

FERNANDO AUGUSTO PACÍFICO

**EFEITO DE DIFERENTES SITUAÇÕES DE ESCUTA NO
MASKING RELEASE PARA A FALA**

RECIFE

2014

FERNANDO AUGUSTO PACÍFICO

**EFEITO DE DIFERENTES SITUAÇÕES DE ESCUTA NO
MASKING RELEASE PARA A FALA**

Dissertação aprovada pelo colegiado do Programa de Pós-graduação em Neuropsiquiatria e Ciências do Comportamento do Centro de Ciências da Saúde da Universidade Federal de Pernambuco, área de neurociências, como requisito para obtenção do grau de Mestre.

Orientador: Profa. Dra. Maria Lúcia Gurgel da Costa

Coorientadora: Profa. Dra. Silvana Maria Sobral Griz

RECIFE

2014

Ficha catalográfica elaborada pela
Bibliotecária: Mônica Uchôa, CRB4-1010

P117e Pacífico, Fernando Augusto.
Efeito de diferentes situações de escuta no masking release para a
fala / Fernando Augusto Pacífico. – Recife: O autor, 2014.
115 f.: il.; 30 cm.

Orientadora: Maria Lúcia Gurgel da Costa.
Dissertação (mestrado) – Universidade Federal de Pernambuco,
CCS. Programa de Pós-Graduação em Neuropsiquiatria e Ciências do
Comportamento, 2014.
Inclui referências, apêndices e anexos.

1. Mascaramento perceptivo. 2. Psicometria. 3. Percepção da fala. 4.
Ruído. I. Costa, Maria Lúcia Gurgel da (Orientadora). II. Título.

612.665 CDD (22.ed.) UFPE (CCS2014-034)

Pró-Reitoria para Assuntos de Pesquisa e Pós-Graduação

Centro de Ciências da Saúde

Programa de Pós-Graduação em Neuropsiquiatria

e Ciências do Comportamento

**RELATÓRIO DA BANCA EXAMINADORA DA DEFESA DE DISSERTAÇÃO DO
MESTRANDO FERNANDO AUGUSTO PACÍFICO**

No dia 28 de fevereiro de 2014, às 9h, no Auditório do Prof^o Paulo Rosas na ADUFEPE da Universidade Federal de Pernambuco, os Professores: Cleide Fernandes Teixeira, Doutora Professora do Departamento de Fonoaudiologia da Universidade Federal de Pernambuco; Otávio Gomes Lins, Doutor Professor do Departamento de Neuropsiquiatria da Universidade Federal de Pernambuco e Marcelo Moraes Valença, Doutor Professor do Departamento de Neuropsiquiatria da Universidade Federal de Pernambuco, componentes da Banca Examinadora, em sessão pública, arguiram o Mestrando FERNANDO AUGUSTO PACÍFICO, sobre a sua Dissertação intitulada “**EFEITO DE DIFERENTES SITUAÇÕES DE ESCUTA NO MASKING RELEASE PARA FALA**”, orientado pela Professora Dr^a. Maria Lúcia Gurgel da Costa. Ao final da arguição de cada membro da Banca Examinadora e resposta do Mestrando, as seguintes menções foram publicamente fornecidas:

Prof^a. Dr^a. Cleide Fernandes Teixeira

Prof. Dr. Otávio Gomes Lins

Prof. Dr. Marcelo Moraes Valença

Prof^a. Dr^a. Cleide Fernandes Teixeira

Prof. Dr. Otávio Gomes Lins

Prof. Dr. Marcelo Moraes Valença
Presidente da Banca Examinadora

UNIVERSIDADE FEDERAL DE PERNAMBUCO

REITOR

Prof. Anísio Brasileiro de Freitas Dourado

VICE-REITOR

Prof. Silvio Romero Marques

PRÓ-REITOR PARA ASSUNTOS DE PESQUISA E PÓS-GRADUAÇÃO

Prof. Francisco Ramos

CENTRO DE CIÊNCIAS DA SAÚDE

DIRETOR

Prof. Nicodemos Teles Pontes Filho

DEPARTAMENTO DE NEUROPSIQUIATRIA

DIRETOR

Prof. José Francisco Albuquerque

**PROGRAMA DE PÓS-GRADUAÇÃO EM NEUROPSIQUIATRIA E CIÊNCIAS
DO COMPORTAMENTO**

COORDENADOR

Profa. Marcelo Moraes Valença

VICE-COORDENADORA

Profa. Sandra Lopes de Souza

CORPO DOCENTE

Profa. Ângela Amâncio dos Santos

Profa. Ana Elisa Toscano

Prof. Amaury Cantilino

Profa. Belmira Lara da S.A.da Costa

Profa. Claudia Jacques Lagranha

Prof. Everton Botelho Sougey

Prof. Gilson Edmar Gonçalves e Silva

Prof. Hilton Justino da Silva

Prof. Hildo Rocha Cirne de Azevedo Filho

Prof. Hugo André de Lima Martins

Prof. João Ricardo Mendes de Oliveira

João Henrique da Costa Silva

Profa. Kátia Karina Monte Silva

Profa. Luciana Patrícia A. de Andrade Valença

Prof. Lúcio Vilar Rabelo Filho

Prof. Luiz Ataíde Junior

Prof. Marcelo Moraes Valença

Prof. Marcelo Cairrão Araújo Rodrigues

Profa. Maria Lúcia de Bustamente Simas

Profa. Maria Lúcia Gurgel da Costa

Prof. Murilo Costa Lima

Prof. Otávio Gomes Lins

Prof. Othon Coelho Bastos Filho

Profa. Paula Rejane Beserra Diniz

Prof. Pedro Augusto Sampaio Rocha Filho

Prof. Raul Manhães de Castro

Profa. Rosana Christine C. Ximenes

Profa. Sandra Lopes de Souza

Profa. Sílvia Regina de Arruda Moraes

A Zelândia, a quem devo minha existência, e a Alice, a quem dedicarei minha existência.

Alfa e Ômega.

AGRADECIMENTOS

A **Deus**, inteligência suprema e causa primeira de todas as coisas.

A **Alice Oliveira Pacífico**, uma benção que o céu me reserva ao coração. Será meu amanhecer e meu anoitecer. Empregarei todas minhas forças em desenvolver sua inteligência e moralidade.

A minha mãe, **Zelândia M^a de Souza Leandro Pacífico**, que abdicando de todo ato egoístico, optou por empregar todas suas forças durante sua existência em prol da felicidade daquele que a vida enviou aos seus cuidados. Serei eternamente grato por cada segundo que me dedicaste e espero um dia conseguir retribuir todo seu amor. És o maior exemplo de renúncia que presenciei nesta vida.

Ao meu pai, **Sérgio Félix Pacífico**, que em nada tendo, tudo me forneceu dentro de suas limitações. Ensinou-me que o trabalho é a escada divina de acesso aos lauréis imarcescíveis do espírito. Espero um dia também poder retribuir tamanha dedicação e renúncia.

Ao meu irmão, **Sérgio Félix Pacífico Jr.**, que tanto tem me ensinado nesta vida. Tens me ensinado a faculdade de compreender, amaparar e acima de tudo amar.

A **Bárbara Soledá Gonçalves de Oliveira**, anjo iluminado que a misericórdia divina enviou para amparar-me. Você constitui para mim a personificação do meu idealismo humano, a realidade viva e insofismável de meus sonhos mais íntimos. Tenho a convicção de haurir em sua convivência um salutar estímulo às minhas energias morais.

A minha orientadora Profa. Dra. **Maria Lúcia Gurgel da Costa**, por ter me acolhido em mais uma jornada. Com sua invejável polivalência que sempre admirei, me ensinou que para a ciência não existe fronteiras. Obrigado por toda compreensão.

A minha coorientadora Profa. Dra. **Silvana Maria Sobral Griz** por sua intercessão decisiva em meu percurso como pesquisador. Mais uma vez permitiste ampliar minha visão científica. Grande virtude é verdadeiramente necessária, para que alguém renuncie à sua personalidade em proveito dos outros. Obrigado por seu auxílio mesmo diante das adversidades. Obrigado Sil.

A Profa. Dra. **Denise Costas Menezes** por sempre tornar o que é complexo em algo compreensível. Agradeço de coração todo esforço e comprometimento empregado em todas as etapas da pesquisa. Sem dúvidas alguma, sem sua intervenção seria impossível realizar este estudo com esse grau de exigência. Muito obrigado mesmo Denis.

A Profa. **Karina Paes Advíncula**, por partilhar todos os momentos desta pesquisa. Suas colaborações na captação dos sujeitos, seu auxílio na coleta dos dados, suas sugestões, e principalmente nossas discussões foram cruciais em todas as etapas desta pesquisa. Ter conhecido você e sua família foi uma das maiores conquistas nestes últimos dois anos. Muito aprendi e muito agradeço por tudo. Muito obrigado Karis.

A Dr. **John H. Grose** um dos maiores pesquisadores que tive a honra de conhecer. Aos meus olhos um modelo de comprometimento, ética e solidariedade. Alicerce principal da pesquisa, sem sua colaboração nada seria possível. Obrigado pelas ideias, sugestões, discussões, enfim, por todo tempo empregado em nossas orientações. Considero seu auxílio um ato devocional na tentativa melhorar a qualidade da pesquisa em países em desenvolvimento, e consecutivamente tornar o mundo menos desigual.

Ao nosso Grupo de Pesquisa em Audiologia pela brilhante parceria construída. Difícil seria explicar em palavras as dificuldades que enfrentamos, no entanto, foram estas que nos tornaram fortes e resilientes. Para mim foi um orgulho ter partilhado cada momento ao lado vocês.

A construção de uma bela obra requer obrigatoriamente momentos de descontração. Agradeço aos meus amigos **Aníbal Sobrinho, Bernardo Kiryllos, Bruno Henrique, Bruno Matos, Camille Siqueira, Cláudio Marcondes, Daniel Tabosa, Eduardo Melo, Ewerton Rodrigues, Felipe Holanda, Geimson Farias, Hanna Raia, Hugo Holanda, Isis Albuquerque, Mohamed Júnior, Rodrigo Raony, Shayna Rodrigues, Thales Lima, Willian Olivares** e tantos outros que acompanharam minha história desde muito antes dela se concretizar. Obrigado por compreender minha ausência e por tornar meus momentos de lazer mais felizes.

Aos funcionários **Alexsandre Cavalcante, Bruno Pereira, Conceição Matoso, Danylo Bruno, Gilberto Sousa Filho, Marilene Costa, João** e demais colegas de trabalho que me acompanharam, compreenderam minhas dificuldades e que acima de tudo sempre me incentivaram nesta importante etapa de minha vida. Agradeço a todos por tornar meus dias melhores.

A minha turma da posneuro, que se tornou uma segunda família. Orgulho-me em dizer que partilhei momentos com cada uma de vocês.

Aos participantes da pesquisa que pela paciência e vontade de ajudar, tanto contribuíram para o conhecimento científico. Sem vocês não haveria ciência.

Ao *National Institute of Health* pelo financiamento de nossa pesquisa, que possibilitou a aquisição dos equipamentos, tornando nosso laboratório uma dos mais bem equipados do país.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudo concedida.

“Feliz de ti se já compreendes o valor do silêncio”

Emmanuel

RESUMO

Embora muito se tenha estudado sobre o *masking release* ou benefício causado pela modificação do mascaramento modulado, estudos recentes tem apontado a relação de dependência entre este e a relação fala-ruído. Dessa forma, o objetivo principal desta dissertação foi investigar o *masking release* em diferentes situações de escuta com o material de fala do teste HINT versão do Português Brasileiro em adultos jovens com audição normal. Foi realizada no período entre julho de 2012 a julho de 2013, no Laboratório de Audiologia da Clínica Escola do Curso de Fonoaudiologia da Universidade Federal de Pernambuco (UFPE), situado no Hospital das Clínicas da UFPE. A pesquisa foi composta por 55 adultos nativos do Português do Brasil com limiars auditivos tonais ≤ 20 dB NA, com idade entre 17 a 35 anos, de ambos os sexos. O primeiro experimento verificou o efeito de diferentes frequências de modulação do mascaramento no *masking release* para a fala. Os resultados mostraram a magnitude do *masking release* para sentenças do HINT na versão do Português Brasileiro 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. O segundo experimento mediu a função desempenho-intensidade em presença de ruído estável (com espectro de fala) e modulado (por uma onda quadrática a uma taxa de 10 Hz). Os resultados mostraram curvas psicométrica menos íngremes quando o ruído era modulado. O terceiro experimento mediu as relações fala-ruído em função da compressão da fala em tempo (0%, 33% e 50% de compressão) em ambos os tipos de ruído. Os limiares aumentaram com o aumento da compressão da fala em tempo em ambas as condições, no entanto, foi mais acentuado para o ruído modulado. Portanto, o benefício do mascaramento modulado diminuiu com o aumento da compressão temporal da fala. Este padrão de resultado sustenta a hipótese de que, em condições de teste, a redução do benefício do mascaramento modulado pode ser devido a RFR elevadas.

Palavras-chave: Mascaramento perceptivo. Função psicométrica. Percepção de fala. Compressão temporal da fala. Ruído.

ABSTRACT

Although much has been studied about the masking release or benefit caused by the modification of the modulated masking, recent studies have pointed to the dependency relationship between this and a speech-noise. Thus, the goal of this thesis was to investigate the masking release in different listening situations with speech material HINT test version of Brazilian Portuguese in young adults with normal hearing. Was performed between July 2012 to July 2013, the Laboratory of Clinical Audiology, Federal University of Pernambuco (UFPE), located at the Hospital of UFPE . The survey consisted of 55 adult native Portuguese of Brazil with hearing limiars Tonai ≤ 20 dB HL, aged 17-35 years, of both sexes. The first experiment evaluated the effect of different frequencies of modulation of the masking release for speech masking. The results showed the magnitude of masking release for the HINT sentences in Brazilian Portuguese version is no different with exchange modulation amplitude between 4 Hz and 32 Hz However, when the rate is increased to 64 Hz, the magnitude of masking release decreases. The second experiment measured performance-intensity functions in both a steady speech-shaped noise masker and a 10 Hz square-wave modulated masker. The results showed that the slope of the performance-intensity function was shallower in the modulated masker than in the steady masker. The third experiment measured speech-to-masker ratios at threshold as a function of time-compression of the speech (0, 33, and 50%) in both the steady and modulated maskers. Thresholds increased with increasing time-compression in both maskers, but more markedly in the modulated masker. The derived modulation masking release therefore declined with increasing time-compression. This pattern of results supports the hypothesis that, under some conditions, failure to benefit from masker modulation could be due to elevated speech-to masker ratios at threshold.

Keywords: Perceptual masking. Psychometric function. Speech perception. Time compressed speech. Noise.

LISTA DE FIGURAS

Artigo Original 2

Figura 1. Percentual de palavras reconhecidas corretamente em função das relações fala-ruído medido em um ruído estável (círculos vazios) e um ruído modulado (quadrados cheios). Cada painel mostra os dados a partir de um sujeito, com o último painel apresenta os dados médios. Também são mostradas melhores funções logísticas de ajuste.....62

Figura 2. Média das funções psicométricas obtidas a partir dos percentuais de palavras reconhecidas corretamente em função das relações fala-ruído medido em um ruído estável (círculos vazios) e um ruído modulado.....63

Figura 3. Limiares de reconhecimento de fala plotados para cada nível de compressão do tempo. Limiares individuais medidos no ruído estável (círculos cheios) e ruído modulado (círculos vazios) são conectados com uma barra vertical. Limiares médios dos grupos também são mostrados para o ruído estável (quadrados cheios) e ruído modulado (quadrados vazios) com barras de erro 1-DP.....64

Artigo Original 1

Figura 1. Média dos limiares de reconhecimento de sentença em função da frequência de modulação do ruído.....104

LISTA DE TABELAS

Artigo Original 1

Tabela 1. Limiares de reconhecimento de sentenças em presença de ruído modulado em diferentes taxas.....104

Tabela 2. Comparação entre os limiares médios de reconhecimento de sentenças obtidos em presença de ruído com diferentes taxas de modulação.....104

Tabela 3. Valores da relação sinal/ruído (RSR) e do *masking release* em diferentes taxas de modulação do ruído mascarante.....105

LISTA DE ABREVIATURAS E SIGLAS

CONEP	Comitê Nacional de Ética em Pesquisa
CT	Compressão do tempo
IRB	<i>Institutional Review Board</i>
HEI	<i>House Ear Institute</i>
HINT	<i>Hearing in Noise Test</i>
HTD	<i>Hearing Test Device</i>
LRS	Limiares de reconhecimento de fala de sentença
NIH	<i>National Institute of Health</i>
RC	Ruído composto
RD	Fala com ruído a direita
RE	Fala com ruído a esquerda
RFR	Relação fala-ruído
RF	Fala com ruído frontal
S	Fala sem ruído
UFPE	Universidade Federal de Pernambuco
TCLE	Termo de Consentimento Livre e Esclarecido

LISTA DE SÍMBOLOS

dB	Decibel
dB A	Decibel ponderado na curva “A”
dB NA	Decibel Nível de Audição
dB NPS	Decibel Nível de Pressão Sonora
Hz	Hertz

SUMÁRIO

1 APRESENTAÇÃO.....	17
2 INTRODUÇÃO.....	19
3 REVISÃO DE LITERATURA	23
3.1 Testes de Fala no Ruído	23
3.2 Hearing in Noise Test.....	25
3.3 Diferentes Medições do Desempenho no Reconhecimento de Fala	29
3.4 Fatores Relacionados aos Estímulos de Teste que Afetam o <i>Masking Release</i>	32
4 HIPÓTESES.....	36
5 MÉTODO.....	37
6 RESULTADOS	43
7 CONCLUSÕES.....	83
REFERÊNCIAS.....	84
APÊNDICE A - TERMO DE CONSENTIMENT LIVRE E ESCLARECIDO	99
APÊNDICE B - ENTREVISTA AUDIOLÓGICA.....	100
APÊNDICE C – ARTIGO ORIGINAL 1: EFFECT OF MODULATION RATE ON MASKING RELEASE FOR SPEECH.....	101
ANEXO A - Aprovação do Comitê de Ética Americano	108
ANEXO B - Aprovação do Comitê de Ética Nacional 1	109
ANEXO C - Aprovação do Comitê de Ética Nacional 2	110
ANEXO D - NORMAS DO PERÍODICO INTERNATIONAL JOURNAL OF AUDIOLOGY	111
ANEXO E - CERTIFICADO DE APRESENTAÇÃO DE TRABALHO NO 21º CONGRESSO BRASILEIRO E 2º IBERO-AMERICANO DE FONOAUDIOLOGIA	115

1 APRESENTAÇÃO

Esta dissertação está associada à 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. Os estudos realizados nesta área tem procurado avançar o conhecimento sobre como um processo normal de envelhecimento compromete o mascaramento temporal. Particularmente esses estudos estão relacionados ao reconhecimento da fala em ambientes com ruído.

Todos os experimentos desta dissertação foram realizados no Laboratório de Audiologia da Clínica Escola do Curso de Fonoaudiologia da Universidade Federal de Pernambuco (UFPE), situado no Hospital das Clínicas da UFPE.

