Original articles

Modulation masking release reduction as a function of time-compressed speech

Fernando Augusto Pacífico[1]
https://orcid.org/0000-0002-5162-0694

Silvana Maria Sobral Griz[2]
https://orcid.org/0000-0002-6759-5964

Denise Costa Menezes[2]
https://orcid.org/0000-0002-2008-4091

Karina Paes Advincula[2]
https://orcid.org/0000-0002-1314-6100

Ana Augusta de Andrade Cordeiro[2]
https://orcid.org/0000-0002-8458-3367

Maria Lucia Gurgel da Costa[2]
https://orcid.org/0000-0001-8614-685X

[1] Universidade Federal de Pernambuco – UFPE, Departamento de Anatomia, Recife, Pernambuco Brasil; Faculdade de Medicina de Olinda, Olinda, Pernambuco, Brasil.

[2] Universidade Federal de Pernambuco – UFPE, Departamento de Fonoaudiologia, Recife, Pernambuco, Brasil.

Conflict of interests: Nonexistent

ABSTRACT

Purpose: to investigate the magnitude of the modulation masking release in sentence recognition as a function of compression level and modulation rate.

Methods: sentences of the Brazilian Portuguese version of the Hearing in Noise Test sentences were used as stimulus. The sentence recognition thresholds were established as a function of speech compression level (0%, 33%, and 50%) in steady and modulated noise at different modulation rates (4, 10, 32 Hz). The analysis of variance was performed for repeated measures, using the 5% significance level.

Results: sentence recognition thresholds were higher for higher compression levels in the different types of noise. However, thresholds were smaller for modulated noises. Also, the magnitude of modulation masking release decreased as speech compression level increased. Nevertheless, no difference was observed in compressed speech between different noise modulation rates, in relation to the speech compression level.

Conclusion: the magnitude of the modulation masking release decreased as the speech time-compression increased. Also, the reductions in modulation masking release, in relation to the speech time-compression level, did not differ between the masking-noise modulation rates (4, 10, and 32 Hz).

Keywords: Perceptual Masking; Speech Perception; Acoustic Stimulation; Acoustics; Hearing

INTRODUCTION

In the last decades, many studies have compared speech recognition in steady noise with speech recognition in modulated noise, both presented with the same speech-to-noise ratio (SNR)[1-3]. In normally hearing subjects, speech recognition performance is substantially better in modulated noise when compared with steady noise – a phenomenon referred to as modulation masking release (MMR)[3].

The MMR can be explained when the intensity levels of the modulated noise are reduced (moments of minimal intensity), providing a listener brief glimpses of speech information to a more favorable SNR[4,5]. The listener's auditory system can temporarily process the masking-noise envelope, with periods when the SNR is less favorable (when noise is modulated in its maximal intensity) and periods when SNR is more favorable (when noise is modulated in its minimal intensity). In other words, the MMR depends in part on the fidelity in which the masking-noise envelope is decoded by the auditory system[1].

Several factors related to masking noise and speech material can change MMR magnitude, such as intensity, interruption rate, cyclic ratio, and modulation depth[6]. The noise modulation rate has been particularly observed to have a significant effect on MMR magnitude. Greater MMR magnitudes for slower modulation rates (at around 10 Hz or lower) have been reported[2,5,7].

Lowest modulation rates have longer durations with smaller noise amplitudes – minimal modulations as compared to higher modulated rates. This provides more time for perception of the target speech – i.e., more time to speech glimpse, contributing to a better speech recognition[2,4]. The MMR is expected to decrease as the modulation rate rises above a given frequency, referred to as the best modulation sensitivity region[1]. For the broadband noise, this frequency is approximately 50 Hz[6].

The MMR magnitude for speech does not vary in relation to different masking-noise modulation rates for 4, 8, 16, and 32 Hz. However, for the modulation rate of 64 Hz and higher, the MMR magnitude is smaller[2]. On the other hand, the MMR for speech remains constant for modulation rates of 2, 10, and 25 Hz. Nevertheless, when a modulation rate of 50 Hz is used, the MMR can be smaller[8].

