

Percepção da fala em presença de ruído competitivo: o efeito da taxa de modulação do ruído mascarante

Effect of modulation rate on masking release for speech

Karina Paes Advíncula¹, Denise Costa Menezes², Fernando Augusto Pacífico³, Silvana Maria Sobral Griz²

RESUMO

Objetivo: Este estudo investigou o efeito das diferentes taxas de modulações do mascaramento na magnitude do *masking release*. **Métodos:** Quinze indivíduos jovens, com audição normal, foram submetidos ao teste de reconhecimento de sentença na presença de ruído, utilizando as listas de sentenças do HINT-Brasil. Foram obtidos limiares de reconhecimento de fala em presença de ruído estável e ruído modulado, em diferentes taxas de modulação (4, 8, 16, 32 e 64 Hz). A magnitude do *masking release* foi obtida para cada modulação e foi realizada a análise comparativa dos resultados. **Resultados:** Os achados demonstraram melhores limiares de reconhecimento de sentenças quando o ruído mascarante foi modulado em 4, 8, 16 e 32 Hz e piores limiares quando o ruído mascarante estava estável e em 64 Hz. No que diz respeito à análise da relação sinal/ruído, foram observados, no presente estudo, maiores valores para as tarefas que envolvem reconhecimento de sentenças com ruído estável, seguidos das tarefas que envolvem reconhecimento de sentenças com ruído modulado em 64 Hz, e menores valores para as tarefas que envolvem reconhecimento de sentenças com ruído modulado em 32, 16, 8 e 4 Hz, respectivamente. **Conclusão:** A magnitude do *masking release* para sentenças não se diferencia com taxas de modulação em amplitude entre 4 e 32 Hz. No entanto, quando a taxa de modulação é elevada a 64 Hz, a magnitude do *masking release* diminui.

Descritores: Audição; Mascaramento perceptivo; Percepção da fala; Testes de discriminação da fala; Ruído

ABSTRACT

Purpose: To investigate the effect of modulation rate on masking release for speech. **Methods:** Fifteen normal hearing subjects participated in the study. They were tested on speech perception in noise using the sentences of the HINT-Brazil. Speech recognition thresholds were obtained in the presence of steady masking and amplitude modulated masking at rates of 4, 8, 16, 32 and 64 Hz. The magnitude of masking release was obtained for each modulation rate, and results were compared. **Results:** Data showed significantly lower speech thresholds when the masking noise was amplitude modulated at rates of 4, 8, 16 and 32 Hz, when compared to thresholds obtained in steady noise or noise modulated at 64 Hz. Speech-to-masker ratios (SMR) were higher in steady masking noise, followed by modulated noise at 64 Hz, and lower for modulated rates of 32, 16, 8 and 4 Hz, respectively. **Conclusion:** The magnitude of masking release for speech does not seem to differ significantly among modulation rates of 4 to 32 Hz. However, for a modulation rate of 64 Hz the magnitude of masking release is significantly reduced.

Keywords: Hearing; Perceptual masking; Speech perception; Speech discrimination tests; Noise

Trabalho realizado na Universidade Federal de Pernambuco – UFPE – Recife (PE), Brasil.

(1) Programa de Pós-Graduação (Doutorado) em Neuropsiquiatria e Ciências do Comportamento, Universidade Federal de Pernambuco – UFPE – Recife (PE), Brasil.

(2) Departamento de Fonoaudiologia, Universidade Federal de Pernambuco – UFPE – Recife (PE), Brasil.

(3) Programa de Pós-Graduação (Mestrado) em Neuropsiquiatria e Ciências do Comportamento, Universidade Federal de Pernambuco – UFPE – Recife (PE), Brasil.

Conflito de interesses: Não

Contribuição dos autores: *KPA* pesquisadora principal, elaboração da pesquisa, elaboração do cronograma, levantamento da literatura, coleta e análise dos dados, redação do artigo, submissão e trâmites do artigo; *DCM* orientadora, elaboração da pesquisa, elaboração do cronograma, análise dos dados, correção da redação do artigo, aprovação da versão final; *FAP* pesquisador colaborador, levantamento da literatura, coleta e análise dos dados, redação e revisão do artigo final; *SMSG* coorientadora, elaboração da pesquisa, elaboração do cronograma, análise dos dados, correção da redação do artigo, aprovação da versão final.