A presente dissertação foi responsável por fornecer: (i) parâmetros de referência, tais como melhor frequência moduladora a ser empregada na modulação do ruído utilizado; (ii) a validação da utilização de um teste de reconhecimento de fala na presença de ruído na versão brasileira na investigação do mascaramento temporal; (iii) hipótese de que a magnitude do benefício proporcionado pela modificação do ruído apresentado é dependente, dentre outros fatores, dos limiares de relação fala-ruído (RFR). Em resumo, os resultados obtidos nesta dissertação ofereceram uma melhor compreensão sobre o processamento temporal da audição, mais especificamente sobre o *masking release*, através da utilização de um teste de reconhecimento de fala na presença de ruído na versão brasileira, corroborando com estudos recentes.

Este achado está de acordo com a noção de que este benefício que se apresenta reduzido em ouvintes com perda auditiva coclear pode ser, em partes, devido a dependência do *masking release* aos limiares de RFR (Bernstein, Grant, 2009; Bernstein, Brungart, 2011; Bernstein et al., 2012; Smits, Festen, 2013).

Outro trabalho futuro deve estender esta abordagem a ouvintes idosos com audição normal, pois se tem observado que este benefício é reduzido quando comparado aos jovens, embora seus limiares de relação fala-ruído sejam equivalentes.

Os resultados obtidos nesta dissertação geraram um artigo original já publicado: (i) “*Effect of modulation rate on masking release for speech*” no periódico *Audiology - Communication Research (ACR)*, no volume 18, número 4, out./dez. 2013; e (ii) um artigo original intitulado “*Modulation masking release using the Brazilian-Portuguese HINT*”

Psychometric functions and the effect of speech time compression”, a ser submetido ao periódico *International Journal of Audiology*.

Atendendo às normas vigentes do Programa de Pós-graduação *Strictu Sensu* em Neuropsiquiatria e Ciências do Comportamento da UFPE para elaboração da dissertação, no presente exemplar os resultados obtidos nesta dissertação são apresentados no formato de artigo original.

2 INTRODUÇÃO

O reconhecimento de fala em presença de ruído mascarante é geralmente melhor se o ruído de mascaramento for modulado do que se for estável (Miller, Licklider, 1950). Este benefício é conhecido como *masking release*, que é a diferença entre os limiares de reconhecimento de fala em presença de mascaramento estável e mascaramento modulado. Uma sugestão de tradução que propomos para esse termo é: *benefício causado pela modificação do ruído mascarante* (Advíncula et al., 2013). No entanto, no presente estudo, e para fins didáticos, utilizamos o termo na língua inglesa. A magnitude do *masking release* depende de uma variedade de fatores, tais como o tipo de material de fala, a natureza da modulação do mascaramento, a relação fala-ruído (RFR), dentre outras.

Todas estas questões têm sido amplamente estudadas em diversas línguas. No entanto, nenhuma pesquisa havia utilizado a fala do Português Brasileiro como estímulo para o estudo do *masking release*. Bevilacqua et al. (2008) publicaram uma versão do Hearing in Noise Test (HINT) no Português-Brasileiro e uma motivação fundamental surgiu para caracterizar o *masking release* utilizando-se este teste, uma vez que permitiria a comparação da versão do Português Brasileiro com as versões em outras línguas onde o teste já houvesse sido padronizado, como, por exemplo, o inglês americano.

Não somente foram pesquisadas as magnitudes do *masking release* com as sentenças do teste HINT na versão do Português Brasileiro na velocidade de fala normal, como foi pesquisado o *masking release* com as sentenças do teste HINT na versão Português Brasileiro em diferentes graus de compressão temporal da fala, para determinar se a manipulação do grau de compressão temporal da fala alvo resultaria numa mudança dos limiares de RFR, e em caso afirmativo, se estas mudanças dos limiares de RFR resultaram numa modificação na magnitude do *masking release*. Neste estudo, fica caracterizado que o limiar de relação fala-ruído (RFR) corresponde à diferença, em dB, entre o valor do limiar de reconhecimento de fala na presença de ruído mascarante e o valor da intensidade do ruído competitivo utilizado.

Uma variedade de fatores relacionada ao estímulo de teste pode influenciar na magnitude do *masking release*. Um deles é o tipo de material de fala utilizado, tais como: estímulos vogal-consoante-vogal (Fullgrabe et al., 2006; Gnansia et al., 2008), sílabas sem sentido (Dubno et al., 2003; Bernstein et al., 2012), palavras monossilábicas (Miller, Licklider, 1950; Stuart, Phillips, 1996), palavras espondeicas. (Dirks, Bower, 1971), e sentenças (Jin, Nelson, 2006; Desloge et al., 2010).

Um segundo fator que pode influenciar na magnitude do *masking release* é a natureza do ruído mascarante. Esses ruídos mascarantes geralmente são ruídos com espectro de fala ou a própria fala competitiva de uma única pessoa ou de várias pessoas falando ao mesmo tempo (Festen, Plomp, 1990; Gustafsson, Arlinger, 1994; Oxenham, Simonson, 2009; Francart et al., 2011). No caso de ruídos modulados produzidos artificialmente, os padrões de modulação geralmente são modulações que se assemelham ao envelope da fala (modulações de onda senoidal ou quadrática), além de ciclos regulares e irregulares incorporados a esses padrões de flutuação (Stuart, Phillips, 1996; George et al., 2006).

Um terceiro fator relacionado ao estímulo que afeta a magnitude do *masking release* é a relação fala-ruído (RFR). Esse efeito é percebido quando se observa as inclinações (*slopes*) das curvas psicométricas do reconhecimento de fala em presença de ruído mascarante estável e em presença de ruído mascarante modulado. Essas curvas não são paralelas. A curva psicométrica para reconhecimento de fala em presença de ruído modulado é tipicamente menos íngreme do que a curva psicométrica para reconhecimento de fala em presença de ruído estável (Dirks, Bower, 1971; Bernstein, Grant, 2009; Oxenham, Simonson, 2009). Vários estudos têm demonstrado que, para os ouvintes com audição normal, o *masking release* diminui com o aumento da RFR (Oxenham, Simonson, 2009; Christiansen, Dau, 2012; Smits, Festen, 2013).

Bernstein e colaboradores (Bernstein, Grant, 2009; Bernstein, Brungart, 2011; Bernstein et al., 2012) e Smits e Festen (2013) levantaram a hipótese de que essa dependência do *masking release* da RFR é em parte a razão na qual a magnitude do *masking release* entre indivíduos de audição normal e indivíduos com perda auditiva coclear difere, uma vez que indivíduos com perda auditiva coclear exibem reduzido *masking release* (Festen, Plomp, 1990; Peters et al., 1998; Jin, Nelson, 2006; Lorenzi et al., 2006; Christiansen, Dau, 2012).

Uma série de possíveis explicações para esse deficit foi discutida na literatura. Para indivíduos com perda auditiva, não somente a audibilidade reduzida é considerada uma possível explicação para a redução do *masking release* (Desloge et al., 2010), mas também uma pobre resolução temporal (Festen, 1987). No entanto, Bernstein e colaboradores (2012) argumentam que a diferença é em grande parte devido ao fato de que os indivíduos com perda auditiva coclear geralmente recebem a fala alvo em uma RFR mais favorável do que os ouvintes com audição normal. Portanto, segundo os autores, as diferentes inclinações das curvas psicométricas (mencionadas anteriormente), poderia explicar a redução de magnitude do *masking release* em indivíduos com perda auditiva.

Essa hipótese demonstrando que, se ouvintes com audição normal recebem a fala em condições que resultam em maiores RFR, o *masking release* dos ouvintes também declina concomitantemente. Dessa forma, manipulações que resultam em alteradas RFR incluem: a quantidade de opções de respostas (*response set size*) (Buss et al., 2009; Bernstein et al., 2012), o uso da fala filtrada como fala alvo (Oxenham, Simonson, 2009; Christiansen, Dau, 2012), e o uso de falantes não nativos como fala alvo (Calandruccio et al., 2014).

Oxenham e Simonson (2009) apontaram que qualquer manipulação que reduza a redundância da fala deve resultar em um aumento no limiar de RFR. Se essa afirmativa estiver correta, um efeito semelhante deve ser observado quando a redundância da fala é alterada por meio da compressão temporal, um dos objetivos desse estudo testar esta hipótese.

A fala comprimida em tempo é uma manipulação que tem sido utilizada em uma variedade de contextos, como por exemplo, na determinação de fontes de déficits de processamento da fala relacionadas com a idade (Gordon-Salant, Fitzgibbons, 2001; Jenstad, Souza, 2007; Grose et al., 2009). A quantificação dessa compressão é normalmente expressa pela proporção do conteúdo da forma da onda que é removido. Por exemplo, 33% de compressão implica em dizer que um terço do tempo da onda original foi removido. Os algoritmos de compressão tipicamente removem segmentos da sentença falada, com o objetivo de aumentar a velocidade de fala, sem afetar significativamente entonação ou gerar outras distorções (através da remoção de períodos completos de *pitch* de vogais sonoras e através da remoção de lacunas entre as palavras). O resultado é que a redundância da fala comprimida é reduzida e o limiar de reconhecimento de fala é aumentado, particularmente quando em presença de ruído.

Acredita-se que os limiares de RFR aumentam em função da compressão temporal da fala e, simultaneamente, que a magnitude do *masking release* diminui. Para tal, um estudo preliminar foi realizado para verificar se o teste HINT na versão do Português Brasileiro apresentaria a dependência esperada da inclinação (*slope*) na curva psicométrica do desempenho-intensidade de acordo com as características do mascaramento (estável x modulado).

Dessa forma, o objetivo principal desta dissertação foi investigar o benefício causado pela modificação do ruído mascarante (*masking release*) em diferentes situações de escuta em adultos jovens com audição normal. Inicialmente, foi investigado o efeito das diferentes taxas de modulações do mascaramento (0, 4, 8, 16, 32 e 64 Hz) na magnitude do benefício do mascaramento modulado em adultos nativos do Português do Brasil com audição normal. Em

seguida foram determinadas as funções desempenho-intensidade para o material de fala do HINT na versão Português Brasileiro na presença de ruído estável e modulado a 10 Hz e suas inclinações em adultos nativos do Português do Brasil com audição normal e os limiares de relação fala-ruído nos diferentes graus de compressão temporal do material de fala do teste HINT na versão Português Brasileiro em adultos nativos do Português do Brasil com audição normal.

3 REVISÃO DE LITERATURA

3.1 Testes de Fala no Ruído

Um dos aspectos mensuráveis mais importantes da função auditiva é a habilidade reconhecer da fala. Por isso, os testes de reconhecimento de fala são de grande importância para o diagnóstico audiológico.

O Limiar de Reconhecimento de Fala e o Índice Percentual de Reconhecimento de Fala são testes que avaliam a habilidade de reconhecer a fala em condições adequadamente controladas. Estes testes são utilizados para medir o desempenho auditivo dos indivíduos em tarefas de reconhecimento de fala utilizando como material de fala estímulos isolados (sem presença de ruído competitivo), como palavras monossílabas, dissílabas e polissílabas. Existem basicamente duas listas de palavras que tem sido mais utilizada na aplicação do teste de reconhecimento de fala (Pen, Mangabeira-Albernaz, 1970; Santos, Russo, 1986). No entanto, outros autores também têm sugerido novas listas para aplicação do teste (Chaves et al., 1999; Harris et al., 2001; Rol et al., 2003).

O reconhecimento de fala é resultante da combinação das pistas acústicas, linguísticas, semânticas e circunstanciais (Gama, 1994). Contudo, em ambientes ruidosos o sujeito pode apresentar inúmeras dificuldades no reconhecimento de fala, pois o número das pistas descritas acima reduz significativamente, levando-o utilizar somente algumas das pistas disponíveis na situação do contexto onde se encontra. Este fato justifica a preocupação em não apenas medir habilidade em reconhecer a fala em condições favoráveis, nas quais os estímulos concorrentes estão controlados, mas também em situações mais próximas da vida real, nas quais as pistas nem sempre estão presentes (Caporali, Silva, 2004).

Uma das principais limitações dos testes da avaliação audiológica básica que avaliam a habilidade de reconhecer a fala citados é o fato destes não serem realizados em condições de comunicação próximas às encontradas no cotidiano, que envolve a percepção da fala em presença de ruído (Henriques, Costa, 2011). Entretanto, o reconhecimento da fala na presença do ruído competitivo é uma das mais complexas tarefas diárias realizadas pelo ouvinte. Para avaliar a habilidade de ouvir em presença de ruído deve-se fazer uso de testes de fala que avaliam como o ouvinte escuta nesta situação.

Como dito anteriormente, alguns fatores devem ser levados em consideração quando se avalia a habilidade de ouvir em presença de ruído: o material de fala utilizado, a presença ou ausência de ruído competitivo (Arieta, 2009), e o tipo de ruído utilizado.

Quanto ao material de fala, sílabas e palavras têm sido utilizadas para medir o desempenho auditivo do indivíduo em tarefas de reconhecimento de fala (Santos, Daniel, Costa, 2009). Entretanto, na avaliação do reconhecimento de fala na presença de ruído, o uso de sentenças é melhor que o uso de palavras, pois as sentenças mais se aproximam das situações reais de comunicação (Bronkhorst, Plomp, 1990).

No Brasil, alguns estudos utilizaram o teste de reconhecimento de fala com ruído branco (Pereira, 1993; Schochat, 1994). Outros autores recomendaram a utilização de ruído competitivo do tipo cafeteria (Costa, 1992), reconhecimento de palavras monossilábicas mascarados por fala competitiva (Costa, 1995), ruído competidor com espectro de fala (Costa et al., 1998) e modulações de amplitude (Costa, 1998) e o ruído cocktail party (Mantelatto, 1998; Mantelatto, Silva, 2000a; Mantelatto, Silva, 2000b; Caporali, Arieta, 2004).

Entretanto, nas últimas décadas diversos testes de reconhecimento de fala na presença de ruído constituídos utilizando-se listas de sentenças têm sido desenvolvidos (Kalikow, Stevens, Elliot, 1977; Hagerman, 1982; Plomp, Mimpen, 1979; Smoorenburg, 1992; Nilsson, Soli, Sullivan, 1994). Por exemplo, no Brasil, o teste Listas de Sentenças em Português (Costa, 1998) foi pioneiro para este fim, cuja confiabilidade foi verificada posteriormente (Freitas, Lopes, Costa, 2005). Neste estudo, observou-se que houve alta confiabilidade entre os limiares de reconhecimento de sentenças no silêncio e na presença de ruído (Freitas, Lopes, Costa, 2005). Diferentes estratégias de aplicação para este teste foram realizadas, como a apresentação dos estímulos de fala e ruído em campo livre, por simular condições semelhantes às encontradas em situações reais de comunicação (Henriques, Miranda, Costa, 2008; Henriques, Costa, 2011), bem como apresentação com fones auriculares (Costa, Daniel, Santos, 2011).

Em 2008, foi desenvolvido o *Hearing in Noise Test* na versão Português Brasileiro (Bevilacqua et al., 2008). Este teste utiliza sentenças para mensurar a habilidade de reconhecer a fala no silêncio e na presença de ruído.

3.2 Hearing in Noise Test

O *Hearing in Noise Test* (HINT) é um teste de reconhecimento de fala em formato *open-set* utilizado para medir os limiares de reconhecimento de fala de sentença (LRSs) no silêncio e as relações fala-ruído em que os LRSs foram obtidos em condições de ruído, sendo inicialmente criado no inglês americano (Nilsson et al., 1994).

Este teste foi desenvolvido em 1994 (Nilsson; Soli; Sullivan, 1994), no *House Ear Institute*, sendo inicialmente comercializado e testado por meio de um compact disk, acoplado ao audiômetro. Em 2003, a empresa *Bio-Logic Systems Corp* criou um *hardware* e um *software* que possibilitaram novas versões do HINT (Duncan, Aarts, 2006). Foi primeiramente testado em ouvintes normais para obtenção de parâmetros de normatização. O teste é composto por 12 listas foneticamente balanceadas, equivalendo a 20 sentenças cada. Cada sentença varia de seis a sete sílabas de comprimento e é classificada em um nível de leitura de primeiro grau. As sentenças HINT foram gravadas por um falante do sexo masculino nativo do inglês americano. O ruído estável foi combinado com o espectro médio de longo prazo das sentenças.

Posteriormente, outras versões do teste HINT foram desenvolvidas em diferentes línguas, sob as mesmas premissas metodológicas e procedimentos, a fim de terem-se medidas comparáveis da habilidade de percepção de fala de maneira que, a depender do objetivo do estudo, permitissem comparações serem generalizadas (Soli, Wong, 2008).

As versões do teste HINT podem ser encontradas em diversas línguas: no inglês americano (Nilsson, Soli, Sullivan, 1994; Vermiglio, 2008), no espanhol da América latina (Otero et al., 2008), no Português Brasileiro (Bevilacqua et al., 2008), no turco (Cekic, Sennaroglu, 2008), no espanhol castelhano (Huarte, 2008), no búlgaro (Lolov, 2008), no francês (Luts et al., 2008), no coreano (Moon et al., 2008), no norueguês (Myhrum, Moen, 2008), no malaio (Quar et al., 2008), no japonês (Shiroma et al., 2008), no francês do Canadá (Vaillancourt et al., 2008), no cantonês (Wong, 2008) e no mandarim (Wong, Huang, 2008; Wong, Liu, Han, 2008; Wong et al., 2007). As semelhanças dos procedimentos e dos materiais de teste fazem das medidas obtidas com o teste HINT comparáveis entre os idiomas (Bevilacqua et al., 2008).

Os materiais de teste foram selecionados a partir de um grande número de sentenças cotidianas simples (336 sentenças para a versão em inglês americano), que foram classificadas de acordo com a naturalidade das sentenças por falantes nativos da língua alvo. Sentenças com baixa redundância de fala foram descartadas. A dificuldade das sentenças do teste foi equalizada utilizando-se os mesmos procedimentos para as diferentes versões linguísticas do teste HINT. Sentenças que eram difíceis de serem percebidas foram descartadas. Todos os materiais de fala, incluindo materiais de língua inglesa e estrangeira, foram gravadas e processadas no mesmo instituto, sob os mesmos procedimentos (Shiroma et al., 2008; Soli, Wong, 2008).

As médias e os desvios-padrão da amostra normativa das versões do teste HINT encontram-se disponíveis, bem como valores normativos para a versão em inglês americano do teste HINT, para as listas de 20 sentenças foram obtidas a partir de três centros de teste, com um total de 67 falantes nativos do inglês americano (Vermiglio, 2008). Os resultados do HINT são apresentados de duas formas, limiares de reconhecimento de sentença no silêncio e limiares de RFRs. Os limiares de reconhecimento de sentença no silêncio são obtidos por meio de procedimento adaptativo padrão para medir uma pontuação de 50% de acerto. Os limiares de RFRs HINT refletem as RFRs em que os ouvintes alcançam uma pontuação de 50% de acerto em condições de ruído. Uma lista de 20 sentenças é utilizada para ambos os limiares. Enfim, o teste HINT possibilita utilizar limiares de reconhecimento de fala com o intuito de avaliar o desempenho da fala na presença do ruído competitivo (GIGUÈRE et al., 2008).