It is necessary to verify whether there is an MMR magnitude difference using low modulation rates and introducing some new factors, such as time-compressed speech. Also, changes in speech materials are related to the MMR magnitude, such as speech redundancy, which refers to multiple coexisting speech cues, including contextual and coarticulatory cues, and other acoustic signals[1]. Hence, any manipulation that reduces speech redundancy will probably increase the SNR threshold[9].

The ideal masking-noise modulation rate may be different for the various speech materials. For instance, the ideal rate for spondaic words was found to be 1 Hz lower than the ideal rate for other words. This difference may be interpreted in terms of the increase in speech redundancy of the spondaic words. For redundant speech materials, the glimpses may be enough to identify the target word.

Speech redundancy can vary in many dimensions – e.g., contextual integrity (high versus low speech predictability), and acoustic integrity (filtered versus unfiltered speech). Another possibility is to change speech redundancy by manipulating the speech time-compression level[1,3]. The speech time-compression level is normally expressed in the percentage in which the original duration of the speech shape is reduced. For example, speech time-compressed of 33% means that the original time of the target speech was reduced by one third, whereas speech time-compressed of 50% means that the original time was reduced by a half[1].

The increase in speech time-compression levels can increase speech recognition thresholds for the different types of noise. Such increase is greater for modulated noise than steady noise. Consequently, MMR magnitude decreases as speech time-compression level increases[1,3].

MMR magnitudes in relation to the speech time-compression levels and the masking-noise modulation rates are relatively well established in the literature when assessed separeted[8-10]. It has not been well established, though, how MMR magnitude behaves in relation to speech time-compression level in different masking modulation rates, especially in older adults.

Older adults have difficult in understanding speech[4], especially when it is degraded in time and presented in noise with different modulations[2,6,7] – which occurs, for example, in broadcast announcements on radio and television. It is believed that information presented at the end of advertisements may not be fully understood by the older adults. Also, results of this study may contribute to planning speech therapy of individuals with auditory change due to aging.

The main goal of this study was to investigate MMR magnitude through sentence recognition threshold, as a function of speech compression level and noise modulation rate. Doing so, speech recognition thresholds were verified, considering: (i) masking noise type and modulation rates; (ii) speech compression level; and (iii) interaction between masking noise type and its modulation rates in different speech compression level.

METHODS

This study was developed in compliance with Resolution no. 466/12 of the *Conselho Nacional de Saúde* (Brazilian National Health Council), and an original project approved by the Research Ethics Committee of the Department of Health Sciences of the *Universidade Federal de Pernambuco* – UFPE, PE, Brazil, under evaluation report no. 137.884.

Participants

A total of 90 young adults participated in this experiment. They were 45 females and 45 males, aged 17 to 28 years (mean 20.8 years), all native Brazilian Portuguese speakers and with normal hearing (pure-tone thresholds ≤25 dB HL for octave frequencies between 250 and 8000 Hz, and interoctaves of 3000 Hz and 6000 Hz in the tested ear). Participants with a history or diagnosis of otologic or neurological diseases were excluded. Changes in the middle ear were excluded in terms of otologic complaint and acoustic immittance was not performed. All the participants agreed to participate in the study, having signed the informed consent form (ICF).

Material

The speech stimuli used were sentences of Brazilian Portuguese version of HINT. They were presented in their original format, without speech compression (speech time compression [STC] = 0%), and in two speech compression levels, in which the presentation time of the sentence was reduced by one third (STC = 33%) or by half (STC = 50%). The time-compressed speech was conducted using an algorithm (owned by iZotope Radius, in Adobe Audition™) that specifies a fixed change in the wave shape duration, maintaining speech realism.

The stimuli were sent to the listeners from a Dell InCore 7 desktop, connected to an RX6 speech signal processor, Tucker-Davis Technologies. They were presented via Sennheiser HD580 earphones to the right ear – since there is no difference between ear's responses in the analysis of time-compressed speech and modulation rate, and also, this study did not approach laterality.

The masking noise had the same frequency spectrum as the original sentences. The steady noise was presented at the fixed intensity of 65 dB SPL, and the amplitude-modulated noise presented modulations performed by a quadratic wave between 65 and 30 dB SPL, with modulation rates of 4, 10, and 32 Hz.

