

Artigos originais

Redução do benefício da modulação do mascaramento em função da compressão temporal da fala

*Modulation masking release reduction as a function of time-compressed speech*Fernando Augusto Pacífico¹<https://orcid.org/0000-0002-5162-0694>Silvana Maria Sobral Griz²<https://orcid.org/0000-0002-6759-5964>Denise Costa Menezes²<https://orcid.org/0000-0002-2008-4091>Karina Paes Advincula²<https://orcid.org/0000-0002-1314-6100>Ana Augusta de Andrade Cordeiro²<https://orcid.org/0000-0002-8458-3367>Maria Lucia Gurgel da Costa²<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.

Fonte de auxílio à pesquisa: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) / Código de financiamento: 001.

Conflito de interesses: Inexistente



Recebido em: 19/06/2020

Aceito em: 22/09/20

Endereço para correspondência:

Silvana Maria Sobral Griz
Universidade Federal de Pernambuco -
Departamento de Fonoaudiologia
Avenida Professor Artur de Sá, s/n
Cidade Universitária
CEP: 50670-420 – Recife, Pernambuco,
Brasil
E-mail: silvana.griz@hotmail.com

RESUMO

Objetivo: investigar a magnitude do benefício da modulação do mascaramento no reconhecimento de sentenças, em função do nível de compressão temporal da fala e da taxa de modulação do ruído.

Métodos: foram utilizadas sentenças do *Hearing in Noise Test* versão Português do Brasil. Foram determinados os limiares de reconhecimento das sentenças em função do nível de compressão temporal da fala (0%, 33% e 50%) em presença de ruído estável e modulado, em diferentes taxas de modulação (4, 10, 32 Hz). Foi realizada uma análise de variância para medidas repetidas, adotando nível de significância de 5%.

Resultados: os limiares de reconhecimento de sentenças foram mais elevados com o aumento do nível de compressão temporal nos diferentes tipos de ruídos, no entanto, foram menores em presença dos ruídos modulados. Além disso, a magnitude do benefício da modulação do mascaramento diminuiu com o aumento do nível de compressão temporal da fala. Contudo, não foi observada diferença na fala comprimida entre as diferentes taxas de modulação do ruído, em função do nível de compressão temporal da fala.

Conclusão: a magnitude do benefício do mascaramento modulado diminuiu com o aumento da compressão temporal da fala e as reduções no benefício do mascaramento modulado em função do nível de compressão temporal da fala não diferiram entre as taxas de modulações do ruído mascarante (4 Hz, 10 Hz e 32 Hz).

Descritores: Mascaramento Perceptivo; Percepção de Fala; Estimulação Acústica; Acústica; Audição

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

INTRODUÇÃO

Nas últimas décadas, muitos estudos têm comparado o reconhecimento de fala em presença de ruído estável com o reconhecimento de fala em presença de ruído modulado, apresentados na mesma relação fala-ruído (RFR)¹⁻³. Em ouvintes com audição normal, o desempenho no reconhecimento de fala é substancialmente melhor em presença de um ruído modulado comparado ao ruído estável, fenômeno referido como *modulation masking release*³, traduzido para o Português como *benefício da modulação do mascaramento* (BMM).

O BMM pode ser explicado pelos momentos na redução dos níveis de intensidade do ruído modulado (momentos de mínima intensidade), proporcionando ao ouvinte breves visualizações da informação de fala, e RFR mais favorável^{4,5}.

O sistema auditivo do ouvinte é capaz de processar temporalmente o envelope do ruído mascarante, havendo períodos em que a RFR é menos favorável (quando o ruído está modulado em sua máxima intensidade) e períodos em que a RFR é mais favorável (quando o ruído está modulado em sua mínima intensidade). Em outras palavras, o BMM depende, em parte, da fidelidade com que o envelope do ruído mascarante é codificado pelo sistema auditivo¹.

Uma variedade de fatores relacionados ao ruído mascarante e ao material de fala pode alterar a magnitude do BMM, que é dependente de várias características do ruído modulado, como a intensidade, a taxa de interrupção, a razão cíclica e a profundidade da modulação⁶. Tem-se observado, em especial, um efeito significativo da taxa de modulação do ruído na magnitude do BMM. Maiores magnitudes do BMM têm sido relatadas para taxas de modulação mais lentas, em torno de 10 Hz ou inferiores^{2,5,7}.