Há de se destacar como os limiares de reconhecimento de sentença no silêncio (condição de silêncio) e os limiares de RFRs (condição de ruído) são obtidos. Para condições silêncio, a primeira sentença em cada lista é apresentada a uma intensidade de 20 dBA. Quando a sentença não é repetida corretamente, a mesma sentença é apresentada várias vezes, e a intensidade é aumentada em passos de 4 dB até que o ouvinte repeta a sentença corretamente. As sentenças subsequentes na lista são apresentadas uma vez cada. Durante as primeiras quatro sentenças, a intensidade da fala é reduzida em 4 dB após uma resposta correta e um aumento de 4 dB depois de uma resposta incorreta. Após a apresentação da quarta sentença, um limiar é estimado tomando a média da: (1) a intensidade final de apresentação da primeira sentença, (2) as intensidades de apresentação da segunda a quarta sentença, e (3) a intensidade em que a quinta sentença seria apresentada (isto é, ou 4 dB mais elevado ou mais reduzido do que a intensidade de apresentação da quarta sentença). A quinta sentença é, em seguida, apresentada a uma intensidade do limiar estimado. Da quinta a vigésima sentença, a intensidade da fala varia em passos de 2 dB de acordo com a resposta do participante. O limiar de reconhecimento de sentença é calculado a partir do nível médio apresentação da quinta a vigésima sentenças e da intensidade em que a sentença vigésima primeira seria apresentada. Para a medição dos limiares de RFRs, o ruído é apresentado em uma intensidade fixa de 65 dBA, e a primeira sentença na lista é apresentada a uma intensidade de 65 dBA. A intensidade da fala varia de acordo com a resposta do ouvinte. O mesmo protocolo adaptativo é utilizado nas condições de ruído, tal como descrito anteriormente para as condições silêncio para determinar o nível do estímulo que o ouvinte

alcançava o percentual de 50% de acerto, valor esse que é convertido para o RFR (limiar de RFR) (Nakamura, Gordon-Salant, 2011).

Em resumo, o HINT possibilita utilizar limiares de reconhecimento de fala com o intuito de avaliar o desempenho da fala na presença do ruído competitivo (Giguère et al., 2008). Através da utilização do teste HINT, torna-se prático e possível confirmar as vantagens da audição direcional binaural, pois os testes podem ser aplicados em diferentes direções e proporcionam medidas que não são detectadas pela audiometria convencional (Nilsson, Soli, Sullivan, 1994).

No Brasil, Bevilacqua et al. (2008) desenvolveram o teste HINT na versão do Português Brasileiro. Para a preparação do teste, 1.700 sentenças foram compiladas por meio de três métodos diferentes. Após uma avaliação de familiaridade das sentenças, foram selecionadas 800 para utilização do estudo inicial. Dois centros participaram da pesquisa: Universidade de São Paulo, Bauru e Universidade Estadual de Campinas, Campinas. As sentenças selecionadas foram gravadas no *House Ear Institute*, em Los Angeles, por um profissional locutor brasileiro do sexo masculino.

Em uma segunda etapa, foi estimada a função desempenho-intensidade por meio da avaliação de doze indivíduos, sendo seis de cada instituição acima citada. A partir do material de fala existente, foram produzidas 6 listas de 50 sentenças, totalizando 300 sentenças (três listas para cada centro), e um ruído mascarante com o mesmo espectro de frequência das sentenças. A intensidade do ruído foi constante em 65 dB A, com as seguintes variações do sinal para taxas de ruído (RFR): - 7 dB, -4 dB e -2 dB, para todas as listas testadas. As palavras corretas foram contadas, resultando em um percentual de inteligibilidade, para cada lista, em cada RFR. A média da inclinação da função desempenho-intensidade para os dois centros foi 11,4% por mudança de dB na RFR. Posteriormente foi realizada a equalização da dificuldade das sentenças.

A partir dos dados da fase de equalização, foram criadas 24 listas de sentenças foneticamente balanceadas. Transcrições fonêmicas codificadas foram inseridas em um programa de computador que contou automaticamente os fonemas em cada sentença de cada lista. O resultado final originou a criação de 24 listas, com 10 sentenças foneticamente balanceadas. Ao final, as 240 sentenças foram redistribuídas em 12 listas com 20 sentenças cada. A inteligibilidade de fala foi avaliada em 29 indivíduos de ambos os sexos, na faixa

etária de 18-45 anos, em quatro situações: silêncio, ruído frontal, o ruído à direita e o ruído à esquerda (Bevilacqua et al., 2008).

O microprocessador HTD (*Hearing Test Device*), versão 7.2, fabricado pela Bio-logic, foi desenvolvido pelo Laboratório de Pesquisas de Aparelhos Auditivos do Departamento de Ciência e Comunicação Humana do HEI (*House Ear Institute*), no ano de 1994. Este equipamento contém o *software* que foi utilizado para conduzir o processo do teste HINT na versão do Português Brasileiro (sentenças e ruído).

O teste HINT na versão do Português Brasileiro pode ser realizado em duas situações: com fones auditivos e em campo livre. Na situação com fones auditivo, o limiar de reconhecimento de fala é obtido por meio da apresentação de 20 sentenças gravadas que simulam quatro condições, a saber: (1) fala sem ruído (S): o sinal de fala gravado é apresentado em ambas as orelhas sem competição; (2) fala com ruído frontal (RF): o sinal e o ruído são apresentados de maneira bilateralmente a simularem esta condição; (3) fala com ruído à direita (RD): o sinal é apresentado na orelha esquerda e o ruído na orelha direita; e, (4) fala com ruído à esquerda (RE): o sinal é apresentado na orelha direita e o ruído na orelha esquerda.

Na situação em campo livre, a fala e o ruído são emitidos por duas caixas de sons devidamente posicionadas, a saber: (1) fala sem ruído (S): a fala sai à frente do sujeito em uma condição de teste sem ruído (0° azimuth); (2) fala com ruído frontal (RF): a fala e o ruído são colocados diretamente em frente ao sujeito em uma condição de ruído em 0° azimuth; (3) fala com ruído a direita (RD): a fala é posicionada à frente do sujeito e o ruído é emitido a 90° a direita do sujeito (Condição ruído à direita - ruído em 90° azimuth); e, (4) fala com ruído a esquerda (RE): o sinal é posicionado à frente do sujeito e o ruído é emitido a 90° a esquerda do sujeito (condição ruído à esquerda - ruído em 270° azimuth).

Em ambas as situações, o *software* elabora automaticamente a situação denominada Ruído Composto (RC), obtido através da seguinte fórmula (Arieta, 2009):

$$FR = \frac{2 (RF + RD + RE)}{4}$$

3.3 Diferentes Medições do Desempenho no Reconhecimento de Fala

O desempenho do reconhecimento de fala em presença de ruído competitivo pode variar de acordo com a forma na qual é avaliado, sendo um dos fatores que influenciam a habilidade de reconhecer a fala na presença do ruído, a tarefa de escuta.

Na mensuração do reconhecimento de fala geralmente são utilizados dois tipos de tarefas de escuta ou paradigmas, aquela com relações fala-ruído (RFR) fixas e aquela com RFR adaptativas variadas. Os resultados de um paradigma com a RFR fixa são geralmente expressos como pontuações percentuais corretas e os resultados de um paradigma adaptativo são dados como a RFR no limiar do reconhecimento de fala, ou seja, limiares de RFR.

Para a determinação de resultados expressos em limiares de RFR, utiliza-se um procedimento adaptativo no qual a intensidade em que um determinado estímulo é apresentado ao sujeito depende de como o sujeito respondeu aos estímulos anteriores (Wetherill, Levitt, 1965; Levitt, 1971; Bode, Carhart, 1973). A utilização do termo “procedimento adaptativo” tem sido associado a métodos que tendem a convergir sobre a intensidade do limiar (ou algum outro ponto alvo), apresentando também a vantagem de não exigir um conhecimento prévio de onde o limiar está localizado, uma vez que os métodos adaptativos tendem a convergir para o limiar, independentemente do ponto de partida.

Neste procedimento, o ponto de partida geralmente tem intensidade mais forte, que vai diminuindo e aumentando de acordo com a resposta do indivíduo, em intervalos predeterminados. Esses intervalos são inicialmente maiores e, em seguida, tornam-se menores na medida em que a intensidade se aproxima do limiar. Sugere-se que sejam utilizados intervalos de 4 dB até que haja uma reversão da resposta e, posteriormente, os intervalos de apresentação dos estímulos sejam de 2 dB entre si, até a obtenção do limiar de reconhecimento (Levitt, Rabiner, 1967). Como resultado, a eficiência e precisão das respostas são maximizadas (Gelfand, 2010).

O procedimento é realizado da seguinte maneira: uma corrida descendente começa com uma resposta positiva; continua a diminuir até que haja uma resposta negativa; por conseguinte, uma reversão. Uma corrida ascendente começa com uma resposta negativa e termina com uma positiva (Wetherill, 1963; Wetherill, Levitt, 1965). O procedimento é continuado por meio de pelo menos seis a oito reversões (exceto o primeiro) e o valor do limiar de reconhecimento de fala é, então, calculado como a média dos valores médios das corridas, ou como a média dos seus picos e depressões (Wetherill, 1963; Wetherill, Levitt, 1965).

No contexto das medidas de reconhecimento de fala na presença de ruído, a função de reconhecimento de fala ou função psicométrica relaciona os percentuais de estímulos corretamente repetidos (plotados no eixo y) em função das diferentes entradas de estímulo ou diferentes relações fala-ruído (plotada no eixo-x). Ou seja, mostra a probabilidade (em porcentagem) de respostas para diferentes intensidades de estímulo (Smits, Festen, 2011). Para isso, torna-se necessária a escolha de intervalos de relações fala-ruído cujo intervalo abranja pontuações de percentuais corretos que variam de aproximadamente 10 a 90% (Nakamura, Gordon-Salant, 2011), para evitar o efeito piso ou efeito teto.

Para avaliar a habilidade de percepção de fala no silêncio e no ruído utilizando o Hearing in Noise Test (HINT) no inglês americano e no japonês, foram utilizados os dois paradigmas de testes, o paradigma procedimento adaptativo e o paradigma de nível fixo. Na avaliação da habilidade de percepção tanto no silêncio quanto no ruído foi utilizado um procedimento adaptativo para determinar o limiar de reconhecimento de fala do ouvinte. O desempenho no reconhecimento de fala foi também avaliado utilizando as RFR fixas, para a construção de funções de desempenho-intensidade, escolhidas para abranger pontuações de percentuais corretos que variam de aproximadamente 10 a 90%. A percentagem de acertos foi determinada dividindo-se o número total de palavras-chave corretamente identificadas pelo número total de palavras-chave da lista de sentenças. Foi observada uma diferença de 2,5 dB entre os limiares obtidos através do procedimento adaptativo e os limiares, estimados em 50% de percentual de acerto no procedimento com níveis fixos de RFR (Nakamura, Gordon-Salant, 2011).

Estudos anteriores (Taylor et al. 1983; Kollmeier et al. 1988; Stillman 1989) corroboram com esses achados, demonstrando que os limiares de reconhecimento de fala obtidos, utilizando-se os procedimentos adaptativos foram diferentes daqueles que foram obtidos utilizando-se o paradigma de nível fixo, mesmo quando os dois métodos tinham as mesmas probabilidades alvo para respostas corretas. Gordon-Salant (1987) relatou que diferentes paradigmas de teste (ou seja, um paradigma procedimento adaptativo contra um paradigma de nível fixo) resultaram em sensibilidade diferente para identificar os efeitos da idade no reconhecimento de fala no ruído entre os ouvintes idosos.

Influenciando a performance do reconhecimento de fala, para pontuações de percentuais corretos, refere-se à forma de contabilizar o acerto/erro da resposta apresentada pelo ouvinte. Pode-se considerar o acerto da sentença completa ou o acerto de partes da sentença (o número de palavras repetidas corretamente em uma sentença). Além disso, testes

de sentenças utilizam diferentes métodos de pontuação, por exemplo, palavras-chave ou *word scoring* (Kollmeier, Wesselkamp, 1997; Killion et al., 2006).

Plomp e Mimpen (1979) desenvolveram um teste de reconhecimento de sentenças no ruído denominado *Dutch speech-in-noise test*, no qual a resposta pode ser considerada correta somente quando a sentença completa é repetida corretamente (ou seja, sentença de pontuação). Versfeld et al. (2000) mostraram que o limiar de reconhecimento de fala do teste desenvolvido por Plomp e Mimpen (1979) é menor para a palavra de pontuação do que para a sentença de pontuação. Terband e Drullman (2008) encontraram efeitos semelhantes quando utilizaram a pontuação da palavra-chave.

No paradigma fixo pode ser considerada, na análise da resposta, a sentença completa ou o número de palavras repetidas corretamente na sentença. No entanto, no paradigma adaptativo, a sentença deve ser repetida completamente de forma correta, sendo marcada como correta ou incorreta, de modo que a regra adaptativa para a fixação da RFR para o próximo item pode ser aplicada.

3.4 Fatores Relacionados aos Estímulos de Teste que Afetam o *Masking Release*

Uma variedade de fatores relacionada ao estímulo de teste pode influenciar na magnitude do *masking release*. São descritos neste estudo o tipo de material de fala, o tipo de mascaramento e a relação fala-ruído.

Um deles é o tipo de material de fala utilizado, tais como: estímulos vogal-consoante-vogal (Fullgrabe et al., 2006; Gnansia et al., 2008), sílabas sem sentido (Dubno et al., 2003; Bernstein et al., 2012), palavras monossilábicas (Miller, Licklider, 1950; Stuart, Phillips, 1996), palavras espondeicas (Dirks, Bower, 1971), e sentenças (Jin, Nelson, 2006; Desloge et al., 2010).

Diferentes tipos de materiais de fala têm sido utilizados na avaliação do *masking release* como citado anteriormente. A diversidade dos materiais utilizados depende dos objetivos dos estudos. Por exemplo, sentenças sem sentido são utilizadas, a fim de minimizar o efeito de contexto no reconhecimento de fala.

Diferenças nos procedimentos relacionados ao material de fala também modificam o desempenho no reconhecimento de fala na presença de ruído, como a utilização de testes *open*

set e *closed set*, uma vez que materiais de fala *open set* são mais difíceis de serem reconhecidos quando comparados a materiais de fala *closed set*.

O tamanho do conjunto do material de teste pode modificar o desempenho no reconhecimento da fala na presença de ruído. Bernstein et al. (2012) avaliaram um método de conjunto de tamanho adaptável para ajustar os níveis de desempenho do reconhecimento de fala na presença de diferentes tipos de ruídos. Na presença de ruído estável os ouvintes com perda auditiva geralmente mostram déficits de desempenho de reconhecimento de fala em relação a ouvintes com audição normal. Com isso, os algoritmos adaptativos de rastreamento convergem para diferentes RFR para ouvintes com audição normal e com perda auditiva, podendo confundir a interpretação dos dados em experimentos que comparam o *masking release* para ouvintes com audição normal e com perda auditiva. Isto porque as diferenças na RFR para a condição linha de base (ruído estável) tem uma influência sobre a quantidade do *masking release* observado.

O método empregado envolveu a identificação de consoante-vogal ou vogal-consoante, com o tamanho do conjunto de respostas adaptativamente ajustadas, para se obter um nível de desempenho percentual correto alvo, num dado nível de RFR no ruído estável. Sua hipótese foi que esse ajuste poderia igualar pontuação de reconhecimento em uma determinada RFR, eliminando a complicação na interpretação dos resultados. Os resultados demonstraram a viabilidade da utilização do tamanho do conjunto do material de teste como uma ferramenta para equalizar os níveis de desempenho entre os dois grupos de ouvintes (com audição normal e com perda auditiva), evitando a confusão relacionada às diferenças na RFR utilizadas na medição do *masking release*.

Outro fator que interfere na magnitude do *masking release* é a natureza do ruído mascarante. Freyman et al. (2004) referem que ouvir a fala na presença de ruído competitivo de fala é uma tarefa perceptual complexa, que tem sido o objeto de vários estudos ao longo dos anos.

Como dito anteriormente e como foco principal desta dissertação, ao longo das últimas décadas, uma grande variedade de estudos tem demonstrado que ouvintes com audição normal mostram melhora na inteligibilidade de fala na presença de ruído modulado em relação ao ruído estável quando os dois ruídos são apresentados ao mesmo nível de pressão sonora de longo do tempo (Miller, Licklider, 1950; Carhart, Tillman, 1970; Festen, Plomp, 1990; Eisenberg et al., 1995; Bacon et al., 1998; Peters et al., 1998; Dubno et al., 2003; George et al., 2006; Jin, Nelson, 2006; Wilson et al., 2007; Bernstein, Grant 2009), um

fenômeno muitas vezes referido como “*release from masking*” ou “*speech masking release*” (Gnansia et al., 2008) ou benefício causado pela modificação do ruído mascarante (Advíncula et al., 2013).

Entre os vários mecanismos conhecidos envolvidos no efeito *masking release*, o *dip listening* e o mecanismo *auditory grouping* parecem desempenhar um papel importante (Fullgrabe et al, 2006). A magnitude do *masking release* depende de algumas características físicas do ruído mascarante, como por exemplo, a sua frequência de modulação (Nelson et al, 2003; Füllgrabe et al, 2006; Rhebergen et al, 2006; Stuart, Phillips, 1996; Summers, Molis, 2004; Rhebergen et al, 2006; George et al, 2006). George e colaboradores (2006) encontraram variações na melhora do reconhecimento da fala entre 15 e 25 dB para frequências de modulações entre 8 e 20 Hz. Tipicamente, tem sido observada maior magnitude do *masking release* para frequências de modulações mais baixas (Desloge et al, 2010), tais como modulações entre 8 Hz e 25 Hz (especialmente em 10 Hz) (Buss et al, 2013; Füllgrabe et al, 2006; Lorenzi et al, 2006). Isso porque frequências de modulações mais baixas produzem maiores espaços temporais de menor amplitude (*dip listening*). Dessa forma, facilitam a percepção da fala (*auditory grouping*) (Shejif, Yost, 2007).

Um terceiro fator relacionado ao estímulo que afeta a magnitude do *masking release* é a relação fala-ruído (RFR). O limiar de reconhecimento de fala tem-se revelado útil quando se compara o desempenho no reconhecimento de fala em condições de escuta que geram diferentes níveis de desempenho em uma mesma RFR. Por exemplo, ouvintes com audição normal normalmente demonstram um melhor desempenho no reconhecimento de fala na presença do ruído modulado que para um ruído estável quando ambos são apresentados em uma mesma RFR (Miller, Licklider, 1950; Festen, Plomp, 1990). Vários estudos têm demonstrado que, para os ouvintes com audição normal, o *masking release* diminui com o aumento da RFR (Oxenham, Simonson, 2009; Christiansen, Dau, 2012; Smits, Festen, 2013). Esse efeito é percebido quando se observa as inclinações (*slopes*) das curvas psicométricas do reconhecimento de fala em presença de ruído mascarante estável e em presença de ruído mascarante modulado. Essas curvas não são paralelas, uma vez que a curva psicométrica para reconhecimento de fala em presença de ruído modulado é tipicamente menos íngreme do que a curva psicométrica para reconhecimento de fala em presença de ruído estável, ou seja, a inclinação da função psicométrica é mais acentuada para o ruído estável (Dirks, Bower, 1971; Bernstein, Grant, 2009; Oxenham, Simonson, 2009). O *masking release* variou com a RFR, sendo maior em RFRs menos favoráveis, onde as duas curvas desviaram mais.