Endereço para correspondência: Karina Paes Advíncula. R. Amaro Coutinho, 531/701, Encruzilhada, Recife (PE), Brasil, CEP: 52041-110.

E-mail: kpadvincula@hotmail.com

Recebido em: 3/1/2013; **Aceito em:** 26/8/2013

INTRODUÇÃO

A habilidade em compreender a fala pode ser considerada um dos mais importantes aspectos a ser avaliado, quando se trata dos distúrbios da comunicação humana. As várias situações de escuta diária requerem que o ouvinte perceba a informação de fala que, muitas vezes, encontra-se mascarada por um ruído de fundo - ruído competitivo simultâneo^(1,2). Esse ruído competitivo se origina em diferentes fontes em um ambiente social e o ouvinte deve ‘separar’ a fala alvo dessa mensagem competitiva^(3,4). Ouvir a fala na presença de ruído competitivo é uma atividade que tem sido motivo de estudo há vários anos^(5,6). Como exemplo, pode-se citar o *Articulation Index* (AI), um teste utilizado em estudos de telefonia, baseado em uma série de experimentos com ruídos mascarantes. Esse teste é capaz de prever índices de inteligibilidade da fala através da análise do espectro da fala e do mascaramento^(7,8).

Por décadas, estudos têm comparado a habilidade de percepção da fala alvo em presença de ruído de fundo com diferentes características físicas⁽¹⁻¹²⁾. Quando uma modificação acústica realizada no ruído (geralmente uma modulação em amplitude ao longo do tempo, ou uma modulação em espectro de frequência) provoca melhor reconhecimento da fala alvo, diz-se que aconteceu o *masking release*. Esse é o termo da língua inglesa que define a melhora no desempenho de reconhecimento da fala em presença de ruído mascarante, quando este foi acusticamente modificado. Uma sugestão de tradução que propomos para esse termo é: *benefício causado pela modificação do ruído mascarante*. No entanto, no presente estudo, e para fins didáticos, utilizamos o termo na língua inglesa.

Estudos têm demonstrado que o desempenho no reconhecimento da fala alvo, para ouvintes com audição normal, é melhor quando o ruído oscila em amplitude ao longo do tempo (*amplitude modulated* – AM), ou em espectro de frequência (*frequency modulated* – FM), ou seja, a percepção da fala é otimizada quando o ruído é modulado, comparando-se ao ruído estável^(12,13). No caso da modulação em amplitude ao longo do tempo, o *masking release* acontece devido à redução dos níveis de intensidade do mascaramento no momento da modulação (modulações mínimas), melhorando, assim, a relação sinal/ruído^(13,14). Essa situação permite que o ouvinte perceba melhor o sinal alvo, exatamente no momento da modulação mínima, quando o ruído mascarante está com sua intensidade mais fraca, ou seja, o ouvinte percebe breves segmentos da fala no momento em que a modulação ocorre e obtém mais informações para reconhecer o que está sendo dito.

A magnitude do *masking release* depende de algumas características físicas do ruído mascarante, como por exemplo, a sua taxa de modulação^(10,13-18). Um estudo⁽¹⁸⁾ encontrou variação na melhora do reconhecimento da fala entre 15 e 25 dB para taxas de modulações entre 8 e 20 Hz. Tipicamente, tem sido observada maior magnitude do *masking release* para taxas de modulações mais baixas⁽¹⁹⁾, tais como modulações entre

8 Hz e 25 Hz (especialmente em 10 Hz)^(13,14,20). Isso porque frequências de modulações mais baixas produzem maiores espaços temporais de menor amplitude. Dessa forma, facilitam a percepção da fala⁽²¹⁾.

A magnitude do *masking release* é tipicamente observada em uma ou outra das seguintes situações: a) melhora, em valores percentuais, do reconhecimento da fala em uma determinada relação sinal/ruído (RSR), em presença de ruído modulado (quando comparado com ruído estável); b) diminuição do valor da RSR em situação de ruído competitivo modulado (quando comparado com ruído estável), em um determinado percentual de acerto no reconhecimento da fala⁽²²⁾. A mensuração do *masking release* para fala é obtida por meio da utilização de material linguístico (listas de palavras e/ou sentenças) de testes de fala, em presença de ruído mascarante. Dentre esses testes, encontra-se o *Hearing in Noise Test* (HINT)⁽²³⁾, que investiga aspectos temporais da audição, mais especificamente, o mascaramento temporal⁽²⁴⁾.