Procedures

The participants were tested in a sound booth and instructed to repeat each sentence as they perceived it. As each sentence was presented to the subject, written sentence simultaneously appeared to the researcher on the computer screen; all words were highlighted in a shaded marking-sensitive rectangle. The researcher used the computer mouse to mark omitted or incorrectly repeated words. However, as proposed in adaptive procedure, in which sentence recognition thresholds converge to 71% of correct answers, one sentence was considered as "correct sentence" or "incorrect sentence". That is, for a sentence to be considered correct, all words had to be precisely and correctly repeated. Any mistake resulted in an "incorrect sentence" score.

After two correct sentences, the following one was presented 2 dB lower; after an incorrect one, the following sentence was presented 2 dB higher. The threshold was established after six reversals had been acquired. The threshold was calculated considering the mean of the four final reversal levels (intensities). The initial type of masking noise to be presented in the test was randomly chosen. For each participant, three speech recognition thresholds were obtained for every masking-noise condition (steady and modulated at 4, 10, or 32 Hz).

MMR was calculated based on the difference between the mean thresholds of speech recognition in steady noise (taken as a reference) and the mean thresholds of sentence recognition in the different masking-noise modulation rates.

The lists were randomly chosen, and each participant did not listen to any of the sentences more than once, to eliminate learning-related variables. Since the sentences were presented without repetition, each subject could participate in one speech time-compression level (0%, 33%, or 50%) and two

masking-noise conditions (0 and 4 Hz, 0 and 10 Hz, or 0 and 32 Hz). Hence, three 10-person groups were necessary for each speech time-compression level, as the set of eight thresholds (four thresholds for each type of noise) described near the maximum number of sentences of the test, without hearing any of the sentences more than once. The adaptive procedure, including stimulus presentation, was controlled by a personalized script MATLAB™.

Statistical analysis

The variables analyzed were the type of masking noise (stable and modulated noise); the speech time-compression level (0%, 33%, and 50%); and the masking-noise modulation rate (0, 4, and 32 Hz). The analysis of variance (ANOVA) was conducted for repeated measures, with one intra-subject factor (the type of masking noise) and two inter-subject factors (speech time-compression level and masking modulation rate). The 5% significance level was used.

The intra-subject analysis enabled to investigate what was the effect type of masking noise (steady and modulated noise) on the speech recognition thresholds. It also made it possible to investigate: (i) interaction between the type of masking noise and the speech time-compression level in the speech recognition thresholds; (ii) interaction between the type of masking noise and the noise modulation rate in the speech recognition thresholds; and (iii) interaction between the type of masking noise, the speech time-compression level, and the noise modulation rate in the speech recognition thresholds.

The inter-subject analysis allowed for the investigation of: (i) the effect of speech time-compression level on thresholds of speech recognition in steady and modulated noise; (ii) the effect noise modulation rate on the thresholds of speech recognition in steady and modulated noise; and (iii) interaction between speech time-compression level and noise modulation rate on thresholds of speech recognition in steady and modulated noise.

The analysis of variance (ANOVA) was conducted for MMR magnitudes, making it possible to investigate: (i) the effect of speech time-compression level on MMR; (ii) the effect of masking-noise modulation rate on MMR; and (iii) the interaction between time-compressed speech and masking-noise modulation rate in MMR. Likewise, the 5% significance level was used.