O limiar de reconhecimento de fala também é utilizado para avaliar o reconhecimento de fala entre diferentes grupos de ouvintes que variam substancialmente no seu desempenho. Por exemplo, ouvintes com perda auditiva podem apresentar desempenhos próximos a 0% de acerto em muitas condições de fala, nas quais os ouvintes com audição normal podem marcar quase 100% de acerto em uma mesma RFR. Vários estudos têm demonstrado que ouvintes com perda coclear exibem reduzido *masking release* (Festen, Plomp, 1990; Peters et al., 1998; Jin, Nelson, 2006; Lorenzi et al., 2006; Christiansen, Dau, 2012). Uma série de razões para isso foi discutida na literatura, incluindo a audibilidade reduzida (Desloge et al., 2010) e pobre resolução temporal (Festen, 1987) dos ouvintes com perda auditiva.

No entanto, Bernstein e colaboradores argumentam que a diferença é em grande parte devido ao fato de que os indivíduos com perda auditiva coclear geralmente recebem a fala alvo em uma RFR mais favorável do que os ouvintes com audição normal. Portanto, segundo os autores, as diferentes inclinações das curvas psicométricas (mencionadas anteriormente), poderia explicar a redução de magnitude do *masking release* em indivíduos com perda auditiva. Eles sustentam essa hipótese demonstrando que, se ouvintes com audição normal recebessem a fala em condições que resultassem em maiores RFR, o *masking release* dos ouvintes também declinaria concomitantemente.

Bernstein e colaboradores (Bernstein, Grant, 2009; Bernstein, Brungart, 2011; Bernstein et al., 2012) e Smits e Festen (2013) levantaram a hipótese de que essa dependência do *masking release* do limiar de RFR é a razão na qual a magnitude do *masking release* entre indivíduos de audição normal e indivíduos com perda auditiva ser diferente, uma vez que indivíduos com perda auditiva coclear exibem reduzido *masking release*.

Oxenham e Simonson (2009) apontaram que qualquer manipulação que reduza a redundância da fala deve resultar em um aumento no limiar de RFR. A fala comprimida em tempo é uma manipulação que tem sido utilizada em uma variedade de contextos, como por exemplo, na determinação de fontes de déficits de processamento da fala relacionadas com a idade (Gordon-Salant, Fitzgibbons, 2001; Jenstad, Souza, 2007; Grose et al., 2009). A quantificação dessa compressão é normalmente expressa pela proporção do conteúdo da forma da onda que é removido. Por exemplo, 33% de compressão implica em dizer que um terço do tempo da onda original foi removido. Os algoritmos de compressão tipicamente removem segmentos da sentença falada, com o objetivo de aumentar a velocidade de fala, sem afetar significativamente entonação ou gerar outras distorções (através da remoção de períodos completos de *pitch* de vogais sonoras e através da remoção de lacunas entre as

palavras). O resultado é que a redundância da fala comprimida é reduzida e o limiar de reconhecimento de fala é aumentado, particularmente quando em presença de ruído.

4. HIPÓTESES

4.1 Experimento 1

A hipótese foi que menores taxas de modulações produzem maior benefício do mascaramento modulado.

4.2 Experimento 2

A hipótese foi de que a inclinação da curva psicométrica em presença de ruído modulado é menor do que a inclinação da curva psicométrica em presença de ruído estável e, por conseguinte, que a magnitude do *masking release* difere em função do limiar de relação fala-ruído.

4.3 Experimento 3

As hipóteses foram que o limiar de relação fala-ruído aumenta à medida que o grau de compressão temporal aumenta, e que o *masking release* diminui à medida que o limiar de relação fala-ruído aumenta.

5 MÉTODO

A presente pesquisa foi aprovada pelo Comitê de Ética americano, *Institucional Review Board*, sob o número 11-1113, pelo Comitê Nacional de Ética em Pesquisa (Comissão Nacional de Ética em Pesquisa – CONEP), sob o número 233/2012 e pelo Comitê de Ética em Pesquisa envolvendo Seres Humanos, do Centro de Ciências da Saúde da Universidade Federal de Pernambuco: 02466612.2.0000.5208. Trata-se de uma pesquisa analítica, do tipo transversal, observacional com amostra por conveniência.

Foi realizada no período entre julho de 2012 a julho de 2013, no Laboratório de Audiologia da Clínica Escola do Curso de Fonoaudiologia da Universidade Federal de Pernambuco (UFPE), situado no Hospital das Clínicas da UFPE, sob a supervisão e orientação de docentes do Departamento de Fonoaudiologia e de outros departamentos.

A pesquisa foi composta por 55 adultos nativos do Português do Brasil sem passado de problemas otológicos e neurológicos; com meatoscopia normal, limiares auditivos tonais ≤ 20 dB NA, nas frequências compreendidas entre 250 e 8000 Hz (incluindo as frequências interoitavas de 3000 e 6000 Hz), com idade entre 17 a 35 anos, de ambos os sexos. A amostra foi dividida de acordo com os experimentos realizados, a saber: 15 participantes no Experimento 1; 10 participantes no Experimento 2; e 30 participantes no Experimento 3.

Os participantes foram recrutados dentre os estudantes da Universidade Federal de Pernambuco (UFPE), os quais foram convidados a participar da pesquisa. Em data agendada, de acordo com a conveniência do participante e dos pesquisadores envolvidos, foram explicados todos os objetivos e procedimentos da pesquisa, incluindo explicação dos riscos e benefícios de sua participação. Após aceitarem participar da pesquisa, os participantes assinam um Termo de Consentimento Livre e Esclarecido – TCLE (APÊNDICE A).

No dia agendado, antes da realização de cada exame, os indivíduos responderam a um questionário semiestruturado (APÊNDICE B) e realizaram os exames audiológicos, a saber: meatoscopia, audiometria tonal e vocal. Após atender aos critérios de inclusão, foi dado prosseguimento aos exames com um teste de reconhecimento de sentença na presença de ruído, utilizando os estímulos do HINT na versão do Português Brasileiro.

Para atingir os objetivos desta dissertação, foram realizados 3 experimentos, descritos a seguir:

A proposta do Experimento 1 foi investigar o efeito das diferentes taxas de modulações do mascaramento (0, 4, 8, 16, 32 e 64 Hz) na magnitude do benefício do mascaramento

modulado em adultos nativos do Português do Brasil com audição normal utilizando as sentenças do teste HINT na versão do Português Brasileiro. O objetivo foi testar a hipótese de menores taxas de modulações produzem uma maior magnitude do *masking release* para a fala.

Os participantes desse primeiro experimento foram 15 adultos jovens (5 mulheres), com idades entre 17-35 anos (média de 21,0 anos). Todos eram falantes nativos do Português do Brasil e apresentaram audição normal (limiões tonais ≤ 20 dB para as frequências de oitava 250-8000 Hz). Nenhum participante relatou história de doença otológica ou neurológica. Todos os sujeitos assinaram um TCLE para a participação no estudo.

O material de teste foi o HINT na versão do Português Brasileiro (Bevilacqua et al., 2008). Esta versão do HINT consiste em 12 listas de 20 sentenças por lista. As sentenças, gravadas originalmente no HEI, nos Estados Unidos.

O ruído mascarante utilizado foi o ruído fornecido com o teste HINT na versão do Português Brasileiro. Este ruído tem o mesmo formato espectral em longo prazo do espectro de fala médio das sentenças que compreende o material teste. O ruído foi apresentado continuamente sob duas condições: estável e modulado em amplitude. Na condição estável, o ruído foi apresentado em intensidade fixa de 65 dB NPS. Na condição de modulação, o ruído foi modulado por uma onda quadrática entre 65 dB NPS e 30 dB NPS, nas diferentes taxas de modulação (4, 8, 16, 32 e 64 Hz).

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. As sentenças foram enviadas à orelha direita dos participantes pelo fone de ouvido Sennheiser HD580, através de uma plataforma de processamento de sinal digital (Tucker-Davis RX6) e sob o controle de um computador que executa um *script* Matlab™ personalizado. Os limiões de reconhecimento de sentença foram medidos utilizando um procedimento adaptativo que converge para 71% de acerto. Os sujeitos foram testados em uma cabina acústica, e orientados a repetir cada sentença da forma como era percebida. Fora da cabina, o pesquisador monitorou a resposta oral do sujeito através de fones de ouvido ligados a um microfone posicionado dentro da cabine. O texto da sentença foi apresentado simultaneamente na tela do computador na frente do experimentador, com todas as palavras em destaque num retângulo sombreado sensível à marcação. O pesquisador usou o mouse do

computador para marcar as palavras que foram omitidas ou repetidas incorretamente. No entanto, para a proposta do procedimento adaptativo, a sentença foi dada em uma pontuação global de “correto” ou “incorreto”, na qual a sentença completa teve de ser repetida com precisão para uma pontuação de “correta” e qualquer erro resultou em uma pontuação de “incorreta”. 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*). 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). A estimativa do limiar foi calculada como a média dos quatro níveis (intensidades) finais de reversão. Para cada participante, quatro estimativas de limiares foram obtidas para uma dada condição de mascaramento. O limiar final foi calculado como a média de todas as estimativas obtidas. 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). 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 proposta do Experimento 2 foi determinar as funções desempenho-intensidade para o material de fala do do teste HINT na versão do Português Brasileiro na presença de ruído estável e modulado a 10 Hz e suas inclinações em adultos nativos do Português do Brasil com audição normal. Foram geradas curvas psicométricas de desempenho de fala em função da intensidade para o mascaramento estável e modulado. O objetivo foi verificar a hipótese de

que a inclinação da curva psicométrica em presença de ruído modulado é menor do que a inclinação da curva psicométrica em presença de ruído estável e, por conseguinte, que a magnitude do *masking release* difere em função da RFR.

Participaram deste experimento 10 adultos jovens (5 do sexo masculino e 05 do sexo feminino), com idades entre 19-24 anos (média de 20,4 anos), todos falantes nativos do Português do Brasil e apresentaram audição normal (limiares tonais ≤ 20 dB para as frequências de oitava 250-8000 Hz). Nenhum participante relatou história de doença otológica ou neurológica. Todos os sujeitos assinaram um TCLE para a participação no estudo.

O material foi o teste HINT na versão do Português Brasileiro (Bevilacqua et al., 2008). Esta versão do HINT consiste em 12 listas de 20 sentenças por lista. As sentenças, gravadas originalmente no HEI, nos Estados Unidos.

O ruído mascarante utilizado foi o ruído fornecido com o teste HINT na versão do Português Brasileiro. Este ruído tem o mesmo formato espectral em longo prazo do espectro de fala médio das sentenças que compreende o material teste. O ruído foi apresentado continuamente sob duas condições: estável e modulado em amplitude. Na condição estável, o ruído foi apresentado em intensidade fixa de 65 dB NPS. Na condição de modulação, o ruído foi modulado por uma onda quadrática entre 65 dB NPS e 30 dB NPS, a uma taxa de modulação em 10 Hz (Desloge et al., 2010; Advíncula et al., 2013).

Para cada condição (ruído estável e ruído modulado), uma lista completa do HINT (20 sentenças) foi apresentada em seis RFR. Para o ruído estável, as RFR foram -14, -11, -8, -5, -2, e 1 dB; para o ruído modulado, as relações foram -23, -20, -17, -14, -11, e -8 dB. Essas relações foram selecionadas com base em estudo piloto como sendo ideais para capturar a porção de forte inclinação da curva psicométrica de desempenho em função da intensidade (obtidas na maioria dos adultos jovens com audição normal). As sentenças foram enviadas à orelha direita dos participantes pelo fone de ouvido Sennheiser HD580, através de uma plataforma de processamento de sinal digital (Tucker-Davis RX6) e sob o controle de um computador que executa um *script* Matlab™ personalizado. Os sujeitos foram testados em cabina acústica e instruídos a repetir em voz alta o máximo possível do que havia percebido de cada sentença, mesmo que a sentença percebida não parecesse fazer sentido gramatical ou semântico. Fora da cabina, o pesquisador monitorou a resposta oral do sujeito através de fones de ouvido ligados a um microfone posicionado dentro da cabine. À medida que cada sentença ia sendo apresentada ao sujeito, o texto da sentença era apresentado simultaneamente na tela do computador na frente do experimentador, com todas as palavras destacadas num retângulo

sombreado sensível a marcação. O experimentador codificou os erros, clicando com mouse do computador sobre as palavras que foram omitidas ou repetidas incorretamente. O programa do computador registrou e rastreou esses erros por palavras e calculou um percentual de acertos no final de cada lista. Cada sujeito ouviu 12 listas selecionadas aleatoriamente: uma lista para cada uma das seis RFR, para cada um dos dois tipos de mascaramento. O desempenho do reconhecimento de fala em presença de ruído estável foi a primeira medida calculada. Para cada tipo de ruído, o teste começou em RFR mais altas.

A proposta do Experimento 3 foi determinar os limiares de relação fala-ruído nos diferentes graus de compressão temporal do material de fala do teste HINT na versão do Português Brasileiro em adultos nativos do Português do Brasil com audição normal. O objetivo foi testar a hipótese de que em ouvintes com audição normal, a magnitude do *masking release* diminui à medida que a RFR aumenta.

Participaram desse experimento 30 adultos jovens (23 do sexo feminino) com idade variando de 17 -25 anos (média de 21,5 anos). Nenhum dos participantes participou do Experimento 1. Todos eram falantes nativos do Português do Brasil e apresentaram audição normal (limiares tonais ≤ 20 dB para as frequências de oitava 250-8000 Hz na orelha de teste), com exceção de um sujeito que apresentou limiar auditivo de 25 dB NA em 8000 Hz. Nenhum sujeito relatou história de doença otológica ou neurológica. Todos concordaram em participar do estudo através da assinatura do TCLE.

Os estímulos de fala utilizados neste experimento foram as sentenças do teste HINT na versão do Português Brasileiro. As sentenças foram apresentadas tanto em seu formato original, sem compressão (compressão do tempo [CT] = 0%), ou em dois níveis de compressão, onde um terço (TC = 33%) ou metade (TC = 50%) da sentença tinham sido removidos. A compressão da fala foi realizada usando o algoritmo proprietário iZotope Radius no Adobe Audition™ que especifica uma mudança estipulada de duração da forma de onda, mantendo o realismo da fala. Tal como acontece com o Experimento 1, o ruído mascarante tinha o mesmo espectro de frequência das sentenças originais. O ruído estável foi apresentado em uma intensidade fixa de 65 dB NPS e o ruído foi modulado em amplitude apresentou modulações realizadas por uma onda quadrática entre 65- e 30-dB SPL, a uma taxa de modulação de 10 Hz. Os estímulos foram enviados através de uma plataforma digital de processamento de sinal (RX6, Tucker-Davis Technologies) e apresentado à orelha direita através de um fone auditivo Sennheiser HD580.

Os limiares de reconhecimento de sentença foram medidos utilizando um procedimento adaptativo que converge para 71% de acerto. Os sujeitos foram testados em uma cabina acústica e orientados a repetir cada sentença da forma como era percebida. Tal como acontece com o Experimento 1 e 2, o texto da sentença foi apresentado simultaneamente na tela do computador na frente do experimentador, 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, a sentença foi dada em uma pontuação global de “correto” ou “incorreto”, na qual a sentença completa teve de ser repetida com precisão para uma pontuação de “correta” e qualquer erro resultou em uma pontuação de “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. A faixa de estimação do limiar continuou até 6 reversões em direção nível do limiar. A estimativa do limiar foi calculada como a média dos quatro níveis (intensidades) finais de reversão. Para cada participante, três estimativas de limiares foram obtidas para uma dada condição de mascaramento, com uma estimativa adicional obtida se o intervalo entre as 3 primeiras fossem igual o maior que 3 dB. O limiar final foi calculado como a média de todas as estimativas obtidas. O primeiro limiar foi determinado com um ruído estável, mas os demais foram randomizados de acordo com o tipo de ruído. Pelo fato das sentenças terem sido apresentadas sem repetição, um único sujeito pode participar de apenas um nível de CT (0%, 33%, ou 50%); ou seja, foi preciso dez indivíduos para cada nível CT. Isso porque o conjunto de oito estimativas de limiares (4 estimativas em cada um dos ruídos) aproximou-se do número máximo de sentenças do teste (sem que o sujeito escutasse qualquer sentença mais de uma vez). O procedimento adaptativo, incluindo a apresentação do estímulo, foi controlado por um *script* Matlab™ personalizado.

6 RESULTADOS

Os resultados da pesquisa originaram dois artigos científicos originais:

6.1. Artigo original 1: Effect of modulation rate on masking release for speech publicado no periódico Audiology - Communication Research (ACR), no volume 18, número 4, out./dez. 2013 (APÊNDICE C).

6.2. Artigo original 2: Modulation masking release using the Brazilian-Portuguese HINT: Psychometric functions and the effect of speech time compression, a ser submetido ao periódico International Journal of Audiology.

5.2 Artigo Original 2: Modulation masking release using the Brazilian-Portuguese HINT: Psychometric functions and the effect of speech time compression (a ser submetido ao periódico *International Journal of Audiology*). Versão em Português.

Benefício do Mascaramento Modulado utilizando o HINT na versão do Português Brasileiro: Função psicométrica e o efeito da compressão temporal da fala

Resumo

Este estudo utilizou o Hearing in Noise Test (HINT) na versão Português Brasileiro para investigar o benefício do mascaramento modulado em amplitude no reconhecimento de fala. O objetivo foi determinar se o benefício do mascaramento modulado variou em função do limiar de reconhecimento de fala em presença de ruído, aqui expresso como relação fala-ruído (RFR). O primeiro experimento mediu a função desempenho-intensidade em presença de ruído estável (com espectro de fala) e modulado (por uma onda quadrática a uma taxa de 10 Hz). Os resultados mostraram curvas psicométrica menos íngremes quando o ruído era modulado. O segundo experimento mediu as relações fala-ruído em função da compressão da fala em tempo (0%, 33% e 50% de compressão) em ambos os tipos de ruído. Os limiares aumentaram com o aumento da compressão da fala em tempo em ambas as condições, no entanto, foi mais acentuado para o ruído modulado. Portanto, o benefício do mascaramento modulado diminuiu com o aumento da compressão temporal da fala. Este padrão de resultado sustenta a hipótese de que, em condições de teste, a redução do benefício do mascaramento modulado pode ser devido a RFR elevadas.

Abstract

This study used the Brazilian-Portuguese Hearing In Noise Test (HINT) to investigate the benefit to speech recognition of listening in a fluctuating background. The goal was to determine whether this benefit, expressed as modulation *masking release*, varied as a function of the speech-to-masker ratio at threshold. The first experiment measured performance-intensity functions in both a steady speech-shaped noise masker and a 10-Hz square-wave

modulated masker. The results showed that the slope of the performance-intensity function was shallower in the modulated masker than in the steady masker. The second experiment measured speech-to-masker ratios at threshold as a function of time-compression of the speech (0, 33, and 50%) in both the steady and modulated maskers. Thresholds increased with increasing time-compression in both maskers, but more markedly in the modulated masker. The derived modulation masking release therefore declined with increasing time-compression. This pattern of results supports the hypothesis that, under some conditions, failure to benefit from masker modulation could be due to elevated speech-to masker ratios at threshold.