As menores taxas de modulação do ruído modulado contêm mais espaço de tempo com menor amplitude do ruído - modulações mínimas, quando comparadas às taxas de modulação do ruído modulado mais altas. Isso possibilita mais tempo para a percepção da fala alvo, ou seja, maior espaço de tempo para as visualizações, contribuindo para um melhor reconhecimento da fala^{2,4}. É esperado que o BMM diminua à medida que a taxa de modulação aumente acima de uma determinada frequência, referida como região de melhor sensibilidade da modulação¹. Para o ruído de banda larga, esta frequência é em torno de 50 Hz⁶.

A magnitude do BMM para fala não varia em função das diferentes taxas de modulação do ruído

mascarante para 4, 8, 16 e 32 Hz. No entanto, para a taxa de modulação de 64 Hz, a magnitude do BMM é reduzida². Por outro lado, o BMM para fala permanece constante para as frequências de modulação do ruído mascarante para taxas de 2 Hz, 10 Hz e 25 Hz. Contudo, quando utilizada uma taxa de modulação do ruído mascarante de 50 Hz, o BMM pode ser reduzido⁸.

Há necessidade de verificar se existe diferenciação da magnitude do BMM utilizando taxas de modulação do ruído mascarante menores com a introdução de algum fator novo, tal como a compressão da fala. Além dos fatores relacionados ao ruído mascarante, modificações do material de fala também estão relacionadas com a magnitude do BMM. Destaca-se a redundância de fala, que se refere à multiplicidade de pistas de fala coexistentes, incluindo pistas contextuais, co-articulatórias e outros sinais acústicos¹. Então, qualquer manipulação que reduza a redundância da fala deve resultar em um aumento do limiar de RFR⁹.

A taxa ideal de modulação do ruído mascarante pode ser diferente para diversos materiais de fala. Por exemplo, a taxa ideal encontrada para palavras espondáicas foi 1 Hz inferior à taxa ideal para outras palavras. Esta diferença pode ser interpretada em função do aumento da redundância de fala das palavras espondáicas. Para materiais de fala redundantes, as visualizações podem ser suficientes para identificar a palavra alvo.

Existem várias dimensões nas quais a redundância de fala pode ser variada, como a integridade contextual - alta previsibilidade contra baixa previsibilidade de fala e a integridade acústica - fala filtrada contra fala não filtrada. Outra possibilidade é alterar a redundância da fala por meio da manipulação do nível de compressão temporal da fala^{1,3}. O nível de compressão temporal da fala é normalmente expresso pela percentagem em que a duração original da forma da onda da fala é removida. Por exemplo, 33% de compressão temporal da fala implica em dizer que um terço do tempo original da fala alvo foi reduzido, enquanto 50% de compressão temporal da fala implica em dizer que metade do tempo original foi reduzido¹.

O aumento do nível de compressão temporal da fala pode acarretar em aumento dos limiares de reconhecimento de fala para os diferentes tipos de ruídos, sendo este aumento maior para o ruído modulado que para ruído estável. Como consequência, há uma redução na magnitude do BMM com o aumento do nível de compressão temporal da fala^{1,3}.

As magnitudes do BMM em função do nível de compressão temporal da fala e em função da taxa de modulação do ruído mascarante estão relativamente bem estabelecidas na literatura quando avaliadas de forma isolada⁸⁻¹⁰. No entanto, como se comporta a magnitude do BMM em função do nível de compressão temporal da fala nas diferentes taxas de modulação do ruído mascarante, especialmente em idosos, não se encontra bem estabelecido.

A importância deste estudo se dá pelo fato dos idosos apresentarem dificuldade de compreender a fala⁴, especialmente se esta for degradada em tempo e se a fala for apresentada em ruídos de diversas modulações^{2,6,7}, a exemplo do que ocorre em veiculação de anúncios de rádio e televisão. Acredita-se que as informações contidas ao final das propagandas podem não ser compreendidas integralmente pelos idosos. Além disso, poderá contribuir para o planejamento terapêutico de indivíduos com alteração auditiva decorrente do envelhecimento.