RESULTS

Figure 1 and Tables 1, 2, and 3 show the results. The mean speech recognition threshold for different modulation rates can be seen for each compression level. The mean speech thresholds for steady noise and modulated noise at 4 Hz are represented with full and empty circles, respectively; for steady noise and modulated noise at 10 Hz, they are respectively represented as full and empty squares; and, for steady noise and modulated noise at 32 Hz, they are respectively represented as full and empty triangles (error bars of 1 SD).
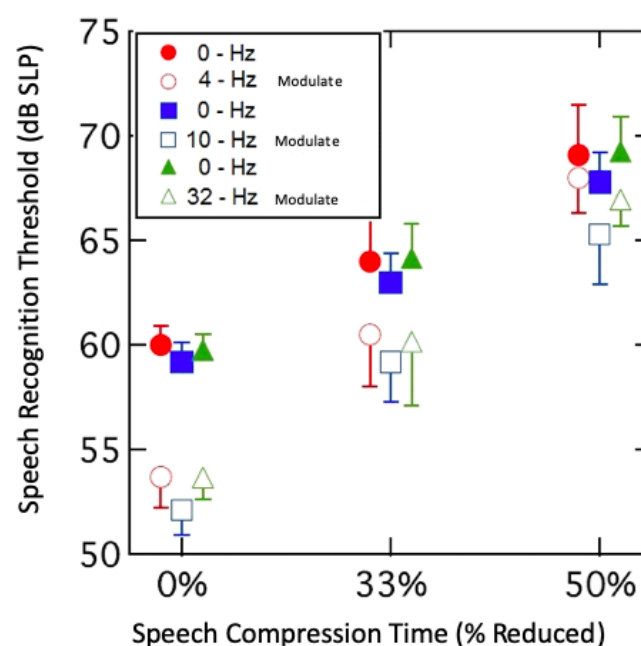
Captions: Mean of the thresholds of recognition in stable (full red circle) and modulated noise at 4 Hz (empty red circle). Mean of the thresholds of recognition in stable (full blue square) and modulated noise at 10 Hz (empty blue square). Mean of the thresholds of recognition in stable (full green triangle) and modulated noise at 32 Hz (empty green triangle). Error bars of 1 SD.

Figure 1. Speech recognition thresholds in the different masking modulation rates for each speech time-compression level

Uncompressed speech (STC = 0%) results show that mean speech recognition thresholds for steady noise and modulated noise at 4 Hz, respectively, 60.0 dB SPL (SD = 0.9 dB) and 53.7 dB SPL (SD = 1.5 dB). The mean speech recognition thresholds for steady and modulated noise at 10 Hz were 59.2 dB SPL (SD = 0.9 dB) and 52.1 dB SPL (SD = 1.2 dB), respectively. Speech recognition thresholds for steady and modulated noise at 32 Hz were, respectively, 59.6 dB

SPL (SD = 0.8 dB) and 53.4 dB SPL (SD = 0.9 dB). Therefore, MMR mean magnitudes were 6.3 dB, 7.1 dB, and 6.2 dB, respectively.

For STC = 33%, speech recognition thresholds means for steady and modulated noise at 4 Hz were, respectively, 64.0 dB SPL (SD = 2.3 dB) and 60.5 dB SPL (SD = 2.5 dB). For steady and modulated noise at 10 Hz, they were 63.0 dB SPL (SD = 1.4 dB) and 59.2 dB SPL (SD = 1.9 dB), respectively. For steady and modulated noise at 32 Hz, they were respectively 64.0 dB SPL (SD = 1.8 dB) and 60.0 dB SPL (SD =

2.9 dB), resulting in an MMR of 3.5, 3.8, and 4.0 dB, respectively.

For STC = 50%, speech recognition thresholds means for steady and modulated noise at 4 Hz were, respectively, 69.0 dB SPL (SD = 2.3 dB) and 67.6 dB SPL (SD = 2.0 dB). For steady and modulated noise at 10 Hz, they were 67.8 dB SPL (SD = 1.4 dB) and 65.3 dB SPL (SD = 2.4 dB), respectively. For steady and modulated noise at 32 Hz they were respectively 69.1 dB SPL (SD = 2.2 dB) and 66.8 dB SPL (SD = 1.1 dB), resulting in 1.4, 2.5, and 2.3 dB of MMR, respectively.