Introdução

O reconhecimento de fala em presença de ruído mascarante é geralmente mais preciso se o ruído de mascaramento for modulado do que se for estável (Miller & Licklider, 1950). Este benefício pode ser quantificado como *masking release*, que é a diferença entre os limiares de reconhecimento de fala em presença de mascaramento estável e mascaramento modulado. Uma sugestão de tradução que propomos para esse termo é: *benefício causado pela modificação do ruído mascarante* (Advíncula et al., 2013). No entanto, no presente estudo, e para fins didáticos, utilizamos o termo na língua inglesa. A magnitude do *masking release* depende de uma variedade de fatores, incluindo o tipo de material de fala, a natureza da modulação do mascaramento e a relação fala-ruído (RFR). Esses aspectos tem sido descrito na literatura. No entanto, nenhuma pesquisa ainda havia utilizado a fala do Português Brasileiro como estímulo. Recentemente, uma versão do Português Brasileiro do Hearing in Noise Test (HINT) foi publicada por pesquisadores brasileiros (Bevilacqua et al., 2008), e uma motivação fundamental para este estudo foi caracterizar o *masking release* utilizando este teste brasileiro. O principal objetivo deste artigo foi determinar os limiares de reconhecimento de fala nas diferentes relações fala-ruído nos diferentes graus de compressão temporal do material de fala do HINT na versão do Português Brasileiro em adultos nativos do Português do Brasil com audição normal. Em caso afirmativo, determinar se estas mudanças dos limiares de RFR resultaram numa mudança sistemática na magnitude do *masking release*. Neste estudo, fica caracterizado que a relação fala-ruído (RFR) corresponde à diferença, em dB, entre o valor do limiar de reconhecimento de fala na presença de ruído mascarante e o valor da intensidade do ruído competitivo utilizado.

Uma variedade de fatores relacionada ao estímulo de teste interfere na magnitude do *masking release*. Um deles é o tipo de material de fala utilizado. Uma amostragem de material de fala que tem sido utilizada para avaliar o *masking release* inclui: estímulos vogal-consoante-vogal (Fullgrabe et al., 2006; Gnansia et al., 2008), sílabas sem sentido (Dubno et al., 2003; Bernstein et al., 2012), palavras monossilábicas (Miller & Licklider, 1950; Stuart & Phillips, 1996), palavras espondaicas. (Dirks & Bower, 1971), e sentenças (Jin & Nelson, 2006; Desloge et al., 2010). Um segundo fator que pode influenciar na magnitude do *masking release* é a natureza do ruído mascarante. Esses ruídos mascarantes geralmente são ruídos com espectro de fala ou a própria fala competitiva de uma única pessoa ou de várias juntas (Festen & Plomp, 1990; Gustafsson & Arlinger, 1994; Oxenham & Simonson, 2009; Francart et al., 2011). No caso de ruídos modulados (nesse caso não se trata de falantes), os padrões de modulação geralmente são modulações no envelope de fala ou modulações em ondas (senoidal ou quadrática). Além disso, ciclos regulares e irregulares foram incorporados a esses padrões de flutuação (Stuart & Phillips, 1996; George et al., 2006). Um terceiro fator relacionado ao estímulo que afeta a magnitude do *masking release* é a relação fala-ruído (RFR). Vários estudos têm demonstrado que, para os ouvintes com audição normal, o *masking release* diminui com o aumento da RFR (Oxenham & Simonson, 2009; Christiansen & Dau, 2012; Smits & Festen, 2013). Esse efeito é percebido quando se observa as inclinações (*slopes*) das curvas psicométricas do reconhecimento de fala em presença de ruído mascarante estável e em presença de ruído mascarante modulado. Essas curvas não são paralelas. A curva psicométrica para reconhecimento de fala em presença de ruído modulado é tipicamente menos íngreme do que a curva psicométrica para reconhecimento de fala em presença de ruído estável (Dirks & Bower, 1971; Bernstein & Grant, 2009; Oxenham & Simonson, 2009).

Bernstein e colaboradores (Bernstein & Grant, 2009; Bernstein & Brungart, 2011; Bernstein et al., 2012) e Smits e Festen (2013) levantaram a hipótese de que essa dependência do *masking release* da RFR é a razão na qual a magnitude do *masking release* entre indivíduos de audição normal e indivíduos com perda auditiva coclear difere. Vários estudos têm demonstrado que ouvintes com perda coclear exibem reduzido *masking release* (Festen & Plomp, 1990; Peters et al., 1998; Jin & Nelson, 2006; Lorenzi et al., 2006; Christiansen & Dau, 2012). Vários são as razões descritas, incluindo a audibilidade reduzida (Desloge et al.,

2010) e a baixa resolução temporal (Festen, 1987) de ouvintes com perda auditiva. No entanto, Bernstein e colaboradores (2012) argumentam que a diferença é em grande parte devido ao fato de que os indivíduos com perda auditiva coclear geralmente percebem a fala alvo em uma RFR mais favorável do que os ouvintes com audição normal. Portanto, as diferentes inclinações das curvas psicométricas, poderia explicar a redução de magnitude do *masking release* em indivíduos com perda auditiva. Eles acreditam nessa hipótese demonstrando que, se ouvintes com audição normal percebem a fala em condições que resultam em maiores RFR, o *masking release* dos ouvintes também declina concomitantemente. Manipulações que resultam em alteradas RFR incluem: a quantidade de opções de respostas (*response set size*) (Buss et al., 2009; Bernstein et al., 2012), o uso da fala filtrada como fala-alvo (Oxenham & Simonson, 2009; Christiansen & Dau, 2012), e o uso de falantes não nativos como fala alvo (Calandruccio et al., 2014). Oxenham e Simonson (2009) apontaram que qualquer manipulação que reduza a redundância da fala deve resultar em um aumento no limiar de RFR. Se essa afirmativa estiver correta, um efeito semelhante deve ser observado quando a redundância da fala é alterada por meio da compressão temporal. O principal objetivo deste estudo é testar esta hipótese.

A fala comprimida em tempo é uma manipulação que tem sido usada em uma variedade de contextos, como por exemplo, na determinação de fontes de déficits de processamento da fala relacionadas com a idade (Gordon-Salant & Fitzgibbons, 2001; Jenstad & Souza, 2007; Grose et al., 2009). A quantificação dessa compressão é normalmente expressa pela proporção do conteúdo da forma da onda que é removido. Por exemplo, 33% de compressão implica em dizer que um terço do tempo da onda original foi removido. Os algoritmos de compressão tipicamente removem segmentos da sentença falada, com o objetivo de aumentar a velocidade de fala, sem afetar significativamente entonação ou gerar outras distorções (através da remoção de períodos completos de *pitch* de vogais sonoras e através da remoção de lacunas entre as palavras). O resultado é que a redundância da fala comprimida é reduzida e o limiar de reconhecimento de fala é aumentado, particularmente quando em presença de ruído. A hipótese do presente trabalho, portanto, é que a os limiares de RFR aumentam em função da compressão temporal da fala e, simultaneamente, que a magnitude do *masking release* diminui. No entanto, antes de testar esta hipótese (experimento principal - Exp. 2), um estudo preliminar foi realizado para verificar se o HINT na versão do Português-Brasilero apresentaria a dependência esperada da inclinação (slope) na curva psicométrica do

desempenho-intensidade de acordo com as características do mascaramento (estável e modulado).

Experimento 1. Curva Psicométrica de Desempenho-Intensidade para o HINT na versão do Português Brasileiro em ruído estável e modulado

A proposta deste experimento foi gerar curvas psicométricas de desempenho de fala em função da intensidade para o material de fala do HINT na versão do Português Brasileiro em condições de mascaramento estável e modulado. O objetivo foi verificar a hipótese principal do principal experimento (Exp. 2), de que a inclinação da curva psicométrica em presença de ruído modulado é menor do que a inclinação da curva psicométrica em presença de ruído estável e, por conseguinte, que a magnitude do *masking release* difere em função da RFR.

Método

Sujeitos

Os participantes desse primeiro experimento foram 10 adultos jovens (5 mulheres), com idades entre 19 -24 anos (média de 20,4 anos). Todos eram falantes nativos do Português do Brasil e apresentaram audição normal (limiões tonais ≤ 20 dB para as frequências de oitava 250-8000 Hz). Nenhum participante relatou história de doença otológica ou neurológica. Todos os sujeitos assinaram um Termo de Consentimento Livre e Esclarecido (TCLE) para a participação no estudo. O projeto foi aprovado pelo Comitê Nacional de Ética em Pesquisa (Comissão Nacional de Ética los Pesquisa - CONEP - (CAAE): 02466612.2.0000.5208).

Estímulo

O material de teste foi o HINT na versão do Português Brasileiro (Bevilacqua et al., 2008). Esta versão do HINT consiste em 12 listas de 20 sentenças por lista. As sentenças, gravadas originalmente no *House Research Institute*, nos Estados Unidos.

O ruído mascarante utilizado foi o ruído fornecido com o teste HINT na versão do Português Brasileiro. Este ruído tem o mesmo formato espectral em longo prazo do espectro de fala

médio das sentenças que compreende o material teste. O ruído foi apresentado continuamente sob duas condições: estável e modulado em amplitude. Na condição estável, o ruído foi apresentado em intensidade fixa de 65 dB NPS. Na condição de modulação, o ruído foi modulado por uma onda quadrática entre 65 dB NPS e 30 dB NPS, a uma taxa de modulação em 10 Hz (Desloge et al., 2010; Advíncula et al., 2013).

Procedimentos

Para cada condição (ruído estável X ruído modulado), uma lista completa do HINT (20 sentenças) foi apresentada em seis RFR. Para o ruído estável, as RFR foram -14, -11, -8, -5, -2, e 1 dB; para o ruído modulado, as relações foram -23, -20, -17, -14, -11, e -8 dB. Essas relações foram selecionadas com base em estudo piloto como sendo ideais para capturar a porção de forte inclinação da curva psicométrica de desempenho em função da intensidade (obtidas na maioria dos adultos jovens com audição normal). As sentenças foram enviadas à orelha direita dos participantes pelo fone de ouvido Sennheiser HD580, através de uma plataforma de processamento de sinal digital (Tucker-Davis RX6) e sob o controle de um computador que executa um *script* Matlab™ personalizado. Os sujeitos foram testados em cabina acústica e instruídos a repetir em voz alta o máximo possível do que havia percebido de cada sentença, mesmo que a sentença percebida não parecesse fazer sentido gramatical ou semântico. Fora da cabina, o pesquisador monitorou a resposta oral do sujeito através de fones de ouvido ligados a um microfone posicionado dentro da cabine. À medida que cada sentença ia sendo apresentada ao sujeito, o texto da sentença era apresentado simultaneamente na tela do computador na frente do experimentador, com todas as palavras destacadas num retângulo sombreado sensível a marcação. O experimentador codificou os erros, clicando com mouse do computador sobre as palavras que foram omitidas ou repetidas incorretamente. O programa do computador registrou e rastreou esses erros por palavras e calculou um percentual de acertos no final de cada lista. Cada sujeito ouviu 12 listas selecionadas aleatoriamente: uma lista para cada uma das seis RFR, para cada um dos dois tipos de mascaramento. O desempenho do reconhecimento de fala em presença de ruído estável foi a primeira medida calculada. Para cada tipo de ruído, o teste começou em RFR mais altas.

Resultados e Discussão

Os resultados são apresentados na Fig. 1. Cada painel mostra os dados de um sujeito, plotando o percentual de reconhecimento correto por palavras contabilizadas em função dos limiares de RFR tanto para o ruído estável (círculos vazios) como para o ruído modulado (quadrados cheios). O último painel mostra a média dos dados. Os dados foram ajustados com as funções logísticas que minimizaram a soma dos quadrados de erro, e estas funções são mostradas em cada painel. A RFR associada com o ponto de 50% de acerto foi identificada, assim como a inclinação da curva. Para o ruído estável, a média dos limiares de RFR em 50% de acerto foi de -8,8 dB (desvio padrão [SD] = 0,86 dB). Este limiar é menor do que geralmente é encontrado quando o teste usa como fala alvo palavras simples (monossilábicas ou espondeicas) (Wilson et al., 2008), mas deve-se destacar que o desempenho de reconhecimento de fala medido neste estudo foi dentro de um contexto de sentença. Boothroyd and Nittrouer (1988) demonstraram que existe uma relação previsível entre a probabilidade de identificar corretamente a palavra e o significado do contexto na qual é apresentada, e que essa relação varia com a RFR. Para o ruído modulado, a média dos limiares de RFR foi de -18,9 dB (DP = 2,1 dB). A diferença em dB entre o ponto de 50% de acerto (no desempenho) identificado com ruído estável versus o ruído modulado foi denominado como a magnitude do *masking release*. A média do *masking release* (dos 10 sujeitos) foi em torno de 10 dB (DP = 1,25). Um teste-t pareado simples indicou que a diferença entre a magnitude do *masking release* entre as duas condições foi significativa ($t[9] = 20.87; p < 0.001$).

O principal interesse deste estudo foi encontrar as diferentes inclinações das curvas psicométricas entre os ruídos estáveis e modulados. Para o ruído estável, a inclinação média foi em torno de 16,2%/dB. Esta é mais acentuada do que o valor de 11,4%/dB medido durante a fase de desenvolvimento do HINT na versão do Português Brasileiro (Bevilacqua et al., 2008), ou o valor médio de 10,3%/dB das 13 versões do HINT publicadas em diferentes línguas (Soli & Wong, 2008). No entanto, estes últimos valores são derivados de funções de desempenho-intensidade medidas por cada sentença correta, enquanto que as curvas apresentadas na Fig. 1 são medidas com base no acerto de palavras (mesmo que dentro do contexto da sentença). Para o ruído modulado, a inclinação média foi em torno de 7,3%/dB. Um Teste-T pareado simples indicou que a inclinação para o ruído estável foi significativamente mais acentuada do que a do ruído modulado ($t [9] = 5,98, p < 0,001$). Esta diferença na inclinação mostra a dependência da magnitude do *masking release* do ponto da

curva psicométrica no qual o limiar é medido. Esses resultados confirmam a suposição subjacente no Experimento 2, de que a magnitude do *masking release* varia com uma função da RFR.

Experimento 2. Benefício do Mascaramento Modulado em função da Compressão da Fala em Tempo

A proposta do Experimento 2 foi variar os limiares de RFR através da manipulação do nível de redundância da fala nas sentenças do HINT na versão do Português Brasileiro, usando a compressão temporal da fala alvo. O objetivo foi testar a hipótese de que em ouvintes com audição normal, a magnitude do *masking release* diminui à medida que a RFR aumenta.

Método

Sujeitos

Participaram desse experimento 30 adultos jovens (23 do sexo feminino) com idade variando de 17 -25 anos (média de 21,5 anos). Nenhum dos participantes participou do Experimento 1. Todos eram falantes nativos do Português do Brasil e apresentaram audição normal (limiares tonais ≤ 20 dB para as frequências de oitava 250-8000 Hz na orelha de teste), com exceção de um sujeito que apresentou limiar auditivo de 25 dB NA em 8000 Hz. Nenhum sujeito relatou história de doença otológica ou neurológica. Todos concordaram em participar do estudo através da assinatura do TCLE. O projeto foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos do Centro de Ciências da Saúde da UFPE: 02466612.2.0000.5208.

Estímulo

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

fala. Um estudo piloto indicou que as sentenças comprimidas nesse estudo foram percebidas (pelos sujeitos) como uma fala rápida tendo atribuído o mesmo *pitch* que as sentenças originais, e foram livres de distorções perceptíveis. Tal como acontece com o Experimento 1, o ruído mascarante tinha o mesmo espectro de frequência das sentenças originais. O ruído estável foi apresentado em uma intensidade fixa de 65 dB SPL e o ruído foi modulado em amplitude apresentou modulações realizadas por uma onda quadrática entre 65 - e 30-dB SPL, a uma taxa de modulação de 10 Hz. Os estímulos foram enviados através de uma plataforma digital de processamento de sinal (RX6, Tucker-Davis Technologies) e apresentado à orelha direita através de um fone de ouvido Sennheiser HD580.

Procedimentos

Os limiares de reconhecimento de sentença foram medidos utilizando um procedimento adaptativo que converge para 71% de acerto. Os sujeitos foram testados em uma cabina acústica, e orientados a repetir cada sentença da forma como era percebida. Tal como acontece com o Experimento 1, o texto da sentença foi apresentado simultaneamente na tela do computador na frente do experimentador, com todas as palavras em destaque num retângulo sombreado sensível à marcação. O pesquisador usou o mouse do computador para marcar as palavras que foram omitidas ou repetidas incorretamente. No entanto, para a proposta do procedimento adaptativo, a sentença foi dada em uma pontuação global de “correto” ou “incorreto”, na qual a sentença completa teve de ser repetida com precisão para uma pontuação de “correta” e qualquer erro resultou em uma pontuação de “incorreta”. Após duas sentenças corretas, o nível de apresentação da próxima sentença era reduzido em 2 dB; após um sentença incorreta, o nível de apresentação da próxima sentença era aumentado em 2 dB. A faixa de estimação do limiar continuou até 6 reversões em direção nível do limiar. A estimativa do limiar foi calculada como a média dos quatro níveis (intensidades) finais de reversão. Para cada participante, três estimativas de limiares foram obtidas para uma dada condição de mascaramento, com uma estimativa adicional obtida se o intervalo entre as 3 primeiras fossem igual o maior que 3 dB. O limiar final foi calculado como a média de todas as estimativas obtidas. O primeiro limiar foi determinado com um ruído estável, mas os demais foram randomizados de acordo com o tipo de ruído. Pelo fato das sentenças terem sido apresentadas sem repetição, um único sujeito pode participar de apenas um nível de CT (0%, 33%, ou 50%); ou seja, foi preciso dez indivíduos para cada nível CT. Isso porque o conjunto de oito estimativas de limiares (4 estimativas em cada um dos ruídos) aproximou-se do

número máximo de sentenças do teste (sem que o sujeito escutasse qualquer sentença mais de uma vez). O procedimento adaptativo, incluindo a apresentação do estímulo, foi controlado por um *script* Matlab™ personalizado.

Resultados e Discussão

Os resultados são apresentados na Fig. 2, onde os limiares de reconhecimento de fala são plotados para cada nível de compressão do tempo. Os limiares individuais medidos com ruído mascarante estável são mostrados como círculos cheios, e os medidos com ruído mascarante modulado são mostrados como círculos vazios; para cada participante, uma barra vertical conecta os respectivos limiares estáveis e modulados. Os limiares médios do grupo para o ruído mascarante estável são apresentados como quadrados cheios, e os limiares encontrados em presença de ruído mascarante modulado são mostrados como quadrados vazios (barras de erro de 1 DP).

Os resultados obtidos com a fala não comprimida (CT = 0%), mostram uma média do limiar de reconhecimento de fala para o ruído estável em 61,4 dB (DP = 1,4 dB). Isso produz um limiar de RFR de -3,6 dB. Esse resultado se compara favoravelmente com a RFR de -4,6 dB encontrada para o percentil dde 50 para a fala e ruído no HINT na versão Português Brasileiro, publicado por Bevilacqua et al. (2008), e para o -3,9 dB que representa a média em 13 versões linguísticas do HINT (Soli & Wong, 2008). A média do limiar de reconhecimento de fala medido com ruído modulado foi 52,7 dB (DP = 2,1 dB). A magnitude média do *masking release* foi, portanto, de 8,7 dB. Para CT = 33%, a média do limiar de reconhecimento de fala no ruído estável foi de 66,0 dB (DP = 1,3 dB), enquanto que o limiar no ruído modulado foi de 61,7 dB (DP = 1,4 dB), produzindo um *masking release* de 4,3 dB. Para CT = 50%, a média do limiar de reconhecimento de fala no ruído estável foi de 70,3 dB (DP = 2,1 dB) e, no ruído modulado, foi de 67,7 dB (DP = 1,4), resultando em 2,6 dB de *masking release*.