Buscando compreender como a taxa de modulação do ruído mascarante e nível de compressão temporal da fala interagem e influenciam o BMM, o objetivo desta pesquisa foi investigar a magnitude do benefício da modulação do mascaramento no reconhecimento de sentenças, em função do nível de compressão temporal da fala e da taxa de modulação do ruído. Para tal, verificaram-se os limiares de reconhecimento de fala, considerando: (i) o tipo de ruído mascarante e as taxas de modulação; (ii) o nível de compressão temporal da fala; (iii) a interação entre o tipo de ruído mascarante e suas taxas de modulação em diferentes níveis de compressão temporal da fala.

MÉTODOS

Este estudo foi elaborado com bases na Resolução nº466/12 do Conselho Nacional de Saúde, e o projeto foi aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal de Pernambuco - UFPE, PE, Brasil, sob Parecer número 137.884.

Participantes

Participaram desse experimento 90 adultos jovens, sendo 45 do sexo feminino e 45 do sexo masculino, com idade entre 17 a 28 anos - média de 20,8 anos, todos falantes nativos do Português do Brasil e com audição normal (limiares tonais ≤ 25 dB NA para as frequências de oitava entre 250-8000 Hz e interoitavas

3000 Hz e 6000 Hz na orelha de teste). Foram excluídos os participantes com histórico ou diagnóstico de doenças otológicas ou neurológicas. As alterações de orelha média foram pesquisadas apenas pela queixa otológica, não tendo sido realizada imitanciometria. Todos concordaram em participar do estudo por meio da assinatura do Termo de Consentimento Livre e Esclarecido (TCLE).

Material

Os estímulos de fala utilizados foram as sentenças do HINT na versão do Português Brasileiro, apresentadas em seu formato original, sem compressão temporal da fala (compressão temporal da fala [CTF] = 0%), e em dois níveis de compressão temporal da fala, nos quais um terço (CTF = 33%) ou metade (CTF = 50%) do tempo de apresentação da sentença haviam sido removidos. A compressão temporal da fala foi realizada usando o algoritmo do proprietário *iZotope Radius*, no *Adobe Audition*TM, que especifica uma mudança estipulada de duração da forma de onda, mantendo o realismo da fala.

Os estímulos foram enviados para o ouvinte por meio do Desktop - Dell InCore 7, conectado ao processador de sinal fala RX6, Tucker-Davis Technologies. Estes foram apresentados à orelha direita por meio de fones auditivos *Sennheiser* HD580, uma vez que para a análise da compressão de fala e da taxa de modulação não há diferença entre as respostas das orelhas, além do fato do estudo não envolver a questão da lateralidade.

O ruído mascarante possuía o mesmo espectro de frequência das sentenças originais. O ruído estável foi apresentado em uma intensidade fixa de 65 dB NPS e o ruído modulado em amplitude apresentou modulações realizadas por uma onda quadrática entre 65 e 30 dB NPS, com taxas de modulação de 4 Hz, 10 Hz e 32 Hz.

Procedimentos

Os participantes foram testados em uma cabina acústica e orientados a repetir cada sentença da forma como era percebida. À medida que cada sentença era apresentada ao sujeito, o texto da sentença também era apresentado ao pesquisador, simultaneamente, na tela do computador, com todas as palavras em destaque num retângulo sombreado sensível à marcação. O pesquisador utilizou o *mouse* do computador para marcar as palavras que foram omitidas ou repetidas incorretamente. No entanto, para a proposta

do procedimento adaptativo, no qual os limiares de reconhecimento de sentença convergem para 71% de acerto, a sentença foi considerada em uma pontuação global de “sentença correta” ou “sentença incorreta”. Ou seja, para ser considerada “sentença correta”, todas as palavras deveriam ser repetidas com precisão, de forma correta. Qualquer erro resultou em uma pontuação “sentença incorreta”.