O HINT é um teste que faz uso de sentenças para medir o limiar de reconhecimento de fala em presença de ruído mascarante. É composto por 25 listas (dez sentenças cada) foneticamente balanceadas e apresentadas em presença concomitante de ruído, para obtenção dos limiares de reconhecimento de fala. Em 2008, um grupo de pesquisadores⁽²⁵⁾ realizou um procedimento similar, utilizando listas foneticamente balanceadas em português, e padronizou o HINT-Brasil.

O HINT-Brasil é composto por 12 listas de 20 sentenças no Português brasileiro e pode ser considerado um instrumento adequado para o estudo da habilidade de reconhecer a fala em presença de ruído competitivo. No entanto, o ruído apresentado no teste encontra-se estável, ou seja, constante em intensidade (amplitude), ao longo do tempo. Através de um processador acústico (ex. *RX6 Turker Davis Technology*), é possível modificar as características desse ruído, tornando-o modulado em amplitude, em diferentes taxas de modulações. Dessa forma, pode-se fazer uso das sentenças do HINT-Brasil para investigação de limiares de reconhecimento de fala em presença de ruído estável e em presença de ruído modulado, em diferentes taxas de modulações.

Com o uso das sentenças do HINT-Brasil, o presente estudo investigou o efeito das diferentes taxas de modulações do mascaramento na magnitude do *masking release*, através da pesquisa do limiar de reconhecimento de sentenças em uma população de adultos com audição normal.

MÉTODOS

Trata-se de estudo de caráter quantitativo, transversal e observacional, realizado no Laboratório de Pesquisa do Departamento de Fonoaudiologia da Universidade Federal de Pernambuco (UFPE), durante o período de julho a agosto de 2012. Este estudo faz parte da pesquisa intitulada *Temporal masking and speech recognition in the aging auditory system*:

US-BRAZIL, financiada pelo *National Institute of Health* (NIH), em parceria com a Universidade da Carolina do Norte, aprovada por Comitê de Ética americano, *Institucional Review Board* (IRB), sob o número 11-1113 e pelo Comitê de Ética em Pesquisa envolvendo Seres Humanos, do Centro de Ciências da Saúde da Universidade Federal de Pernambuco CEP/CCS-UFPE, sob o número 233/2012.

A amostra foi composta por 15 participantes, recrutados por conveniência, sendo nove do gênero feminino e seis do gênero masculino. A faixa etária dos participantes variou entre 17 e 35 anos (média idade de 21 anos). Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

O estudo teve como critérios de inclusão: ausência de queixa, histórico ou diagnóstico de alterações otológica, neurológica e/ou psiquiátrica; meatoscopia sem alterações; audição normal na orelha testada - limiares auditivos tonais de até 20 dB NA nas frequências de oitavas de 250 Hz a 8000 Hz, incluindo as frequências interoitavas de 3000 Hz e 6000 Hz⁽²⁶⁾.

Os participantes foram submetidos ao teste de reconhecimento de sentença na presença de ruído, utilizando as 12 listas do HINT-Brasil, cada uma contendo 20 sentenças, gravadas por um locutor nativo do gênero masculino. As sentenças foram originalmente gravadas no *House Research Institute* (HRI), USA⁽²⁵⁾.

Por meio do processador acústico RX6 (*Tucker-Davis Technologies*), o ruído foi modulado por uma onda quadrática, variando em intensidade entre 65 e 30 dB NPS, nas diferentes taxas de modulação (4, 8, 16, 32 e 64 Hz). O ruído estável foi apresentado a 65 dB NPS. O ruído utilizado possui o envelope de espectro de frequências que se assemelha ao envelope do espectro de frequências das sentenças utilizadas no teste (*speech-shaped noise*).