Este padrão de dados sugere três aspectos interessantes: (1) para todos os sujeitos, os limiares no ruído modulado são melhores do que no ruído estável, significando a ocorrência do

masking release; (2) os limiares geralmente pioram quando a quantidade de CT aumenta, e (3) a elevação do limiar com o aumento da CT é maior para o ruído modulado do que para o ruído estável, resultando numa redução do *masking release* com o aumento de CT. Para avaliar isso, os dados foram submetidos à análise de variância (ANOVA) com um fator entre sujeitos (nível CT) e um fator intra-sujeitos (tipo do ruído). A análise revelou um efeito significativo do nível de CT ($F[2,27] = 172,9$; $p < 0,001$), um efeito significativo do tipo de ruído mascarante ($F[1,27] = 313,3$; $p < 0,001$) e uma interação significativa entre esses fatores ($F[2,27] = 37,3$; $p < 0,001$). Comparações emparelhadas planejadas indicaram que o efeito do tipo de ruído foi significativo em cada nível de nível de CT ($p < 0,001$). Este resultado indica que a os limiares de RFR aumentaram com o aumento da compressão do tempo para ambos os ruídos (estável e modulado). No entanto, a RFR aumentou mais intensamente para o ruído modulado do que para o ruído estável. Para determinar se a magnitude do *masking release* diminuiu com o aumento da compressão do tempo, outro teste de variância-ANOVA foi realizado em todos os níveis de CT. A análise revelou um efeito significativo do nível de CT ($F[2,27] = 37,3$; $p < 0,001$) e contrastes lineares indicaram que a magnitude do BMM foi significativamente diferente em cada um dos três níveis de compressão do tempo ($p < 0,05$).

Os resultados do Experimento 2, portanto, suportam a hipótese de que a magnitude do *masking release* para a fala depende dos limiares de RFR. No entanto, deve ser alertado de que esta conclusão baseia-se no pressuposto de que as funções de desempenho-intensidade são independentes do nível de CT. (As curvas psicométricas apresentadas no Experimento 1 foram realizadas apenas com a fala não comprimida). É improvável que essa suposição seja completamente válida. Por exemplo, Hosoi et al. (1999) mediram as funções de desempenho-intensidade para sentenças Japonesas sem sentido no silêncio e descobriu que, enquanto as curvas psicométricas de desempenho-intensidade para a fala não comprimida (CT = 0%) e para fala comprimida em um terço (CT = 33%) foram sobrepostas, a curva para fala comprimida em 50% tende a ser menos íngreme. Ciente dessa ressalva, os presentes achados são, no entanto, de acordo com a noção de que o *masking release* reduzido observado em ouvintes com perda auditiva coclear pode ser devido a elevados limiares de RFR associados com os ouvintes com perda (Bernstein & Grant, 2009; Bernstein & Brungart, 2011; Bernstein et al., 2012; Smits & Festen, 2013). Ainda é uma questão de debate atual saber se este é o único fator que contribui para o *masking release* reduzido em ouvintes com perda coclear (Christiansen & Dau, 2010). O que é claro, contudo, é que esse mecanismo para a redução do

masking release não é aplicável a todos os casos de reduzidos *masking release*. Por exemplo, Grose et al. (2009) comparou o *masking release* em ouvintes mais jovens e mais velhos com audiogramas relativamente normais e encontrou reduzido *masking release* nos ouvintes mais velhos. No entanto, as RFR não diferiram entre os dois grupos etários para o ruído estável; o reduzido *masking release* foi inteiramente devido aos limiares elevados de RFR com a presença de ruído modulado. Portanto, possivelmente mecanismos adicionais estão envolvidos na redução do *masking release* para fala relacionada à idade.

Conclusões

O objetivo deste estudo foi medir o *masking release* utilizando o HINT na versão do Português Brasileiro com um foco particular sobre a determinação da dependência do *masking release* dos limiares de RFR. O Experimento 1 determinou as curvas de desempenho- intensidade em ambos ruídos estável e modulado a 10 Hz. Os resultados demonstraram que a inclinação da curva psicométrica foi menos íngreme no ruído modulado que no ruído estável, confirmando a hipótese de que a magnitude do *masking release* depende da RFR. O objetivo do Experimento 2 foi manipular os limiares de RFR através da compressão temporal da fala alvo. Os resultados confirmaram que os limiares de RFR aumentaram com o aumento da compressão em tempo da fala alvo, porém mais ainda no ruído modulado que no ruído estável. Como consequência, a magnitude do *masking release* diminuiu com o aumento da compressão de tempo. Essa constatação apoia a noção de que a redução do *masking release* em ouvintes com perda auditiva coclear é devido, em parte, a elevados limiares de RFR, e esse fator parece não ser o motivo do reduzido *masking release* encontrado em ouvintes mais velhos com audiogramas relativamente normais.

Agradecimentos

Este trabalho foi apoiado por subsídios R03DC012278 e R01DC001507 do NIH NIDCD.

Referências

- Advincula, K.P., Menezes, D.C., and Pacífico, F.A., Griz, S.M.S. (2013) "Percepção da fala em presença de ruído competitivo: o efeito da taxa de modulação do ruído mascarante" *Audiol., Commun. Res.* 18(4):238-244.
- Bernstein, J. G., and Brungart, D. S. (2011). "Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio," *J Acoust Soc Am* **130**, 473-488.
- Bernstein, J. G., and Grant, K. W. (2009). "Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners," *J Acoust Soc Am* **125**, 3358-3372.
- Bernstein, J. G., Summers, V., Iyer, N., and Brungart, D. S. (2012). "Set-size procedures for controlling variations in speech-reception performance with a fluctuating masker," *J Acoust Soc Am* **132**, 2676-2689.
- Bevilacqua, M. C., Banhara, M. R., Da Costa, E. A., Vignoly, A. B., and Alvarenga, K. F. (2008). "The Brazilian Portuguese hearing in noise test," *Int J Audiol* **47**, 364-365.
- Boothroyd, A., and Nittrouer, S. (1988). "Mathematical treatment of context effects in phoneme and word recognition," *J Acoust Soc Am* **84**, 101-114.
- Buss, E., Whittle, L. N., Grose, J. H., and Hall, J. W., 3rd (2009). "Masking release for words in amplitude-modulated noise as a function of modulation rate and task," *The Journal of the Acoustical Society of America* **126**, 269-280.
- Calandruccio, L., Buss, E., and Hall, J. W. (2014). "Effects of linguistic experience on the ability to benefit from temporal and spectral masker modulation."
- Christiansen, C., and Dau, T. (2012). "Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise," *J Acoust Soc Am* **132**, 1655-1666.
- Desloge, J. G., Reed, C. M., Braidá, L. D., Perez, Z. D., and Delhorne, L. A. (2010). "Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise," *J Acoust Soc Am* **128**, 342-359.

- Dirks, D. D., and Bower, D. R. (1971). "Influence of pulsed masking on spondee words," *J Acoust Soc Am* **50**, 1204-1207.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2003). "Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing," *J Acoust Soc Am* **113**, 2084-2094.
- Festen, J. M. (1987). "Speech-reception threshold in fluctuating background sound and its possible relation to temporal auditory resolution," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhoff, Dordrecht, The Netherlands).
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *Journal of the Acoustical Society of America* **88**, 1725-1736.
- Francart, T., van Wieringen, A., and Wouters, J. (2011). "Comparison of fluctuating maskers for speech recognition tests," *Int J Audiol* **50**, 2-13.
- Fullgrabe, C., Berthommier, F., and Lorenzi, C. (2006). "Masking release for consonant features in temporally fluctuating background noise," *Hear Res* **211**, 74-84.
- George, E. L., Festen, J. M., and Houtgast, T. (2006). "Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners," *J Acoust Soc Am* **120**, 2295-2311.
- Gnansia, D., Jourdes, V., and Lorenzi, C. (2008). "Effect of masker modulation depth on speech masking release," *Hear Res* **239**, 60-68.
- Gordon-Salant, S., and Fitzgibbons, P. J. (2001). "Sources of age-related recognition difficulty for time-compressed speech," *J Speech Lang Hear Res* **44**, 709-719.
- Grose, J. H., Mamo, S. K., and Hall, J. W., 3rd (2009). "Age effects in temporal envelope processing: speech unmasking and auditory steady state responses," *Ear Hear* **30**, 568-575.
- Gustafsson, H. A., and Arlinger, S. D. (1994). "Masking of speech by amplitude-modulated noise," *Journal of the Acoustical Society of America* **95**, 518-529.

- Hosoi, H., Tsuta, Y., Nishida, T., Murata, K., Ohta, F., Mekata, T., and Kato, Y. (1999). "Variable-speech-rate audiometry for hearing aid evaluation," *Auris, nasus, larynx* **26**, 17-27.
- Jenstad, L. M., and Souza, P. E. (2007). "Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults," *J Speech Lang Hear Res* **50**, 1123-1138.
- Jin, S. H., and Nelson, P. B. (2006). "Speech perception in gated noise: the effects of temporal resolution," *J Acoust Soc Am* **119**, 3097-3108.
- Lorenzi, C., Husson, M., Ardoint, M., and Debrulle, X. (2006). "Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception," *Int J Audiol* **45**, 487-495.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *Journal of the Acoustical Society of America* **22**, 167-173.
- Oxenham, A. J., and Simonson, A. M. (2009). "Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference," *J Acoust Soc Am* **125**, 457-468.
- Peters, R. W., Moore, B. C., and Baer, T. (1998). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," *Journal of the Acoustical Society of America* **103**, 577-587.
- Smits, C., and Festen, J. M. (2013). "The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: II. Fluctuating noise," *J Acoust Soc Am* **133**, 3004-3015.
- Soli, S. D., and Wong, L. L. (2008). "Assessment of speech intelligibility in noise with the Hearing in Noise Test," *Int J Audiol* **47**, 356-361.
- Stuart, A., and Phillips, D. P. (1996). "Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners," *Ear Hear* **17**, 478-489.
- Wilson, R. H., McArdle, R., and Roberts, H. (2008). "A comparison of recognition performances in speech-spectrum noise by listeners with normal hearing on PB-50,

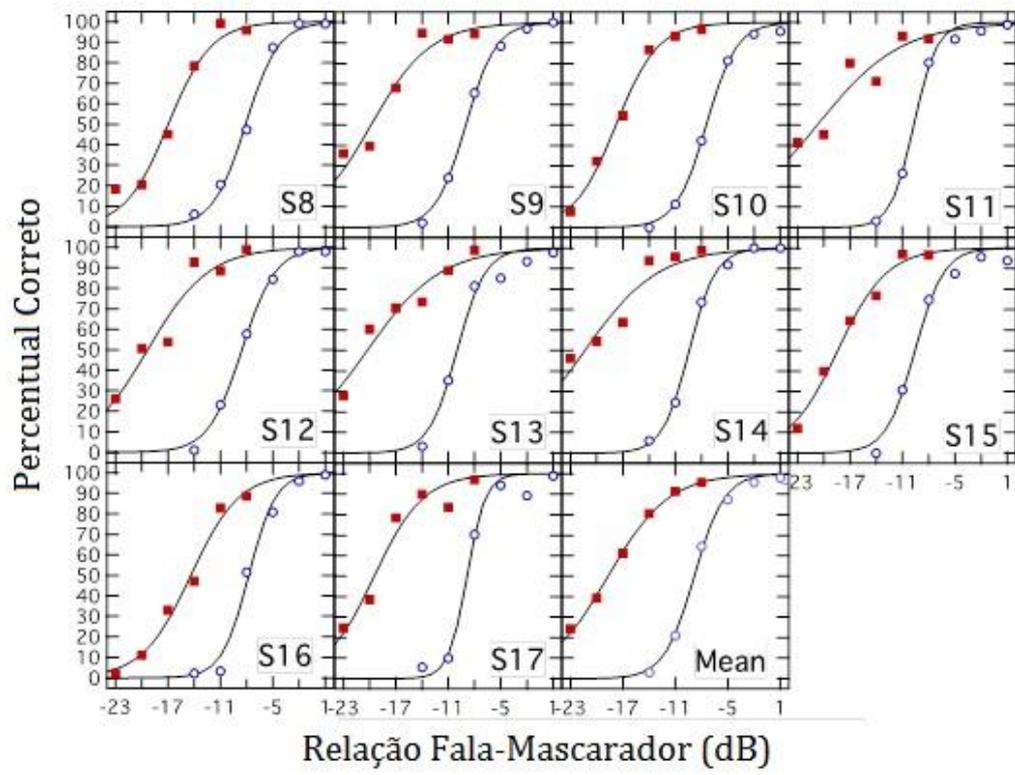
CID W-22, NU-6, W-1 spondaic words, and monosyllabic digits spoken by the same speaker," *J Am Acad Audiol* **19**, 496-506.

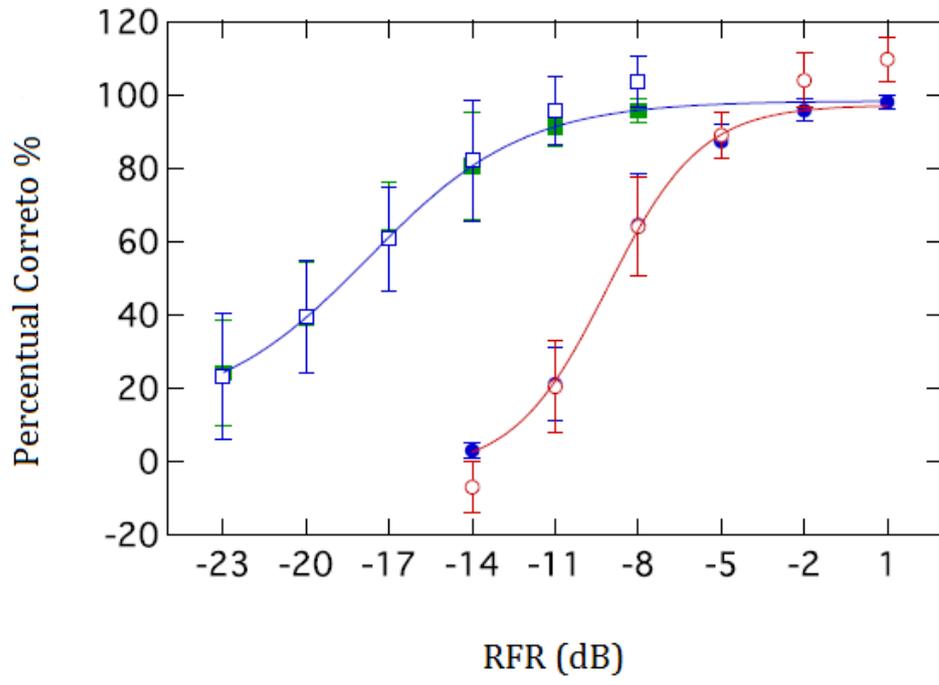
Legendas das Figuras

Fig. 1. Percentual de palavras reconhecidas corretamente em função das relações fala-ruído medido em um ruído estável (círculos vazios) e um ruído modulado (quadrados cheios). Cada painel mostra os dados a partir de um sujeito, com o último painel apresenta os dados médios. Também são mostradas melhores funções logísticas de ajuste.

Fig. 2. Média das funções psicométricas obtidas a partir dos percentuais de palavras reconhecidas corretamente em função das relações fala-ruído medido em um ruído estável (círculos vazios) e um ruído modulado.

Fig. 3. Limiares de reconhecimento de fala plotados para cada nível de compressão do tempo. Limiares individuais medidos no ruído estável (círculos cheios) e ruído modulado (círculos vazios) são conectados com uma barra vertical. Limiares médios dos grupos também são mostrados para o ruído estável (quadrados cheios) e ruído modulado (quadrados vazios) com barras de erro 1-DP.





5.2 Artigo Original 2: Modulation masking release using the Brazilian-Portuguese HINT: Psychometric functions and the effect of speech time compression (a ser submetido ao periódico International Journal of Audiology). Versão em Inglês.

Modulation masking release using the Brazilian-Portuguese HINT: Psychometric functions and the effect of speech time compression.

Abstract

This study used the Brazilian-Portuguese Hearing In Noise Test (HINT) to investigate the benefit to speech recognition of listening in a fluctuating background. The goal was to determine whether this benefit, expressed as modulation masking release, varied as a function of the speech-to-masker ratio at threshold. The first experiment measured performance-intensity functions in both a steady speech-shaped noise masker and a 10-Hz square-wave modulated masker. The results showed that the slope of the performance-intensity function was shallower in the modulated masker than in the steady masker. The second experiment measured speech-to-masker ratios at threshold as a function of time-compression of the speech (0, 33, and 50%) in both the steady and modulated maskers. Thresholds increased with increasing time-compression in both maskers, but more markedly in the modulated masker. The derived modulation masking release therefore declined with increasing time-compression. This pattern of results supports the hypothesis that, under some conditions, failure to benefit from masker modulation could be due to elevated speech-to masker ratios at threshold.

Introduction

Speech recognition in noise is usually more acute if the masking noise is modulated than if it is steady (Miller & Licklider, 1950). This benefit can be quantified as modulation masking release [MMR], which is the difference in speech recognition threshold measured in the

steady and modulated maskers. The magnitude of MMR depends on a variety of stimulus factors – expanded on below – including the type of speech material, the nature of the modulating masker, and the speech-to-masker ratio. However, we begin by noting that none of this research has been done using Brazilian-Portuguese speech. Recently, a Brazilian-Portuguese version of the Hearing In Noise Test (HINT) has become available (Bevilacqua et al., 2008), and an underlying motivation for this study was to characterize MMR using this new (non-English) test. The primary research focus, however, was to determine the dependence of MMR on speech-to-masker ratio using the novel stimulus manipulation of varying the time compression of the target speech.

A variety of stimulus factors affect the magnitude of MMR. One factor is the type of speech material used. A sampling of speech material that has been used to assess the benefit of a modulated masker includes: vowel-consonant-vowel stimuli (e.g., Fullgrabe et al., 2006; Gnansia et al., 2008), nonsense syllables (e.g., Dubno et al., 2003; Bernstein et al., 2012), monosyllabic words (e.g., Miller & Licklider, 1950; Stuart & Phillips, 1996), spondee words (e.g., Dirks & Bower, 1971), and sentences (e.g., Jin & Nelson, 2006; Desloge et al., 2010). A second stimulus factor that influences MMR is the nature of the masker. Maskers range from noise (usually speech-shaped noise [SSN]) to single- or multi-talker competitors (e.g., Festen & Plomp, 1990; Gustafsson & Arlinger, 1994; Oxenham & Simonson, 2009; Francart et al., 2011). For modulated noise (non-talker) maskers, the modulation patterns range from single- or multi-talker speech envelopes to sinusoidal or square-wave modulators. In addition, both regular and irregular duty cycles have been incorporated into these fluctuation patterns (Stuart & Phillips, 1996; George et al., 2006). A third stimulus factor that affects the magnitude of MMR is the speech-to-masker ratio at threshold. Several studies have demonstrated that, for normal-hearing listeners, the benefit of masker modulation decreases as the speech-to-masker ratio increases (Oxenham & Simonson, 2009; Christiansen & Dau, 2012; Smits & Festen, 2013). The basis for this effect is that the slopes of the psychometric functions for speech recognition in steady and modulated maskers are not parallel. Rather, the psychometric function for speech recognition in a modulated masker is typically shallower than that in a steady masker (e.g., Dirks & Bower, 1971; Bernstein & Grant, 2009; Oxenham & Simonson, 2009).