Após duas sentenças corretas, o nível de apresentação da próxima sentença era reduzido em 2 dB; após uma sentença incorreta, o nível de apresentação da próxima sentença era aumentado em 2 dB. O limiar foi determinado após seis reversões serem adquiridas. O limiar final foi calculado—considerando-se a média dos quatro níveis (intensidades) finais de reversão. A escolha do tipo de ruído mascarante inicial a ser apresentado no teste foi feita de forma randômica. Para cada participante, três limiares de reconhecimento de fala foram obtidos para cada condição de ruído mascarante (estável e modulado a 4 Hz, 10 Hz ou 32 Hz).

O cálculo do BMM foi realizado a partir da diferença entre a média dos limiares de reconhecimento de fala em presença de ruído estável, tomado como referência, e a média dos limiares de reconhecimento de sentenças em presença das diferentes taxas de modulação do ruído mascarante.

A escolha das listas foi feita de forma aleatória, bem como o participante não escutou qualquer sentença mais de uma vez, a fim de eliminar variáveis relacionadas ao fenômeno de aprendizagem. Pelo fato das sentenças terem sido apresentadas sem repetição, cada sujeito pôde participar de apenas um nível de compressão temporal de fala (0%, 33%, ou 50%) e duas condições de ruído mascarante (0 Hz e 4 Hz, 0 e 10 Hz, ou 0 e 32 Hz). Isso fez com que fossem necessários três grupos de 10 participantes para cada nível de compressão temporal da fala, pois o conjunto de oito limiares (quatro limiares para cada tipo de ruído) aproximou-se do número máximo de sentenças do teste, sem que escutasse qualquer sentença mais de uma vez. O procedimento adaptativo, incluindo a apresentação do estímulo, foi controlado por um *script MATLAB™* personalizado.

Análise Estatística

As variáveis analisadas foram o tipo de ruído mascarante - ruído estável e ruído modulado; o nível de compressão do temporal da fala - 0%, 33% e 50%; e a taxa de modulação do ruído mascarante - 0 Hz, 4 Hz e 32 Hz. Foi realizada uma análise de variância (ANOVA) para medidas repetidas, com um fator

intra-sujeitos (tipo de ruído mascarante) e dois fatores inter-sujeitos (nível de compressão temporal da fala e taxa de modulação do mascaramento), adotando nível de significância de 5%.

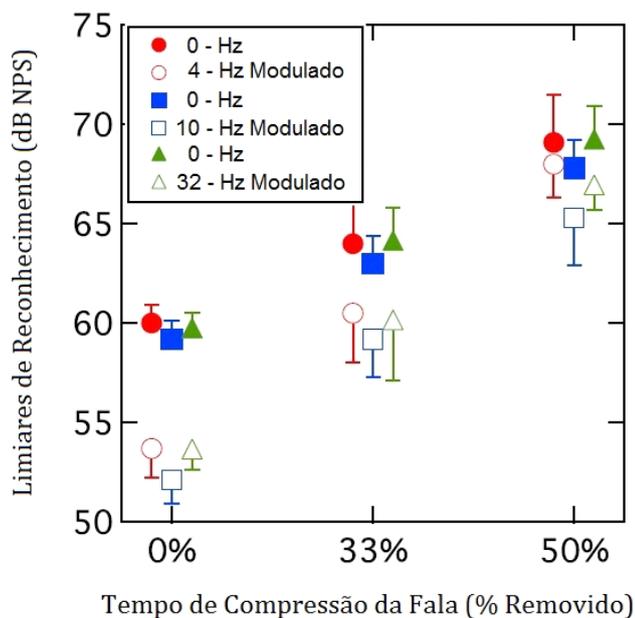
A análise intra-sujeitos permitiu investigar o efeito do tipo de ruído mascarante (ruído estável e modulado) nos limiares de reconhecimento de fala; bem como investigar: (i) a interação entre o tipo de ruído mascarante e o nível de compressão temporal da fala nos limiares de reconhecimento de fala; (ii) a interação entre o tipo de ruído mascarante e a taxa de modulação do ruído nos limiares de reconhecimento de fala; e (iii) a interação entre o tipo de ruído mascarante, o nível de compressão temporal da fala e a taxa de modulação do ruído nos limiares de reconhecimento de fala.

