust. Soc. Am.* 82, 69–79.
- Moore, B.C., Shailer, M.J., 1991. Comodulation masking release as a function of level. *J. Acoust. Soc. Am.* 90, 829–835.
- Nagarajan, S.S., Wang, X., Merzenich, M.M., Schreiner, C.E., Johnston, P., Jenkins, W.M., Miller, S., Tallal, P., 1998. Speech modifications algorithms used for training language learning-impaired children. *IEEE Trans. Rehabilitation Eng.* 6, 257–268.
- Peters, R.W., Moore, B.C., Baer, T., 1996. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* 103 (1), 577–587.
- Pichora-Fuller, K., Cheesman, M., 1997. Preface to the special issue on hearing and aging. *J. Speech Lang. Hear R.* 21, 75–79.
- Plomp, R., 1994. Comments on “Evaluating a speech reception threshold model for hearing-impaired listeners. *J. Acoust. Soc. Am.* 96, 586–589.
- Rosen, S., 1992. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical transactions of the royal society London series B. Biol. Sci.* 336, 367–373.
- Schneider, B.A., Pichora-Fuller, M.K., Kowalchuk, D., Lamb, M., 1994. Gap detection and the precedence effect in young and old adults. *J. Acoust. Soc. Am.* 95, 980–991.
- Shannon, R., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal envelope cues. *Science* 270, 303–304.
- Sneha, S., Hemanth, N., 2015. Effect of Deep Band Modulation and Noise on Phrase Perception in Older Adults with and without Hearing Loss (Unpublished Dissertation: Submitted to University of Mysore).
- Sommers, M.S., Gehr, S.E., 1998. Auditory suppression and frequency selectivity in older and younger adults. *J. Acoust. Soc. Am.* 103, 1067–1074.
- Souza, P.E., Turner, C.W., 1994. Masking of speech in young and elderly listeners with hearing loss. *J. Speech Hear Res.* 37, 655–661.
- Stuart, A., Phillips, D.P., 1996. Word recognition in continuous and interrupted broadband noise by young normal hearing, older normal hearing, and presbycusis listeners. *Ear Hear* 17, 478–489.
- Turner, C.W., Souza, P.E., Forget, L.N., 1994. Use of temporal envelope cues in speech recognition by normal and hearing impaired listeners. *J. Acoust. Soc. Am.* 97, 2568–2576.
- Veloso, K., Hall, J.W., Grose, J.H., 1990. Frequency selectivity and comodulation masking release in adults and in 6-year-old children. *J. Speech Hear Res.* 33, 96–102.