Bernstein and colleagues (Bernstein & Grant, 2009; Bernstein & Brungart, 2011; Bernstein et al., 2012) and Smits and Festen (2013) have hypothesized that this dependence of MMR on the speech-to-masker ratio at threshold is the predominant reason for the difference in MMR magnitude between listeners with normal hearing and listeners with cochlear hearing loss. Several studies have demonstrated that listeners with cochlear loss exhibit reduced benefit from masker modulation (e.g., Festen & Plomp, 1990; Peters et al., 1998; Jin & Nelson, 2006; Lorenzi et al., 2006; Christiansen & Dau, 2012). A number of reasons for this have been suggested including reduced audibility (Desloge et al., 2010) and poor temporal resolution (Festen, 1987). However, Bernstein and colleagues argue that the difference is largely due to the fact that the listeners with cochlear hearing loss generally receive the target speech at a higher speech-to-masker ratio at threshold than do listeners with normal hearing. Given the differing slopes of the psychometric functions, this leads to a reduced MMR magnitude in the impaired listeners. They support their hypothesis by demonstrating that if normal-hearing listeners receive speech under conditions that result in higher speech-to-masker ratios at threshold, the listeners' MMR concomitantly declines. Manipulations that have resulted in altered speech-to-masker ratios at threshold include varying response set size (Buss et al., 2009; Bernstein et al., 2012), filtering the speech (Oxenham & Simonson, 2009; Christiansen & Dau, 2012), and testing non-native speakers of the target speech (Calandruccio et al., 2014). Oxenham and Simonson (2009) pointed out that any manipulation that reduced speech redundancy should result in an increased speech-to-masker ratio at threshold. If correct, a similar effect should be observed when speech redundancy is manipulated by means of time compression. The main purpose of this study is to test this hypothesis.

Time-compression of speech is a manipulation that has been used in a variety of contexts such as determining sources of age-related speech processing deficits (e.g., Gordon-Salant & Fitzgibbons, 2001; Jenstad & Souza, 2007; Grose et al., 2009). Quantification of time compression is usually expressed in terms of the proportion of the time-waveform content that is excised. For example, 33%-time-compression implies that one-third of the original time waveform has been removed. Compression algorithms typically remove segments of the spoken phrase with the goal of increasing speech rate without markedly affecting pitch contour or generating other distortions (e.g., by removing complete pitch periods of voiced vowels and portions of the inter-word gaps). The net result is that the redundancy of the time-compressed speech is reduced and the speech recognition threshold is increased, particularly

in masking. The hypothesis of this study, therefore, is that the speech-to-masker ratio at threshold increases as a function of time compression and, concomitantly, that the magnitude of MMR decreases. However, prior to testing this hypothesis in the main experiment (Exp. 2), a preliminary experiment was undertaken to verify that the new Brazilian-Portuguese HINT exhibits the expected dependency of the slope of the performance-intensity function on the modulation characteristics of the masker.

Exp. 1. Performance-Intensity functions for Brazilian-Portuguese HINT in steady and modulated noise

The purpose of this experiment was to generate performance-intensity functions for Brazilian-Portuguese speech under conditions of steady and modulated masking. The goal was to verify the assumption underlying the main experiment (Exp. 2) that the slope of the performance-intensity function measured in a modulated masker is shallower than that measured in a steady masker and, therefore, that the magnitude of MMR differs as a function of speech-to-masker ratio.

Method

Subjects

The participants were 10 young adults (5 female) ranging in age from 19 -24 yrs (mean 20.4 yrs). All were native speakers of Brazilian Portuguese and had normal hearing as documented by pure-tone thresholds ≤ 20 dB HL for the octave frequencies 250 – 8000 Hz. None reported any history of otologic or neurologic disorder. Each subject gave signed consent for participation in the study, which was approved by the Brazilian National Committee of Ethics in Research (*Comitê de ética em Pesquisa com Seres Humano do Centro de Ciências da Saúde da UFPE*: 02466612.2.0000.5208).

Stimuli

The test material was the Brazilian-Portuguese HINT (Bevilacqua et al., 2008). This version of the HINT consists of 12 lists of 20 sentences per list. The sentences, originally recorded at the House Research Institute, U.S.A., were resampled to 24,414 Hz and scaled to have equal RMS levels across all sentences.

The masker was the SSN supplied with the HINT. This noise has the same spectral shape as the long-term average speech spectrum (LTASS) of the sentences comprising the test material. The masker was presented continuously under two conditions: steady and modulated. In the steady condition, the masker was presented at a constant level of 65 dB SPL. In the modulated condition, the masker was square-wave modulated between 65 dB SPL and 30 dB SPL at a rate of 10 Hz (cf. Desloge et al., 2010).

Procedure

For each masker type, 1 complete HINT list (20 sentences) was presented at each of 6 speech-to-masker ratios. For the steady masker, the speech-to-masker ratios were -14, -11, -8, -5, -2, and 1 dB; for the modulated masker, the ratios were -23, -20, -17, -14, -11, and -8 dB. These ratios were selected based on pilot listening as being optimal for capturing the steeply sloping portion of the performance-intensity function in most normal-hearing young adults. The sentences were output to the right phone of a Sennheiser HD580 headset through a digital signal processing platform (Tucker-Davis RX6) under the control of a computer running a custom Matlab™ script. The subject listened monaurally within a single-walled sound-attenuating booth and was instructed to repeat aloud as much of the perceived sentence as possible, even if the perceived sentence did not appear to make grammatical or semantic sense. Outside the booth, the experimenter monitored the oral response of the subject through headphones linked to a microphone within the booth. As each sentence was presented to the subject, the text of the sentence was simultaneously displayed on the computer screen in front of the experimenter, with each word highlighted in a position-sensitive shaded rectangle. The experimenter coded the errors by clicking the computer mouse on the words that were omitted or repeated incorrectly. The computer program registered and tracked these word-wise errors and computed a percent correct score at the end of each list. Each subject listened to 12 randomly-selected lists: 1 list for each of 6 speech-to-masker ratios for each of 2 masker

types. Performance in the steady masker was measured first and, for each masker type, testing began at one of the higher speech-to-masker ratios.

Results and Discussion

The results are displayed in Fig. 1. Each panel shows the data from one subject, plotting percent correct *at the word level* against speech-to-masker ratio for both the steady masker (unfilled circles) and the modulated masker (filled squares). The final panel displays the mean data. The data were fitted with logistic functions that minimized the sum of squares error, and these functions are shown in each panel. From these functions, the speech-to-masker ratio associated with the estimated 50% correct point was identified, as well as the slope of the function. For the steady masker, the mean speech-to-masker ratio at the 50% threshold was -8.8 dB (standard deviation [SD] = 0.86 dB). This threshold is lower than that is usually measured for single words, such as monosyllabic or spondaic words (Wilson et al., 2008), but recall that the word-wise recognition performance measured here was within a sentence context. Boothroyd and Nittrouer (1988) have shown that there is a predictable relationship between the probability of correctly identifying a word and the meaningfulness of the context in which it is presented, and that this relationship varies with speech-to-masker ratio. For the modulated masker, the speech-to-masker ratio at threshold was -18.9 dB (SD = 2.1 dB). The dB difference between the 50% correct points measured in the steady and modulated maskers was taken as the magnitude of MMR. Across the ten subjects, the average MMR was about 10 dB (SD = 1.52 dB). A paired-sample t-test indicated that the magnitude of MMR was significant ($t[9] = 20.87$; $p < 0.001$).

Of primary interest to this study was the finding that the slopes of the psychometric functions differed between the steady and modulated maskers. For the steady masker, the mean slope was about 16.2%/dB. This is steeper than the 11.4%/dB measured during the development phase of the Brazilian-Portuguese HINT (Bevilacqua et al., 2008), or the value of 10.3%/dB averaged across 13 language versions of the HINT (Soli & Wong, 2008). However, these latter values are derived from performance-intensity functions measured at the sentence level whereas the functions in Fig. 1 are measured at the word level (within sentence context). For

the modulated masker, the mean slope was about 7.3%/dB. A paired-sample t-test indicated the slope for the steady masker was significantly steeper than that of the modulated masker ($t[9] = 5.98$, $p < 0.001$). This difference in slope implicates a dependence of the derived MMR on the point on the psychometric function at which threshold is measured. This expected finding confirms the assumption underlying Exp. 2 that the magnitude of MMR varies as a function of the signal-to-masker ratio at threshold at which performance is measured.

Exp. 2. Modulation masking release as a function of speech time compression

The purpose of Exp. 2 was to vary the speech-to-masker ratio at threshold by systematically manipulating the level of speech redundancy in the Brazilian-Portuguese HINT sentences using time compression. The goal was to test the hypothesis that, in normal-hearing listeners, the magnitude of MMR decreases as the speech-to-masker ratio at threshold increases.

Method

Subjects

The participants were 30 young adults (23 female) ranging in age from 17 -25 yrs (mean 21.5 yrs). None of the participants had taken part in Exp. 1. All were native speakers of Brazilian Portuguese and had normal hearing as documented by pure-tone thresholds ≤ 20 dB HL for the octave frequencies 250 – 8000 Hz in the test ear, except for one subject who had a threshold of 25 dB HL at 8000 Hz. None reported any history of otologic or neurologic disorder. Each subject gave signed consent for participation in the study, which was approved by the Brazilian National Committee of Ethics in Research (*Comitê de ética em Pesquisa com Seres Humano do Centro de Ciências da Saúde da UFPE: 02466612.2.0000.5208*).

Stimuli

The speech stimuli used in this experiment were the sentences of the Brazilian-Portuguese HINT. The sentences were either presented in their original, uncompressed format (time

compression [TC] = 0%), or at two levels of compression where either one-third (TC = 33%) or half (TC = 50%) of the sentence waveform had been removed. Time compression was undertaken off-line using the proprietary iZotope Radius algorithm within Adobe Audition™ that specifies a stipulated change in waveform duration while maintaining speech realism. Pilot listening indicated that the resulting time-compressed sentences were perceived as rapid speech having the same pitch attributes as the original sentences, and were otherwise free of noticeable distortions. As with Exp. 1, the masker was a noise with the same LTASS as the original sentences. The masker was either output continuously at a level of 65 dB SPL or was square-wave modulated between 65- and 30-dB SPL at a rate of 10 Hz. Stimuli were output via a digital signal processing platform (RX6, Tucker-Davis Technologies) and presented to the right ear through a Sennheiser HD580 headphone.

Procedure

Speech recognition thresholds were measured using an adaptive procedure that converged on the 71% correct point. The subject was seated inside a single-walled, sound-attenuating booth and repeated back each sentence as it was presented. As with Exp. 1, the text of the sentence was simultaneously displayed on the computer screen in front of the experimenter, with each word highlighted in a position-sensitive shaded rectangle. The experimenter used the computer mouse to tag the words that were omitted or repeated incorrectly. However, for the purposes of the adaptive procedure, the sentence was given an overall score of ‘correct’ or ‘incorrect’ wherein the complete sentence had to be repeated accurately for a score of ‘correct’ and any error resulted in a score of ‘incorrect.’ Following two correct sentences in a row, the presentation level of the next sentence was reduced by 2 dB; following one incorrect sentence, the presentation level of the next sentence was increased by 2 dB. A threshold estimation track continued until 6 reversals in level direction had occurred. The threshold for that track was taken as the mean of the final 4 reversal levels. For each subject, three threshold estimates were obtained for a given masker condition, with an additional estimate obtained if the range of the first 3 exceeded 3 dB. The final threshold for that condition was taken as the mean of all estimates obtained. The first threshold estimation track was always undertaken with a steady masker, but the masker type for subsequent tracks was quasi-randomized across replications. Because sentences were presented without repetition, a single subject could participate in only one level of TC (0, 33, or 50%); i.e., 10 subjects per TC

level. This was because the collection of 6 - 8 threshold estimates (up to 4 estimates in each of the steady and modulated maskers) approached the maximum number of threshold tracks that could be obtained without the subject hearing any given sentence more than once. The adaptive procedure, including stimulus presentation, was controlled by a custom Matlab™ script.

Results and Discussion

The results are displayed in Fig. 2 where speech recognition thresholds are plotted for each level of time compression. Individual thresholds measured in the steady masker are shown as filled circles and those measured in the modulated masker are shown as unfilled circles; for each subject, a vertical bar connects the respective steady and modulated thresholds. Group mean thresholds for the steady masker are shown as filled squares while those for the modulated masker are shown as unfilled squares (1 SD error bars).

Dealing first with the uncompressed speech (TC = 0%), the mean speech recognition threshold measured in the steady masker was 61.4 dB (SD = 1.4 dB). This yields a speech-to-masker ratio at threshold of -3.6 dB. This compares favorably with the speech-to-masker ratio of -4.6 dB tabulated for the 50th percentile performance for co-located speech and noise for the Brazilian-Portuguese HINT found by Bevilacqua et al. (2008) and to the -3.9 dB which represents the average across 13 language versions of the HINT (Soli & Wong, 2008). The mean speech recognition threshold measured in the modulated masker was 52.7 dB (SD = 2.1 dB). The average magnitude of MMR was therefore 8.7 dB. For TC = 33%, the average speech recognition threshold in the steady masker was 66.0 dB (SD = 1.3 dB), whereas the threshold in the modulated masker was 61.7 dB (SD = 1.4 dB), yielding a MMR of 4.3 dB. For TC = 50%, the average speech recognition threshold in the steady masker was 70.3 dB (SD = 2.1 dB) and, in the modulated masker, was 67.7 dB (SD = 1.4) resulting in 2.6 dB of MMR.

This data pattern therefore suggests three features: (1) for all subjects, thresholds in the modulated masker are lower than in the steady masker signifying the occurrence of MMR; (2) thresholds generally worsen as the amount of TC increases; and (3) the elevation in threshold with increasing TC is greater for the modulated masker than for the steady masker, resulting in a reduction in MMR with increasing TC. To assess this, the data were submitted to an analysis of variance (ANOVA) with one between-subjects factor (*TC level*) and one within-subjects factor (*masker type*). The analysis indicated revealed a significant effect of *TC level* ($F[2,27] = 172.9$; $p < 0.001$), a significant effect of *masker type* ($F[1,27] = 313.3$; $p < 0.001$), and a significant interaction between these factors ($F[2,27] = 37.3$; $p < 0.001$). Planned pairwise comparisons indicated that the effect of *masker type* was significant at each level of *TC level* ($p < 0.001$). This result indicates that the speech-to-masker ratio required for criterion speech recognition increased with increasing time compression for both the steady and modulated maskers. However, the speech-to-masker ratio increased more rapidly for the modulated masker than for the steady masker. To determine whether the derived MMR decreased with increasing time compression, a second ANOVA was undertaken on the derived MMR magnitudes across TC level. The analysis revealed a significant effect of *TC level* ($F[2,27] = 37.3$; $p < 0.001$), and linear contrasts indicated that MMR magnitude was significantly different across each of the three levels of time compression ($p < 0.05$).

The results of Exp. 2 therefore support the hypothesis that the magnitude of MMR for speech depends on the speech-to-masker ratio at threshold. However, it should be cautioned that this conclusion is based on the assumption that the performance-intensity functions are independent of TC level. (Performance-intensity functions measured in Exp. 1 were for the uncompressed speech only.) It is unlikely that this assumption is completely valid. For example, Hosoi et al. (1999) measured performance-intensity functions for nonsense Japanese phrases in quiet and found that, whereas the performance-intensity functions for uncompressed (TC = 0%) and TC = 33% speech overlapped, the function for TC = 50% speech tended to be shallower. Cognizant of this caveat, the present findings are nevertheless in line with the notion that the reduced MMR observed in listeners with cochlear hearing loss could be due to the elevated speech-to-masker ratios at threshold associated with the impaired listeners (Bernstein & Grant, 2009; Bernstein & Brungart, 2011; Bernstein et al., 2012; Smits & Festen, 2013). Whether this is the only factor contributing to the reduced MMR in cochlear loss listeners is a matter of current debate (Christiansen & Dau, 2012). What is clear,

however, is that this mechanism for reduced MMR is not applicable to all cases of reduced benefit of masker modulation. For example, Grose et al. (2009) compared MMR in younger and older listeners with relatively normal audiograms and found reduced MMR in the older listeners. However, the speech-to-masker ratios at threshold did not differ between the two age groups for the steady masker; the reduced MMR was entirely due to elevated speech-to-masker thresholds in the modulated masker. Thus, age-related reductions in speech MMR are likely to involve additional mechanisms. In this context, we are currently investigating age-related differences in temporal masking as contributing to the MMR effect.

Summary

The goal of this study was to measure MMR using the Brazilian-Portuguese HINT with a particular focus on determining the dependence of MMR on speech-to-masker ratio at threshold. Exp. 1 measured performance-intensity functions in both a steady and 10-Hz modulated noise having the same LTASS as the target speech material. The results demonstrated that the slope of the function was shallower in the modulated masker than in the steady masker, confirming the assumption that the derived magnitude of MMR depends on the speech-to-masker ratio at threshold. The purpose of Exp. 2 was to adjust the speech-to-masker ratio at threshold using the novel stimulus manipulation of varying the time-compression of the target speech. The results confirmed that speech-to-masker ratio at threshold increased with increasing time-compression of the target speech, but more so in the modulated masker than in the steady masker. As a consequence, the magnitude of MMR decreased with increasing time-compression. Whereas this finding supports the notion that reduced MMR in listeners with cochlear hearing loss is due, in part, to elevated speech-to-masker ratios at threshold, this factor appears not to account for the reduced MMR in older listeners with relatively normal audiograms.

Acknowledgments

This work was supported by grants R03DC012278 and R01DC001507 from the NIH NIDCD.

Bibliography

- Advincula, K.P., Menezes, D.C., and Pacífico, F.A., Griz, S.M.S. (2013) "Percepção da fala em presença de ruído competitivo: o efeito da taxa de modulação do ruído mascarante" *Audiol., Commun. Res.* 18(4):238-244.
- Bernstein, J. G., and Brungart, D. S. (2011). "Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio," *J Acoust Soc Am* **130**, 473-488.
- Bernstein, J. G., and Grant, K. W. (2009). "Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners," *J Acoust Soc Am* **125**, 3358-3372.
- Bernstein, J. G., Summers, V., Iyer, N., and Brungart, D. S. (2012). "Set-size procedures for controlling variations in speech-reception performance with a fluctuating masker," *J Acoust Soc Am* **132**, 2676-2689.
- Bevilacqua, M. C., Banhara, M. R., Da Costa, E. A., Vignoly, A. B., and Alvarenga, K. F. (2008). "The Brazilian Portuguese hearing in noise test," *Int J Audiol* **47**, 364-365.
- Boothroyd, A., and Nittrouer, S. (1988). "Mathematical treatment of context effects in phoneme and word recognition," *J Acoust Soc Am* **84**, 101-114.
- Buss, E., Whittle, L. N., Grose, J. H., and Hall, J. W., 3rd (2009). "Masking release for words in amplitude-modulated noise as a function of modulation rate and task," *The Journal of the Acoustical Society of America* **126**, 269-280.
- Calandruccio, L., Buss, E., and Hall, J. W. (2014). "Effects of linguistic experience on the ability to benefit from temporal and spectral masker modulation."
- Christiansen, C., and Dau, T. (2012). "Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise," *J Acoust Soc Am* **132**, 1655-1666.