A análise inter-sujeitos possibilitou investigar: (i) o efeito do nível de compressão temporal da fala nos limiares de reconhecimento de fala em presença de ruído estável e modulado; (ii) o efeito da taxa de modulação do ruído nos limiares de reconhecimento de fala em presença de ruído estável e modulado; e (iii) a interação entre o nível de compressão temporal da fala e a taxa de modulação do ruído nos limiares de reconhecimento de fala em presença de ruído estável e modulado.

Por fim, foi realizada uma análise de variância (ANOVA) para as magnitudes do BMM, possibilitando investigar: (i) o efeito do nível de compressão temporal da fala no BMM; (ii) o efeito da taxa de modulação do ruído mascarante no BMM; e (iii) a interação entre a compressão temporal da fala e a taxa de modulação do ruído mascarante no BMM. Também foi adotado o nível de significância de 5%.

RESULTADOS

Os resultados são apresentados na Figura 1 e Tabelas 1, 2 e 3. Observa-se que a média dos limiares de reconhecimento de fala nas diferentes taxas de modulação do mascaramento é apresentada para cada nível de compressão temporal da fala. A média dos limiares encontrados em presença de ruído estável comparados ao ruído modulado a 4 Hz é apresentada, respectivamente, como círculos cheios e vazios; em presença de ruído estável comparado ao ruído modulado a 10 Hz é apresentada, respectivamente, como quadrados cheios e vazios; e por fim, em presença de ruído estável comparado ao ruído modulado a 32 Hz é apresentada, respectivamente, como triângulos cheios e vazios (barras de erro de 1 DP).



Legenda: Média dos limiares de reconhecimento em presença de ruído estável (círculo vermelho cheio) e modulado a 4 Hz (círculo vermelho vazio). Média dos limiares de reconhecimento em presença de ruído estável (quadrado azul cheio) e modulado a 10 Hz (quadrado azul vazio). Média dos limiares de reconhecimento em presença de ruído estável (triângulo verde cheio) e modulado a 32 Hz (triângulo verde vazio). Barras de erro de 1DP.

Figura 1. Limiares de reconhecimento de fala nas diferentes taxas de modulação do mascaramento para cada nível de compressão temporal da fala

Os resultados obtidos com a fala não comprimida (CTF = 0%) mostram uma média dos limiares de reconhecimento de fala em presença de ruído estável e de ruído modulado a 4 Hz, respectivamente de

60,0 dB NPS (DP= 0,9 dB) e 53,7 dB NPS (DP= 1,5dB). A média dos limiares de reconhecimento de fala em presença de ruído estável e modulado a 10 Hz foram de 59,2 dB NPS (DP= 0,9 dB) e 52,1 dB NPS (DP= 1,2 dB), respectivamente. Por fim, os limiares de reconhecimento de fala na presença de ruído estável e modulado a 32 Hz foram, respectivamente, de 59,6 dB NPS (DP = 0,8dB) e 53,4 dB NPS (DP = 0,9 dB). A magnitude média do benefício do mascaramento modulado (BMM) foi, portanto, de 6,3 dB; 7,1 dB e 6,2 dB, respectivamente.

Para CTF= 33%, a média dos limiares de reconhecimento de fala no ruído estável e modulado a 4 Hz foram, respectivamente, de 64,0 dB NPS (DP= 2,3dB) e 60,5 dB NPS (DP= 2,5dB). Para o ruído estável e modulado a 10 Hz foram de 63,0 dB NPS (DP= 1,4 dB) e 59,2 dB NPS (DP= 1,9dB), respectivamente. Por fim, para o ruído estável e modulado a 32 Hz foram, respectivamente, de 64,0dB NPS (DP = 1,8dB), e 60,0dB NPS (DP = 2,9dB), produzindo um BMM de 3,5dB, 3,8 e 4,0 dB, respectivamente.

Para CTF = 50%, a média dos limiares de reconhecimento de fala na presença de ruído estável e modulado a 4 Hz foi, respectivamente, de 69,0 dB NPS (DP= 2,3dB) e 67,6 dB NPS (DP= 2,0dB). Para o ruído estável e o ruído modulado a 10 Hz foi de 67,8 dB NPS (DP= 1,4 dB) e 65,3 dB NPS (DP= 2,4dB), respectivamente. Para o ruído estável e modulado a 32 Hz foi, respectivamente, de 69,1dB NPS (DP = 2,2 dB) e 66,8dB NPS (DP = 1,1dB), resultando em 1,4dB, 2,5 e 2,3 dB de BMM, respectivamente.