Os participantes foram posicionados dentro de uma cabina acústica e orientados a escutar e repetir as sentenças alvo, na presença do ruído de fundo competitivo, exatamente da forma como escutaram. O examinador, posicionado fora da cabina acústica, monitorou e registrou as respostas por meio do *software* Matlab (Matrix Laboratory®), versão R2012a.

O sinal de fala e o ruído de fundo competitivo foram enviados via plataforma de processamento digital (RX6, Tucker-Davis Technologies®) e apresentados monoauralmente, por meio de fone auditivo Sennheiser HD580, à orelha direita ou à melhor orelha, quando os limiares obtidos entre as orelhas variaram mais que 5 dB.

A resposta foi considerada correta quando a sentença foi repetida na íntegra. Qualquer diferença entre a sentença do teste e a sentença emitida pelo participante, foi computada como erro. Por exemplo, alterações na utilização de artigos, conjugação verbal e inclusão ou omissões de palavras, mesmo que não modificassem o significado original, foram consideradas como erros.

Cada participante foi submetido a duas, das seis condições de mascaramento (ruído em 0 Hz, 4 Hz, 8 Hz, 16 Hz, 32 Hz e

64 Hz). Há que se destacar que a modulação em 0 Hz corresponde ao ruído estável.

Foram formados três grupos, com cinco participantes cada, a saber: G1, para a condição 0 e 4 Hz; G2, para a condição 8 e 16 Hz e G3, para a condição 32 e 64 Hz. O limiar de reconhecimento de sentença de cada participante foi obtido por meio da média de quatro medições e o resultado para cada condição de teste, pela média dos limiares dos cinco participantes.

O *masking release* foi definido através da comparação entre o limiar de reconhecimento de sentenças em presença de ruído estável (tomado como referência) e o limiar de reconhecimento de sentenças em presença de ruído modulado (para as diversas condições de modulação).

O limiar de reconhecimento de sentenças foi determinado por procedimento adaptativo que converge para o ponto 71% da função psicométrica. Em um procedimento adaptativo, a intensidade em que um estímulo é apresentado depende de como o sujeito respondeu aos estímulos anteriores⁽²⁷⁻²⁹⁾.

Para a pesquisa do limiar, a intensidade inicial utilizada foi sempre superior ao limiar de reconhecimento esperado, sendo de 60 dB NPS para todas as taxas de modulação e 70 dB NPS na condição do ruído estável. Adotou-se o método descendente-ascendente transformado (*two down - one up*)⁽²⁸⁾. Nesse método, para cada duas respostas corretas consecutivas, diminui-se a intensidade do sinal em intervalos pré-estabelecidos e, a cada resposta incorreta, o nível da apresentação da sentença seguinte é aumentado. No presente estudo, foi utilizado o intervalo de 2 dB entre as apresentações.

O limiar de reconhecimento das sentenças foi obtido após seis reversões de resposta. Considera-se reversão a mudança no padrão de respostas apresentadas pelo indivíduo. Em outras palavras, um curso descendente começa com uma resposta positiva até que haja uma resposta negativa (reversão) e um curso ascendente começa com uma resposta negativa e termina com uma positiva (reversão).

Cada sentença foi apresentada somente uma única vez para o mesmo participante, a fim de eliminar variáveis relacionadas ao fenômeno de aprendizagem. A ordem de apresentação das diferentes condições de mascaramento, bem como a escolha das listas, ocorreu de forma aleatória.

O período de duração do teste foi de aproximadamente 50 minutos, sendo interrompido algumas vezes, quando os sujeitos referiram cansaço e/ou desconforto, para que tais fatores não interferissem em sua atenção e, por conseguinte, no desempenho do exame.

A análise estatística foi realizada no *software* SPSS, com o objetivo de examinar o efeito das diferentes modulações do mascaramento na magnitude do *masking release* em uma população de adultos com audição normal. A significância desses dados foi avaliada por meio da análise de variância, utilizando-se o modelo linear geral de efeitos mistos, onde a frequência de modulação foi designada como o efeito fixo e os participantes foram designados como o efeito aleatório, contando

com medições repetidas entre participantes e entre pares de condição. Foi utilizado o teste F, com nível de significância de 95%, para verificar o efeito da frequência de modulação. A comparação Pairwise entre as frequências de modulação foi ajustada para comparações múltiplas, através da comparação das médias dos limiares de cada taxa de modulação.

