+ MODEL
ARTICLE IN PRESS
+ MODEL

Available online at www.sciencedirect.com

ScienceDirect

Journal of Otology xx (2016) 1—7

www.journals.elsevier.com/journal-of-otology/

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.

Keywords: Temporal; Envelope; Modulation; Older adults

Contents

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/

Please cite this article in press as: Shetty, H.N., Temporal cues and the effect of their enhancement on speech perception in older adults — A scoping review, Journal of Otology (2016), http://dx.doi.org/10.1016/j.joto.2016.08.001

+ MODEL

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 (Sommers 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

Please cite this article in press as: Shetty, H.N., Temporal cues and the effect of their enhancement on speech perception in older adults − A scoping review, Journal of Otology (2016), http://dx.doi.org/10.1016/j.joto.2016.08.001