Tabela 1. Média dos limiares de reconhecimento de fala na presença de ruído estável e modulado, e benefício da modulação do mascaramento para cada um dos três níveis de compressão temporal da fala e para cada uma das três taxas de modulação do ruído mascarante

	Média (DP) dB NPS	Média (DP) dB NPS	dB
	Ruído Estável	Ruído Modulado a 4 Hz	BMM
CTF = 0%	60,0 (0,9)	53,7 (1,5)	6,3
CTF = 33%	64,0 (2,3)	60,5 (2,5)	3,5
CTF = 50%	69,0 (2,3)	67,6 (2,0)	1,4
	Ruído Estável	Ruído Modulado a 10 Hz	BMM
CTF = 0%	59,2 (0,9)	52,1 (1,2)	7,1
CTF = 33%	63,0 (1,4)	59,2 (1,9)	3,8
CTF = 50%	67,8 (1,4)	65,3 (2,4)	2,5
	Ruído Estável	Ruído Modulado a 32 Hz	BMM
CTF = 0%	59,6 (0,8)	53,4 (0,9)	6,2
CTF = 33%	64,0 (1,8)	60,0 (2,9)	4,0
CTF = 50%	69,1 (2,2)	66,8 (1,1)	2,3

Legenda:

CTF = 0%, Compressão Temporal da Fala em 0%; CTF = 33%, Compressão Temporal da Fala em 33%; CTF = 50%, Compressão Temporal da Fala em 50%

A média dos limiares de reconhecimento de sentença em presença de ruído estável e nas diferentes taxas de modulação do mascaramento, bem como o BMM para os três níveis de compressão temporal da fala aponta para três achados: (1) os limiares de reconhecimento de fala na presença de ruído modulado são menores que os limiares de reconhecimento de fala na presença de ruído estável, significando a ocorrência de um BMM positivo nos três níveis de compressão temporal da fala (0%, 33% e 50%); (2) Os limiares de reconhecimento de fala na presença de ruído estável e modulado pioram à medida que o nível de compressão temporal da fala aumenta; (3) A magnitude do aumento do limiar de reconhecimento de fala em função do aumento do nível de compressão temporal da fala é maior quando esses limiares são obtidos em presença de ruído modulado comparado ao ruído estável, resultando numa redução no BMM com o aumento do nível de compressão temporal da fala.

Como observado na Tabela 2, os principais resultados da análise revelaram: (1) um efeito principal significativo do tipo de ruído mascarante ($F[1,81] = 350,290$; $p < 0,001$), indicando que os limiares de reconhecimento de fala em presença de ruído estável são mais elevados que os limiares em presença de ruído modulado, resultando na ocorrência do benefício do mascaramento modulado (BMM); (2) um efeito

principal significativo do nível de compressão temporal da fala (CTF) ($F[2,81] = 457,838$; $p < 0,001$), indicando que os limiares de reconhecimento de fala aumentam com o aumento do nível de compressão temporal da fala para ambos os tipos de ruído mascarante (estável e modulado); (3) uma interação significativa entre o tipo de ruído mascarante e o nível de compressão temporal da fala (CTF) ($F[2,81] = 34,485$; $p < 0,001$), indicando que a diferença entre os limiares de reconhecimento de fala na presença de ruído estável e modulado dependem do nível de compressão temporal da fala (diminui com o aumento da CTF), isto é, os limiares de reconhecimento de fala aumentaram com o aumento do nível de compressão temporal da fala para ambos os ruídos mascarantes, no entanto, foi observado um maior aumento dos limiares de reconhecimento de fala para ruído modulado que para o ruído estável; (4) nenhuma interação entre o tipo de ruído mascarante e a taxa de modulação ($F[2,81] = 7,594$; $p < 0,001$), demonstrando que a diferença entre os limiares em presença de ruído estável e modulado não dependem da taxa de modulação do ruído mascarante; (5) nenhuma interação entre o CTF e a taxa de modulação do mascaramento ($F[4,81] = 0,166$; $p = 0,955$); e (6) nenhuma interação entre tipo de ruído mascarante, a CTF e a taxa de modulação ($F[4,81] = 0,260$; $p = 0,903$).