RESULTADOS

As médias dos limiares de reconhecimento de sentenças nas diferentes condições de mascaramento estão apresentados na Figura 1.

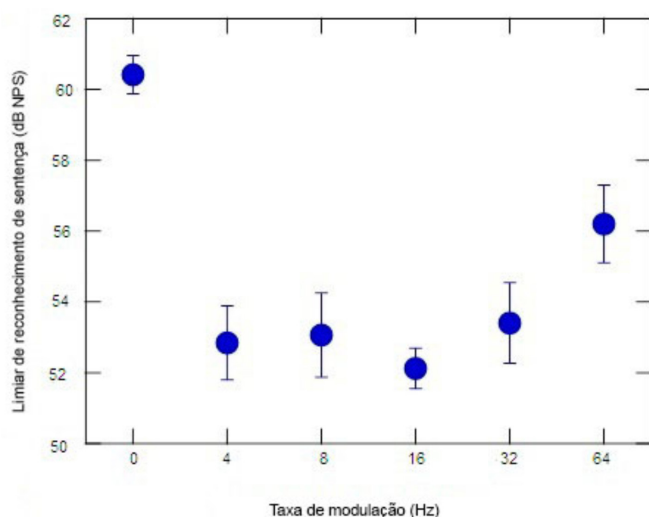


Figura 1. Média dos limiares de reconhecimento de sentença em função da frequência de modulação do ruído

Os limiares de reconhecimento de sentenças foram descritos em função da taxa da modulação do ruído, denominando-se 0 Hz o ruído estável. Observou-se que os limiares de reconhecimento de sentenças em presença de ruído estável (0 Hz) apresentaram valores maiores ($M=60,42$ dB NPS). Com relação aos limiares de reconhecimento de sentenças em presença de ruído modulado, foram obtidas as médias $M=52,84$ dB NPS, $M=53,06$ dB NPS, $M=52,12$ dB NPS e $M=53,40$ dB NPS para a modulação de 4 Hz, 8 Hz, 16 Hz e 32 Hz, respectivamente. Para a modulação em 64 Hz, foram observados valores de limiares maiores do que os das demais modulações ($M=56,20$ dB NPS). No entanto, menores, quando comparados aos valores obtidos em presença de ruído estável (Tabela 1).

Na análise da comparação das médias dos limiares de reconhecimento de sentenças obtidos em presença de ruído modulado, por frequência de modulação, verificamos que o efeito da frequência de modulação foi significativo ($F(5,6.5) = 207.4$; $p < 0,01$) (Tabela 2).

A comparação Pairwise entre as frequências de modulação, ajustada para comparações múltiplas, resultou no seguinte padrão ($p < 0,05$): todos os limiares obtidos em presença de ruído modulado (entre 4 Hz e 64 Hz) foram melhores do que

Tabela 1. Limiares de reconhecimento de sentenças em presença de ruído modulado em diferentes taxas

Taxa de modulação do ruído	Média (dB NPS)	DP	Valor mínimo	Valor máximo
0 Hz	60,42	0,242	59,749	61,091
4 Hz	52,84	0,464	51,551	54,129
8 Hz	53,06	0,534	51,576	54,544
16 Hz	52,12	0,254	51,415	52,825
32 Hz	53,40	0,509	51,987	54,813
64 Hz	56,20	0,492	54,834	57,566

Legenda: DP = desvio-padrão

Tabela 2. Comparação entre os limiares médios de reconhecimento de sentenças obtidos em presença de ruído com diferentes taxas de modulação

Taxa de modulação (Hz)	Taxa de modulação (Hz)	Diferença média	DP	Significância**
0	4	7,58*	0,398	0,000
	8	7,36*	0,587	0,000
	16	8,30*	0,350	0,000
	32	7,02*	0,563	0,000
4	64	4,22*	0,548	0,000
	0	-7,58*	0,398	0,000
	8	-0,22	0,708	0,764
	16	0,72	0,529	0,221
8	32	-0,56	0,689	0,440
	64	-3,36*	0,676	0,001
	0	-7,36*	0,587	0,000
	4	0,22	0,708	0,764
16	64	-3,14*	0,726	0,003
	0	-8,30*	0,350	0,000
	4	-0,72	0,529	0,221
	8	-0,94	0,649	0,221
32	64	-1,28	0,738	0,657
	0	-4,08*	0,554	0,000
	4	0,56	0,689	0,440
	8	0,34	0,738	0,657
64	16	1,28	0,569	0,066
	0	-2,80*	0,742	0,020
	4	-4,22*	0,548	0,000
	8	3,36*	0,676	0,001
	16	3,14*	0,726	0,003
	32	4,08*	0,554	0,000
	64	2,80*	0,742	0,020