Table 1. Mean of the thresholds of speech recognition in steady and modulated noise, and modulation masking release for each of the three speech time-compression levels and each of the three masking-noise modulation rates

	Mean (SD) dB SPL	Mean (SD) dB SPL	dB
	Steady Noise	**Modulated Noise at 4 Hz**	**MMR**
STC = 0%	60.0 (0.9)	53.7 (1.5)	6.3
STC = 33%	64.0 (2.3)	60.5 (2.5)	3.5
STC = 50%	69.0 (2.3)	67.6 (2.0)	1.4
	Steady Noise	**Modulated Noise at 10 Hz**	**MMR**
STC = 0%	59.2 (0.9)	52.1 (1.2)	7.1
STC = 33%	63.0 (1.4)	59.2 (1.9)	3.8
STC = 50%	67.8 (1.4)	65.3 (2.4)	2.5
	Steady Noise	**Modulated Noise at 32 Hz**	**MMR**
STC = 0%	59.6 (0.8)	53.4 (0.9)	6.2
STC = 33%	64.0 (1.8)	60.0 (2.9)	4.0
STC = 50%	69.1 (2.2)	66.8 (1.1)	2.3

Captions:
STC = 0%, time-compressed speech by 0%; STC = 33%, time-compressed speech by 33%; STC = 50%, time-compressed speech by 50%

The mean of sentence recognition thresholds for steady noise in the different masking modulation rates, as well as MMR for the three speech time-compression levels, point to three findings: (1) speech recognition thresholds for modulated noise are lower than speech recognition thresholds for steady noise – which means a positive MMR in all speech time-compression levels (0%, 33%, and 50%); (2) speech recognition thresholds for steady and modulated noise worsen as the speech time-compression level increases; (3) the increase in speech recognition thresholds magnitude in relation to the increase in the speech time-compression level is greater when the thresholds are obtained in modulated noise when compared with steady noise – resulting in a reduced MMR as the speech time-compression level increased.

As observed in Table 2, the main results of the analysis revealed: (1) a significant main effect of masking noise type (F[1.81] = 350.290; p < 0.001), indicating that speech recognition thresholds for steady noise are higher than the thresholds for modulated noise – resulting in MMR; (2) a significant main effect of speech time-compression level (STC) (F[2.81] = 457.838; p < 0.001), indicating that the speech recognition thresholds increase along with speech time-compression level for both types of masking noise (steady and modulated); (3) a significant inter-action between masking noise type of and speech time-compression level (STC) (F[2.81] = 34.485; p < 0.001), indicating that the difference between speech recognition thresholds for steady and modulated noise depend on the speech time-compression level (it decreases as the STC increases) – i.e., speech

recognition thresholds increased as the speech time-compression level increased for both masking noises. However, a greater increase was observed for speech recognition thresholds for modulated noise than steady noise; (4) no interaction between masking noise type and modulation rate (F[2.81] = 7.594; p < 0.001), demonstrating that difference between the thresholds for steady and modulated noise does not depend on the masking-noise modulation rate; (5) no interaction between the STC and the masking modulation rate (F[4.81] = 0.166; p = 0.955); and (6) no interaction between masking noise type, the STC, and the modulation rate (F[4.81] = 0.260; p = 0.903).

Table 2. Analysis of the effect of the type of masking noise, time-compressed speech, masking modulation rate, and interaction between the variables, in relation to the sentence recognition thresholds

Sources of Variation	F	p-value
Type of masking noise	350.290	0.000*
Time-compressed speech	457.838	0.000*
Type of masking noise Time-compressed speech	34.485	0.000*
Type of masking noise Masking modulation rate	0.949	0.391
Time-compressed speech Masking modulation rate	0.166	0.955
Type of masking noise Time-compressed speech Masking modulation rate	0.260	0.903

* Significant values (p < 0.05) Analysis of variance (ANOVA) with repeated measures
Captions: F = F-value (a test used by ANOVA to statistically assess the equality between the means)

MMR magnitudes were also submitted to analysis of variance (ANOVA) (Table 3). It revealed a significant main effect of STC level (F[2.81] = 34.485; p < 0.001), though it did not reveal MMR masking modulation rate (F[2.81] = 0.949; p = 0.391); the interaction between these two factors (TC level and noise modulation rate) was not significant (F[4.81] = 0.260; p = 0.903). These results indicate that MMR magnitude decreases as the TC increases, although it did not differ between all masking-noise modulation rates.