Tabela 2. Análise do efeito do tipo de ruído mascarante, da compressão temporal da fala, da taxa de modulação do mascaramento e da interação entre as variáveis sobre os limiares de reconhecimento de sentenças

Fontes de Variação	F	Significância
Tipo de ruído mascarante	350,290	0,000*
Compressão Temporal da Fala	457,838	0,000*
Tipo de ruído mascarante Compressão Temporal da Fala	34,485	0,000*
Tipo de ruído mascarante Taxa de Modulação do Mascaramento	0,949	0,391
Compressão Temporal da Fala Taxa de Modulação do Mascaramento	0,166	0,955
Tipo de ruído mascarante Compressão Temporal da Fala Taxa de Modulação do Mascaramento	0,260	0,903

* Valores significantes ($p < 0,05$) Análise de variância (ANOVA) com medidas repetidas

Legenda: F = Valor de F (teste utilizado pela ANOVA para testar estatisticamente a igualdade entre médias)

As magnitudes do benefício do mascaramento modulado também foram submetidas a uma análise de variância (ANOVA) (Tabela 3). A análise revelou um efeito principal significativo do nível de CTF ($F[2,81] = 34.485$; $p < 0,001$), mas não revelou um efeito da taxa de modulação do mascaramento no BMM ($F[2,81] =$

0.949 ; $p = 0,391$); a interação entre esses dois fatores (nível CT e a taxa de modulação do ruído) não foi significativa ($F[4,81] = 0.260$; $p = 0,903$). Estes resultados indicam que a magnitude do BMM diminui com o aumento do CT, mas que esta não diferiu entre as três taxas de modulação do ruído mascarante.

Tabela 3. Análise da compressão temporal da fala, da taxa de modulação do mascaramento e da interação entre as variáveis sobre os benefícios da modulação do mascaramento

Fontes de Variação	F	Significância
Compressão Temporal da Fala	34.485	0,000*
Taxa de Modulação do Mascaramento	0,949	0,391
Compressão Temporal da Fala Taxa de Modulação do Mascaramento	.260	0,903

* Valores significantes ($p < 0,05$) Análise de variância (ANOVA) com medidas repetidas ou múltiplas **Legenda:** F = Valor de F (teste utilizado pela ANOVA para testar estatisticamente a igualdade entre médias)

DISCUSSÃO

É sabido que limiares de reconhecimento de fala em presença de ruído estável são mais elevados que limiares em presença de ruído modulado, denotando a existência de BMM¹⁻³. O fenômeno do BMM também foi observado com a utilização do material linguístico do Português Brasileiro^{2,3}, podendo indicar que o sistema auditivo funciona de maneira semelhante para os sons verbais, não havendo distinção entre padrões linguísticos entre as diferentes línguas³.

A redução da magnitude do BMM em função do aumento do nível de compressão temporal da fala já foi verificada anteriormente^{1,3}, demonstrando que os limiares de reconhecimento de fala dependiam tanto do tipo de ruído mascarante, se estável ou modulado, como do nível de compressão temporal da fala, comprimida ou com compressão temporal.

No presente estudo, foi observado um aumento dos limiares de reconhecimento de fala para a compressão temporal da fala a 33% e 50%, no entanto, com maior aumento para o ruído estável do que para o ruído modulado. Como resultado, a magnitude da BMM foi maior para fala não comprimida do que para a fala comprimida em diferentes níveis de compressão. Entretanto observou-se menor BMM com o aumento do nível de compressão temporal da fala podendo estar relacionado à redundância da fala e à relação fala-ruído (RFR).

Ao se restringir pistas acústicas, devido à compressão temporal, houve uma redução na

redundância do sinal de fala, visto que a facilidade em reconhecer a fala deve-se, em parte, à redundância intrínseca do sistema auditivo e à redundância extrínseca do sinal de fala.