* Valores significativos ($p < 0,05$) – Teste F

** ajustado para comparações múltiplas

Legenda: DP = desvio-padrão

os limiares obtidos em presença do ruído estável; os limiares obtidos em presença de ruído modulado entre 4 Hz e 32 Hz não diferiram entre si; os limiares obtidos em presença de ruído modulado em 64 Hz diferiram dos limiares obtidos com uso das demais modulações.

Na análise dos resultados dos limiares de reconhecimento de sentenças em presença de ruído estável e modulado, expressos em função da relação sinal/ruído (RSR) e magnitude do *masking release*, observou-se menor RSR para o ruído estável (0 Hz) e para o ruído modulado, em 64 Hz. Para os ruídos modulados em 4 Hz, 8 Hz, 16 Hz e 32 Hz, a RSR foi maior, porém não houve diferença entre eles (Tabela 3).

Com relação aos resultados do *masking release*, constatamos maiores valores para os ruídos com modulação entre 4 Hz e 32 Hz e menores valores para modulação de 64 Hz. Esse padrão indica a presença do *masking release* em todas as modulações testadas, com magnitude reduzida, entretanto, quando foi utilizada a modulação de 64 Hz, em comparação com as modulações entre 4 Hz e 32 Hz (Tabela 3).

Tabela 3. Valores da relação sinal/ruído (RSR) e do *masking release* em diferentes taxas de modulação do ruído mascarante

Taxa de modulação do ruído (Hz)	RSR (dB)	<i>Masking release</i> (dB)
0	- 4,58	--
4	- 12,16	7,58
8	- 11,95	7,37
16	- 12,88	8,20
32	- 11,60	7,02
64	-8,8	4,22

DISCUSSÃO

Sabe-se que a diferença no desempenho do reconhecimento da fala em presença de ruído varia de acordo com a natureza do ruído mascarante^(2,9,13,14). No que se refere à modulação do ruído, esta pode ser de espectro de frequência ou de amplitude em função do tempo^(12,14,21,24).

Nos ruídos com modulação de espectro de frequência, o que ocorre é que o espectro da fala do ruído oscila e, em alguns momentos, diverge do espectro de frequência da fala alvo. Embora alguns segmentos da fala alvo possam ser completamente mascarados pelo ruído modulado em frequência, outros não são. Os segmentos não mascarados são mais facilmente percebidos pelo ouvinte e ajudam a melhorar o desempenho no reconhecimento da fala alvo. Existe, portanto, a diminuição no valor da relação sinal/ruído, ou seja, o limiar de reconhecimento da fala em presença de ruído modulado diminui, quando comparado ao limiar de reconhecimento da fala em presença de ruído estável⁽²⁴⁾.

Quando a modulação é na amplitude do ruído ao longo do tempo, a intensidade do ruído é que oscila, permitindo que o

ouvinte perceba mais informações da fala alvo nos momentos de intensidades mais baixas. Mais uma vez, o limiar de reconhecimento da fala diminui, quando comparado ao limiar determinado em presença de ruído estável. Como mencionado anteriormente, essa melhora no desempenho do reconhecimento da fala em presença de ruídos modificados, em sua natureza, é chamada de *masking release*, na língua inglesa^(10,13-21,23).

A modulação do ruído pode ainda divergir em relação a seu “formato”. Pesquisas vêm utilizando modulação sinusoidal e modulação quadrática. Estudos mostram que, nos casos em que a modulação é sinusoidal, há menos aproveitamento da diminuição do nível da amplitude do ruído, fazendo com que o ouvinte perceba menos pistas acústicas da fala alvo. Nos casos em que a modulação é quadrática, o ouvinte consegue perceber maior quantidade de pistas de fala e apresentar melhor desempenho no reconhecimento da fala^(12-14,21,24).