Table 3. Analysis of time-compressed speech, masking modulation rate, and interaction between the variables, in relation to the modulation masking release

Sources of Variation	F	p-value
Time-compressed speech	34.485	0.000*
Masking modulation rate	0.949	0.391
Time-compressed speech Masking modulation rate	.260	0.903

* Significant values (p < 0.05) Analysis of variance (ANOVA) with repeated or multiple measures
Captions: F = F-value (a test used by ANOVA to statistically assess the equality between the means)

DISCUSSION

Speech recognition thresholds for steady noise are higher than such thresholds for modulated noise, indicating presence of MMR[1-3]. MMR phenomenon was also observed in Brazilian Portuguese linguistic material[2,3], suggesting the auditory system works similarly for verbal sounds, with no distinction between linguistic patterns for different languages[3].

The reduced MMR magnitude in relation to the increased speech time-compression level has already been previously verified[1,3], demonstrating that the speech recognition thresholds depend on both masking-noise type (if steady or modulated) and speech time-compression level (compressed or uncompressed).

In the present study, an increase in speech recognition thresholds was observed for time-compressed speech at 33% and 50%. However, they increased more for steady noise than for modulated noise. As a result, MMR magnitude was greater for uncompressed speech as compared to compressed speech, at different compression levels. Nonetheless, a smaller MMR was observed with increase in speech-time-compression level, which may be related to speech redundancy and SNR.

When limiting acoustic cues, due to time compression, the speech signal redundancy was decreased, as it was easier to recognize speech in part because of the intrinsic redundancy of the auditory system and the extrinsic redundancy of the speech signal.

The intrinsic redundancy of the central auditory nervous system results from bilateral representation of each ear in the brain hemispheres, thalamic nuclei, by route of the crossed pathways, inter- and intra-hemisphere connections, and projections in primary and secondary cortical areas. The extrinsic redundancy results from the acoustic signal, due to the countless cues that help the listener identify the speech signals – e.g., the intensity, time, and duration of the syllables; semantic and syntactic cues; familiarity with and use of the vocabulary; and frequency band of the sequential phonemes[11].

For speech recognition to be efficient, it is frequently not necessary that all the acoustic cues be present, given the integrity of intrinsic redundancy. Nevertheless, when speech is uttered in an unfavorable listening environment (noisy and/or reverberant), these cues (extrinsic redundancy) become highly valuable for speech recognition[11].

In summary, manipulating speech time-compression level decreased speech redundancy, and changed contextual integrity (high versus low speech predictability) and acoustic integrity (filtered versus unfiltered speech)[12]. As the speech time-compression level increased, the existing speech cues decreased, both for steady and modulated noise. That is, the available speech cues were repressed during the moments of minimal modulated noise intensity. As a consequence, speech recognition thresholds in noise increased[13].

Moreover, any manipulation that reduces speech redundancy must increase SNR[9,13]. Manipulations in speech time-compression level decreased the number of speech cues. For compressed speech to be understood, it was necessary to increase speech intensity, both in steady and modulated noise. Since speech recognition thresholds increased in relation to speech time-compression level, with a greater increase for modulated noise than for steady noise, a reduction in MMR magnitude was verified with increase in speech time-compression level[3].

For normal hearing subjects, MMR decreases as SNR increases[9]. This effect in the increase of SNR is perceived when differences are observed in the slopes of the speech recognition psychometric curves, for both steady and modulated noise[9,10,13].

For normal hearing subjects, the higher the SNR, the smaller the MMR. Hence, different slopes of speech recognition psychometric curves in steady and modulated noise can explain by the fact that normal hearing subjects have a reduced MMR[13,14]. This happens because of increasing in SNR in relation to decreasing in speech redundancy – as observed in this study, in which the speech time-compression level was manipulated.

Furthermore, studies comparing different modulation rates have demonstrated lower masking-noise modulation rates do not produce significant changes in MMR magnitude when speech is not time-compressed[7,8]. On the other hand, higher masking-noise modulation rates are similar to steady masking noises in terms of perceptual characteristics[1], as the listener has a short time to benefit from temporal minimal intensity spaces of the masking noise. This makes it more difficult to perceive speech acoustic cues.

CONCLUSIONS

Reduction in MMR for time-compressed speech did not occur, due to the masking-noise modulation rate, probably because the modulation rates used were lower than 50 Hz, as in the uncompressed speech. In other words, it was observed that MMR magnitude did not differ between the masking-noise modulation rates (4, 10, and 32 Hz), in any of the speech time-compression levels (0%, 33%, and 50%).