A redundância intrínseca do sistema nervoso auditivo central é dada pela representação bilateral de cada orelha nos hemisférios cerebrais, núcleos reles, pelas vias cruzadas, pelas conexões inter e intra-hemisféricas e pelas projeções em áreas corticais primárias e secundárias. A redundância extrínseca é dada pelo sinal acústico devido às inúmeras pistas existentes que auxiliam o ouvinte a identificar os sinais de fala, a exemplo da intensidade, tempo e duração das sílabas; pistas semânticas e sintáticas; familiaridade e uso do vocabulário e faixa de frequência dos fonemas em sequência¹¹.

Para que o reconhecimento de fala ocorra de forma eficiente, não é frequentemente necessário que todas as pistas acústicas estejam presentes, tendo em vista a integridade da redundância intrínseca. No entanto, quando a fala é dita em um ambiente desfavorável de escuta (ruidoso e/ou reverberante), estas pistas (redundância extrínseca) passam a ser de grande valia para o reconhecimento de fala¹¹.

Em síntese, a manipulação do nível de compressão temporal da fala diminuiu a redundância de fala, modificou a integridade contextual (alta previsibilidade contra baixa previsibilidade de fala) e a integridade acústica (fala filtrada contra fala não filtrada)¹². Durante o aumento do nível de compressão temporal da fala

ocorreu uma diminuição das pistas de fala existentes, tanto para o ruído estável como para o ruído modulado. Isto é, as pistas de fala disponíveis se tornaram reprimidas durante os momentos de mínima intensidade do ruído modulado e, como consequência, houve uma elevação dos limiares de reconhecimento de fala na presença de ruído¹³.

Ademais, qualquer manipulação que reduza a redundância da fala deve resultar em um aumento da relação fala-ruído^{9,13}. Manipulações no nível de compressão temporal da fala causaram diminuições da quantidade de pistas de fala existentes. Para que a fala comprimida fosse compreendida, houve necessidade de aumentar a intensidade da fala, tanto na presença de ruído estável quanto na presença de ruído modulado. Como os limiares de reconhecimento aumentaram em função do nível de compressão temporal da fala, com maior aumento para o ruído modulado do que para o ruído estável, verificou-se uma redução na magnitude do BMM com o aumento do nível de compressão temporal da fala⁹.

Verifica-se, para ouvintes com audição normal, que o BMM diminui com o aumento da relação fala-ruído⁹. Esse efeito do aumento da relação fala-ruído é percebido quando se observam as diferenças das inclinações (*slopes*) das curvas psicométricas do reconhecimento de fala, para o ruído estável e para o modulado^{9,10,13}.

Para ouvintes com audição normal, quanto maior for a relação fala-ruído, menor é o benefício do mascaramento modulado. Dessa forma, as diferentes inclinações das curvas psicométricas para o reconhecimento de fala, em presença de ruído estável e modulado, podem explicar o fato de que indivíduos com audição normal apresentam uma redução no BMM^{13,14}. Isto ocorre porque há o aumento da RFR em função da diminuição na redundância da fala, como observado neste estudo, em que houve manipulação do nível de compressão temporal da fala.

Além disso, estudos comparando diferentes taxas de modulações têm demonstrado que taxas de modulações do ruído mascarante mais lentas não produzem alterações significantes na magnitude do BMM, quando a fala não está comprimida em tempo^{7,8}. Por outro lado, taxas de modulações de ruído mascarante mais elevadas se assemelham, em características perceptuais, a ruídos mascarantes estáveis¹, pois o tempo em que o ouvinte pode se beneficiar dos espaços temporais de mínima intensidade do ruído mascarante são curtos, dificultando a percepção das pistas acústicas de fala.

CONCLUSÕES

A redução do BMM para a fala comprimida em tempo não ocorreu devido à taxa de modulação do ruído mascarante, provavelmente porque foram utilizadas taxas de modulações abaixo de 50 Hz, de maneira semelhante à fala não comprimida. Em outras palavras, observou-se que a magnitude do BMM não diferiu entre as taxas de modulações do ruído mascarante estudadas (4 Hz, 10 Hz e 32 Hz) em nenhum dos níveis de compressão temporal da fala (0%, 33% e 50%).

REFERÊNCIAS

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.000000000000191
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, Pacífico 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.