Neste estudo, foi utilizado o ruído modulado em amplitude e de formato quadrático (*squared-shaped noise*). Os resultados aqui encontrados, quando comparados aos resultados de um estudo que utilizou ruído modulado sinusoidalmente em 10 Hz⁽²⁾, mostraram maior magnitude do *masking release* para o ruído quadrático. O estudo de Hall e colaboradores mostra a magnitude do *masking release* em torno de 5 dB e a magnitude aqui encontrada foi em torno de 7/8 dB.

Um estudo⁽¹⁾ demonstrou benefícios de 4 a 8 dB para tarefas de reconhecimento de sentenças, em ouvintes com audição normal. Em outro estudo⁽²²⁾, observou-se melhora no reconhecimento de sentenças para ouvintes normais, com limiares que variaram de 6 a 10 dB. Dois estudos, em 2006^(17,18), relataram magnitudes do *masking release* entre 15 e 25 dB. É importante salientar que a magnitude do benefício do mascaramento modificado é também indicada pelo aumento da relação sinal/ruído⁽²²⁾ (Tabela 3).

Neste estudo, ao se investigar o efeito da taxa de modulação na magnitude do *masking release*, observou-se maior magnitude para modulações entre 4 Hz e 32 Hz, demonstrando tendência ao melhor desempenho de reconhecimento de sentenças em presença de ruídos mascarantes com menores taxas de modulação. Esses resultados são consistentes aos descritos na literatura, desde os primeiros estudos na área⁽³⁰⁾ até estudos mais recentes⁽¹³⁾.

Uma explicação para esses achados é que ruídos com alta taxa de modulação se assemelham, em características perceptuais, a ruídos estáveis, pois o tempo em que o ouvinte pode se beneficiar da menor intensidade do ruído é curto, dificultando a percepção das pistas acústicas de fala⁽²⁰⁾. Seguindo esse raciocínio, quanto maior for a taxa de modulação, menos pistas acústicas são percebidas pelo ouvinte, pois os intervalos de tempo em intensidade mínima são menores. Consequentemente, a magnitude do *masking release* diminui.

Um estudo desenvolvido em 1950, já apresentou melhor desempenho no reconhecimento da fala em presença de ruído mascarante com taxa baixa de modulação – 10 Hz⁽³⁰⁾. Os autores

observaram que os limiares de reconhecimento de fala são semelhantes para ruído mascarante estável e ruídos mascarantes com alta taxa de modulação. Os resultados do presente estudo concordam com o padrão descrito no estudo mencionado anteriormente⁽³⁰⁾, ou seja, para o ruído mascarante estável e para o ruído mascarante com modulação de 64 Hz, observou-se que os limiares de reconhecimento de fala foram maiores do que ruídos com outras modulações.

Os resultados obtidos com uso de ruído com modulação de 4 Hz, 8 Hz, 16 Hz e 32 Hz, não tiveram diferença significativa entre si e mostraram-se semelhantes ao que se tem observado quando se utiliza a modulação de 10 Hz^(2,8,13,19). No que diz respeito à análise da relação sinal/ruído (Tabela 3), constatou-se que a RSR pode ser utilizada para se observar e comparar a magnitude do *masking release*.

CONCLUSÃO

Os achados do estudo são pioneiros no que diz respeito ao reconhecimento de sentenças do HINT-Brasil em presença de ruído com diferentes modulações de amplitude. Com este estudo foi possível concluir que a magnitude do *masking release* para sentenças não se diferencia com taxas de modulação em amplitude entre 4 Hz e 32 Hz. No entanto, quando a taxa é elevada a 64 Hz, a magnitude do *masking release* diminui.

AGRADECIMENTOS

Agradecemos o auxílio do Professor Dr. John H. Grose, da Universidade da Carolina do Norte em Chapel Hill. Essa pesquisa foi financiada por parte do fomento concedido pelo Instituto Nacional de Saúde nos Estados Unidos (*U.S. National Institutes of Health* - R03-DC012278) ao Dr. John H. Grose.