REFERENCES

1. Grose JH, Mamo SK, Hall JW. Age effects in temporal envelope processing: Speech Unmasking and Auditory Steady State Responses. Ear and Hearing [Internet]. 2009 Oct [cited 2020 Feb 3]; 30(5):568-75. Available from: https://insights.ovid.com/article/00003446-200910000-00009

2. Advíncula KP, Menezes DC, Pacífico FA, Griz SMS. Effect of modulation rate on masking release for speech. Audiol. Commun. Res. [Internet]. 2013 [cited 2020 May 21]; 18(4):238-44. Available from: https://www.scielo.br/pdf/acr/v18n4/en_03.pdf

3. Grose JH, Griz S, Pacífico FA, Advíncula KP, Menezes DC. Modulation masking release using the Brazilian-Portuguese HINT: psychometric functions and the effect of speech time compression. Int J Audiol [Internet]. 2015 Apr 3 [cited 2020 Jun 1];54(4):274-81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25630394

4. Billings CJ, Penman TM, McMillian GP, Ellis E. Electrophysiology and perception of speech in noise in older listeners: effects of hearing impairment & age. Ear Hear. 2015;36(6):710-22. DOI:10.1097/AUD.0000000000000191

5. Maamor N, Billings C. Cortical signal-in-noise coding varies by noise type, signal-to-noise ratio, age, and hearing status. Neurosci Lett. 2017;636:258-64. DOI:10.1016/j.neulet. 2016.11.020

6. Advíncula KP, Menezes DC, Pacifico FA, Costa MLG, Griz SMS. Age effects in temporal auditory processing: modulation masking release and forward masking effect. Audiol. Commun. Res. [Internet]. 2018 [cited 2020 May 21]; 23:e1861. Available from: https://www.scielo.br/pdf/acr/v23/2317-6431-acr-23-e1861.pdf

7. Tanner MA, Spitzer ER, Hyzy JP, Grose JH. Masking release for speech in modulated maskers: electrophysiological and behavioral measures. Ear Hear [Internet]. 2019 Jul/Aug. [cited 2020 May 21]; 40(4):1009-15. Available from: https://pubmed.ncbi.nlm.nih.gov/30557224/

8. Dubno JR, Horwitz AR, Ahlstrom JB. Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. J Acoust Soc Am [Internet]. 2003 Apr [cited 2020 May 29];113(4Pt1):2084-94. Available from: http://scitation.aip.org/content/asa/journal/jasa/113/4/10.1121/1.1555611

9. Grose JH, Menezes DC, Poter HL, Griz S. Masking period patterns & forward masking for speech-shaped noise: age-related effects. Ear Hear. 2016;37(1):48-54. DOI:10.1097/AUD.0000000000000200

10. Dirks DD, Wilson RH, Bower DR. Effect of pulsed masking on selected speech materials. J Acoust Soc Am [Internet]. 1969 Oct [cited 2020 Apr 15]; 46(4B):898-906. Available from: http://www.ncbi.nlm.nih.gov/pubmed/5824033

11. Yan Z, Dong X, Qunyan R, Shengming G, Fuyuan M. An effective speech compression based on syllable division. ASA. 2017;29:1-8. DOI: 10.1121/2.0000480

12. Carina P, Anastasios S, Mart VD, Deniz B. Effects of additional low-pass–filtered speech on listening effort for noise-band–vocoded speech in quiet and in noise. Ear Hear. 2018;40(1):3-17. DOI: 10.1097/AUD.0000000000000587.

13. Sobon KA, Taleb NM, Buss E, Grose JH, Calandruccio L. Psychometric function slope for speech-in-noise and speech-in-speech: effects of development and aging. J Acoust Soc Am. 2019;145(4): doi.org/10.1121/1.5097377.

14. Hall JW, Buss E, Grose JH. Factors affecting the development of speech recognition in steady and modulated noise. J Acoust Soc Am. 2016;139(5): doi.org/10.1121/1.4950810.