REFERÊNCIAS

1. Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*. 1990;88(4):1725-36.
2. Hall JW, Buss E, Grose JH, Roush PA. Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear Hear*. 2012;33(3):340-8.
3. Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am*. 2001;109(3):1101-9.
4. Brungart DS, Simpson BD, Scott KR, Ericson MA. Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am*. 2001;110(5):2527-38.
5. Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J Acoust Soc Am*. 2004;115(5):2246-56.
6. Jacob RTS, Monteiro NFG, Molina SV, Bevilacqua MC, Lauris JRP, Moret ALM. Percepção da fala em crianças em situação de ruído. *Arq Int Otorrinolaringol*. 2011;15(2):163-7.
7. Bernstein JG, Grant KW. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2009;125(5):3358-72.
8. Buss E, Hall JW3rd, Grose JH. Spectral integration of synchronous and asynchronous cues to consonant identification. *J Acoust Soc Am*. 2004;115(5):2278-85.
9. Gustafsson HA, Arlinger SD. Masking of speech by amplitude-modulated noise. *J Acoust Soc Am*. 1994;95(1):518-29.
10. Nelson PB, Jin SH, Carney AE, Nelson DA. Understanding speech in modulated interference: cochlear implant users and normal hearing listeners. *J Acoust Soc Am*. 2003;113(2):961-8.
11. Anderson ES, Nelson DA, Kreft H, Nelson PB, Oxenham AJ. Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. *J Acoust Soc Am*. 2011;130(1):364-75.
12. Buss E, Whittle LN, Grose JH, Hall JW3rd. Masking release for words in amplitude-modulated noise as a function of modulation rate and task. *J Acoust Soc Am*. 2009;126(1):269-80.
13. Buss E, He S, Grose JH, Hall JW3rd. The monoaural temporal window based on masking period pattern data in school-aged children and adults. *J Acoust Soc Am*. 2013;133(3):1586-97.
14. Füllgrabe C, Berthommier F, Lorenzi C. Masking release for consonant features in temporally fluctuating background noise. *Hear Res*. 2006;211(1-2):74-84.
15. Stuart A, Phillips DP. Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. *Ear Hear*. 1996;17(6):478-89.
16. Summers V, Molis MR. Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. *J Speech Lang Hear Res*. 2004;47(2):245-56.
17. Rhebergen KS, Versfeld NJ, Dreschler WA. Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. *J Acoust Soc Am*. 2006;120(6):3988-97.
18. George EL, Festen JM, Houtgast T. Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2006;120(4):2295-311.
19. Desloge JG, Reed CM, Braida LD, Perez ZD, Delhorne LA. Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise. *J Acoust Soc Am*. 2010;128(1):342-59.
20. Lorenzi C, Husson M, Ardoint M, Debruelle X. Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception. *Int J Audiol*. 2006;45(9):487-95.
21. Sheff S, Yost WA. Discrimination of start phase with sinusoidal envelope modulation. *J Acoustic Soc Am*. 2007;121(2):84-9.
22. Kwon BJ, Perry TT, Wilhelm CL, Healy, EW. Sentence recognition in noise promoting or suppressing masking release by normal-hearing and cochlear-implant listeners. *J Acoust Soc Am*. 2012;131(4):3111-9.

23. Nilsson M, Soli SD, Sullivan, JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am.* 1994;95(2):1085-99.
24. Grose JH, Mamo SK, Hall JW3rd. Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear.* 2009;30(5):568-75.
25. Bevilacqua MC, Banhara MR, Da Costa EA, Vignoly AB, Alvarenga KF. The Brazilian Portuguese hearing in noise test. *Int J Audiol.* 2008;47(6):364-5.
26. American National Institute (ANSI). (2004). American National Standard Specification for Audiometers. ANSI S3.6-2004. New York: ANSI, 2004.
27. Wetherill GB, Levitt H. Sequential estimation of points a psychometric function. *Br J Math Stat Psychol.* 1965;18:1-10.
28. Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am.* 1971;49(2):467-77.
29. Bode DL, Carhart R. Measurements of articulation functions using adaptive test procedures. *IEE Trans Audiol Electroacoustic.* 1973;21:196-201.
30. Miller GA, Licklider JCR. The intelligibility of interrupted speech. *J Acoust Soc Am.* 1950;22(2):167-73.