- Desloge, J. G., Reed, C. M., Braida, L. D., Perez, Z. D., and Delhorne, L. A. (2010). "Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise," *J Acoust Soc Am* **128**, 342-359.
- Dirks, D. D., and Bower, D. R. (1971). "Influence of pulsed masking on spondee words," *J Acoust Soc Am* **50**, 1204-1207.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2003). "Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing," *J Acoust Soc Am* **113**, 2084-2094.
- Festen, J. M. (1987). "Speech-reception threshold in fluctuating background sound and its possible relation to temporal auditory resolution," in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhoff, Dordrecht, The Netherlands).
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *Journal of the Acoustical Society of America* **88**, 1725-1736.
- Francart, T., van Wieringen, A., and Wouters, J. (2011). "Comparison of fluctuating maskers for speech recognition tests," *Int J Audiol* **50**, 2-13.
- Fullgrabe, C., Berthommier, F., and Lorenzi, C. (2006). "Masking release for consonant features in temporally fluctuating background noise," *Hear Res* **211**, 74-84.
- George, E. L., Festen, J. M., and Houtgast, T. (2006). "Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners," *J Acoust Soc Am* **120**, 2295-2311.
- Gnansia, D., Jourdes, V., and Lorenzi, C. (2008). "Effect of masker modulation depth on speech masking release," *Hear Res* **239**, 60-68.
- Gordon-Salant, S., and Fitzgibbons, P. J. (2001). "Sources of age-related recognition difficulty for time-compressed speech," *J Speech Lang Hear Res* **44**, 709-719.
- Grose, J. H., Mamo, S. K., and Hall, J. W., 3rd (2009). "Age effects in temporal envelope processing: speech unmasking and auditory steady state responses," *Ear Hear* **30**, 568-575.

- Gustafsson, H. A., and Arlinger, S. D. (1994). "Masking of speech by amplitude-modulated noise," *Journal of the Acoustical Society of America* **95**, 518-529.
- Hosoi, H., Tsuta, Y., Nishida, T., Murata, K., Ohta, F., Mekata, T., and Kato, Y. (1999). "Variable-speech-rate audiometry for hearing aid evaluation," *Auris, nasus, larynx* **26**, 17-27.
- Jenstad, L. M., and Souza, P. E. (2007). "Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults," *J Speech Lang Hear Res* **50**, 1123-1138.
- Jin, S. H., and Nelson, P. B. (2006). "Speech perception in gated noise: the effects of temporal resolution," *J Acoust Soc Am* **119**, 3097-3108.
- Lorenzi, C., Husson, M., Ardoint, M., and Debrulle, X. (2006). "Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception," *Int J Audiol* **45**, 487-495.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *Journal of the Acoustical Society of America* **22**, 167-173.
- Oxenham, A. J., and Simonson, A. M. (2009). "Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference," *J Acoust Soc Am* **125**, 457-468.
- Peters, R. W., Moore, B. C., and Baer, T. (1998). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," *Journal of the Acoustical Society of America* **103**, 577-587.
- Smits, C., and Festen, J. M. (2013). "The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: II. Fluctuating noise," *J Acoust Soc Am* **133**, 3004-3015.
- Soli, S. D., and Wong, L. L. (2008). "Assessment of speech intelligibility in noise with the Hearing in Noise Test," *Int J Audiol* **47**, 356-361.
- Stuart, A., and Phillips, D. P. (1996). "Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners," *Ear Hear* **17**, 478-489.

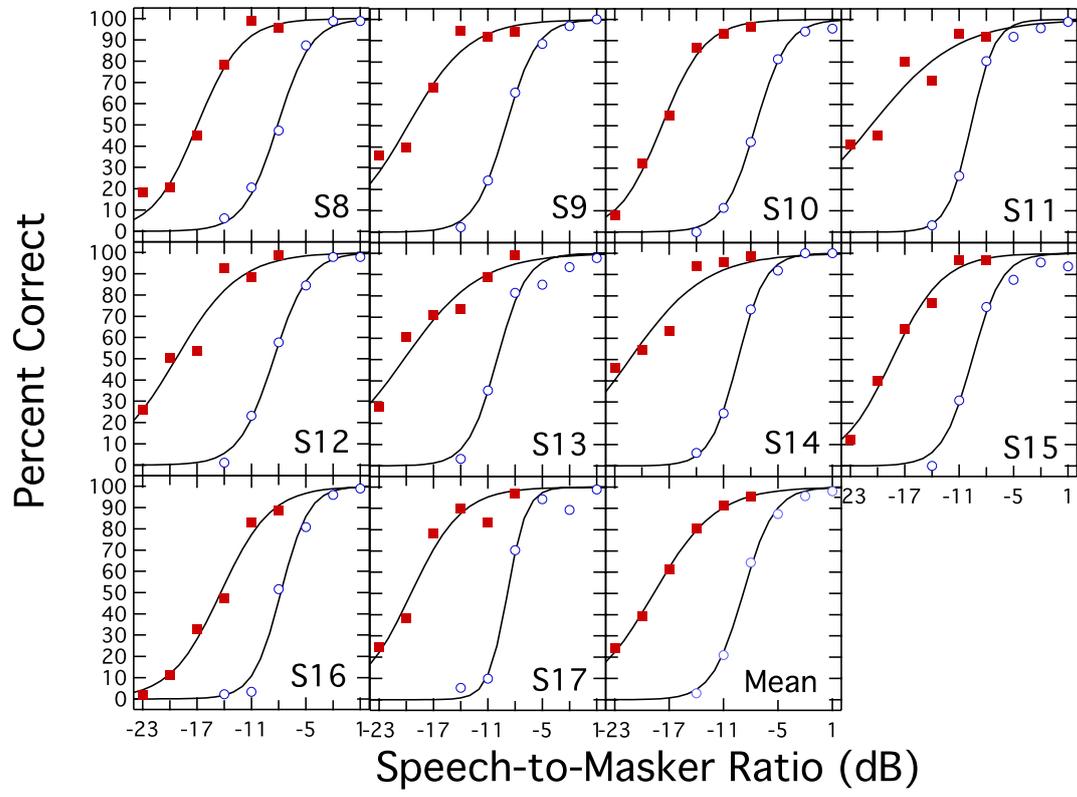
Wilson, R. H., McArdle, R., and Roberts, H. (2008). "A comparison of recognition performances in speech-spectrum noise by listeners with normal hearing on PB-50, CID W-22, NU-6, W-1 spondaic words, and monosyllabic digits spoken by the same speaker," *J Am Acad Audiol* **19**, 496-506.

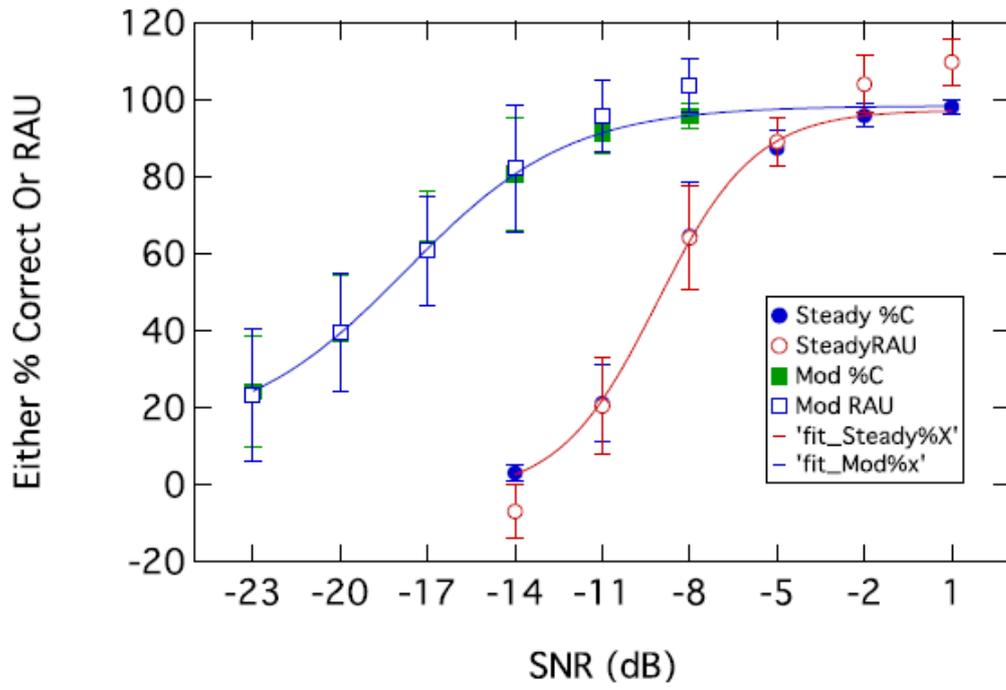
Figure Legends

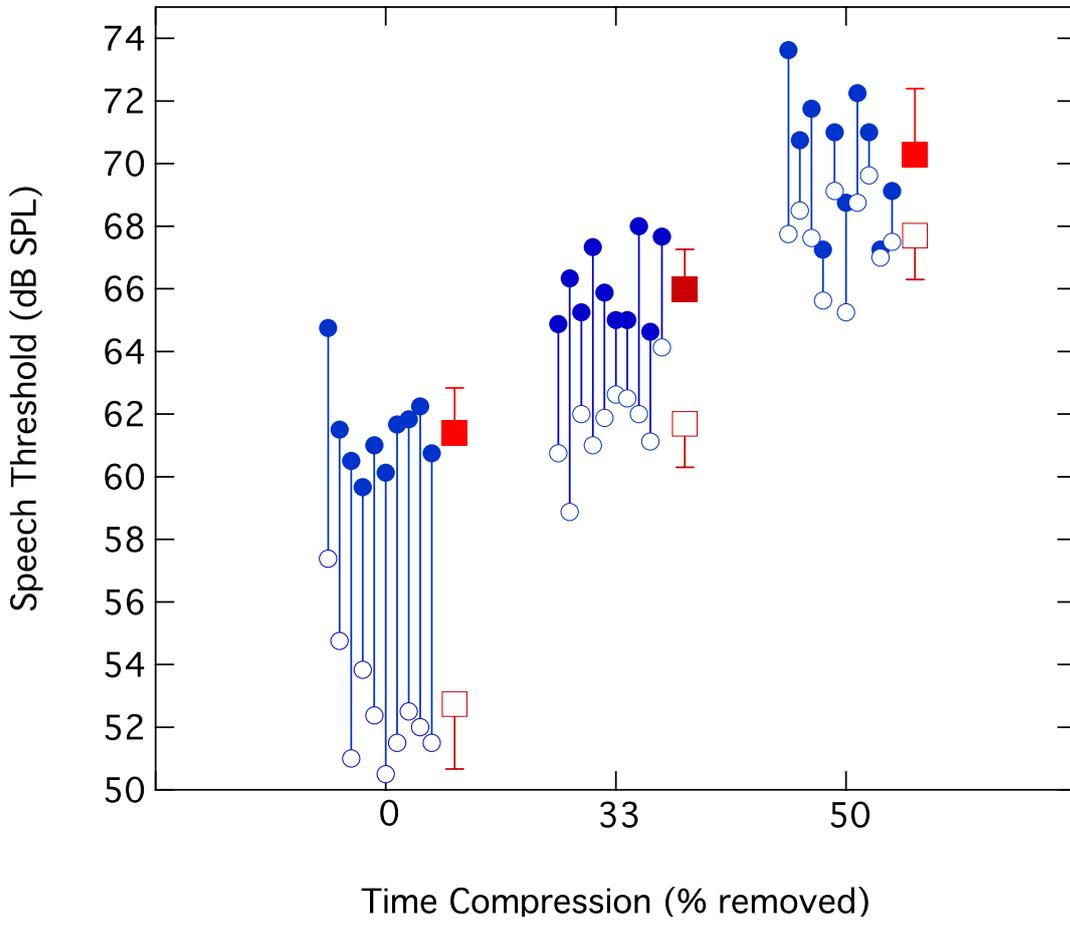
Fig. 1. Percent correct word recognition as a function of speech-to-masker ratio measured in a steady masker (unfilled circles) and a modulated masker (filled squares). Each panel shows the data from one subject, with the final panel displaying the mean data. Also shown are best-fitting logistic functions.

Fig. 2. Average psychometric functions obtained from the percent correct word recognition as a function of speech-to-masker ratio measured in a steady masker (unfilled circles) and a modulated masker (filled squares).

Fig. 3. Speech recognition thresholds plotted for each level of time compression. Individual thresholds measured in the steady masker (filled circles) and modulated masker (unfilled circles) are connected with a vertical bar. Group mean thresholds are also shown for the steady masker (filled squares) and modulated masker (unfilled squares) with 1-SD error bars.







7 CONCLUSÕES

O presente estudo apresenta achados pioneiros no que diz respeito ao benefício causado pela modificação do ruído mascarante (*masking release*) em diferentes situações de escuta. Com este estudo foi possível:

(1) verificar que a magnitude do *masking release* para sentenças do HINT na versão do Português Brasileiro 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;

(2) verificar que a função psicométrica para o reconhecimento da fala com as sentenças do HINT na versão brasileira foi significativamente menos íngreme para o ruído modulado do que para o ruído estável, apresentando um *masking release* de 10 dB que é semelhante à de outros resultados publicados para a fala em inglês. Este achado demonstra que o teste HINT na versão do Português Brasileiro é eficaz para a medição do *masking release*;

(3) verificar que a compressão temporal da fala das sentenças do HINT na versão do Português Brasileiro resultou em uma mudança do limiar de relação fala-ruído, com uma mudança associada no *masking release*. Esses dados mostraram que à medida que o grau de compressão temporal da fala aumentava, o limiar de RFR aumentava para ambas as condições de mascaramento. No entanto, o aumento no limiar de RFR foi relativamente mais íngreme para as condições de mascaramento modulado de tal modo que o *masking release* derivado diminuiu com o aumento da RFR, confirmando assim que a magnitude do *masking release* é dependente da medida de referência de RFR (RFR de base).

O trabalho futuro extenderá esta abordagem aos ouvintes idosos com audição normal, pois, para a fala não comprimida, temos mostrado que seu *masking release* é reduzido, embora a seu limiar de RFR para o ruído estável é equivalente ao de jovens ouvintes.

REFERÊNCIAS

ADVINCULA, K. P.; MENEZES, D. C.; PACÍFICO, F. A.; GRIZ, S. M. S. Percepção da fala em presença de ruído competitivo: o efeito da taxa de modulação do ruído mascarante. **Audiology – Communication Research**, v.18, n.4, p.238-244, 2013.

AMERICAN NATIONAL INSTITUTE (ANSI). American National Standard Specification for Audiometers. ANSI S3.6-2004. New York: ANSI, 2004.

ANDERSON, E. S.; NELSON, D. A.; KREFT, H.; NELSON, P. B.; OXENHAM, A. J. Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users. **The Journal of the Acoustical Society of America**, v.130, n.1, p.364-75, 2011.

ARIETA, A. M. Teste de percepção da fala HINT - Brasil, em normo-ouvintes e usuários de aparelhos auditivos - Atenção à saúde auditiva. 2009. 131 f. Dissertação (Mestrado) – Programa de Pós-Graduação em saúde coletiva. Faculdade de Ciências Médicas da Universidade Estadual de Campinas, Campinas, 2009.

BACON, S. P.; OPIE, J. M., MONTOYA, D. Y. The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds. **Journal of Speech Language and Hearing Research**, v.41, p.549-563, 1998.

BEVILACQUA, M. C.; BANHARA, M. R.; DA COSTA, E. A.; VIGNOLY, A. B.; ALVARENGA, K. F. The Brazilian Portuguese hearing in noise test. **International Journal of Audiology**, v.47, n.6, 364-365, 2008.

BERNSTEIN, J. G.; GRANT, K. W. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. **The Journal of the Acoustical Society of America**, v.125, 3358-3372, 2009.

BERNSTEIN, J. G.; SUMMERS, V.; IYER, N.; BRUNGART, D. S. Set-size procedures for controlling variations in speech-reception performance with a fluctuating masker. **The Journal of the Acoustical Society of America**, 132, 2676-2689, 2012.

BERNSTEIN, J. G.; BRUNGART, D. S. Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. **The Journal of the Acoustical Society of America**, v.130, 473-488, 2011.

BUSS, E.; WHITTLE, L. N.; GROSE, J. H., HALL, J. W. 3RD. Masking release for words in amplitude-modulated noise as a function of modulation rate and task. **The Journal of the Acoustical Society of America**, v.126, p.269-280, 2009.

BODE, D. L.; CARHART, R. Measurements of articulation functions using adaptive test procedures. **IEEE Trans Audio Electroacoust**, v.AU-21, n.3, p.196-201, 1973.

BOOTHROYD, A.; NITTROUER, S. Mathematical treatment of context effects in phoneme and word recognition. **The Journal of the Acoustical Society of America**, v.84, p.101-114, 1988.

BRONKHORST, A. W.; PLOMP, R. A. Clinical test for the assessment of binaural speech perception in noise. **Audiology**, v.29, p.275- 85, 1990.

BRUNGART, D. S. Informational and energetic masking effects in the perception of two simultaneous talkers. **The Journal of the Acoustical Society of America**, v.109, n.3, p.1101-9, 2001.

BRUNGART, D. S.; SIMPSON, B. D.; SCOTT, K. R.; ERICSON, M. A. Informational and energetic masking effects in the perception of multiple simultaneous talkers. **The Journal of the Acoustical Society of America**, v.110, n.5, p.2527-38, 2001.

BUSS, E.; HALL, J. W. 3RD; GROSE, J. H. Spectral integration of synchronous and asynchronous cues to consonant identification. **The Journal of the Acoustical Society of America**, v.115, n.5, p.2278-85, 2004.

BUSS, E.; WHITTLE, L. N.; GROSE, J. H.; HALL, J. W. 3RD. Masking release for words in amplitude-modulated noise as a function of modulation rate and task. **The Journal of the Acoustical Society of America**, v.126, n.1, p.269-80, 2009.

BUSS, E.; HE, S.; GROSE, J. H.; HALL, J. W. 3RD. The monaural temporal window based on masking period pattern data in school-aged children and adults. **The Journal of the Acoustical Society of America**, v.133, n.3, p.1586-97, 2013.

CALANDRUCCIO, L.; BUSS, E.; HALL, J. W. Effects of linguistic experience on the ability to benefit from temporal and spectral masker modulation. 2014.

CAPORALI, S. A.; ARIETA, A. M. Reconhecimento de fala no ruído: estudo comparativo entre grupos com e sem queixa de percepção de fala. **Revista da Sociedade Brasileira de Fonoaudiologia**, v.9, n.3, 2004.

CAPORALI, S. A.; SILVA, J. A. Reconhecimento de fala no ruído em jovens e idosos com perda auditiva. **Revista Brasileira de Otorrinolaringologia**, v.70, n.4, p.525-32, 2004.

CARHART, R.; TILLMAN, T. W. Interaction of competing speech signals with hearing loss. **Archives of Otolaryngology**, v.91, p.273-279, 1970.

CEKIC, S.; SENNAROGLU, G. The Turkish hearing in noise test. **International Journal of Audiology**, 2008; v.47, n.6, p.366-8, 2008.

CHAVES, A. D.; NEPOMUCENO, L. A.; ROSSI, A. G.; MOTA, H. B.; PILLON, L. Reconhecimento de fala: uma descrição de resultados obtidos em função do número de sílabas dos estímulos. **Pró-Fono Revista de Atualização Científica**, v.1, n.1, p:53-8, 1999.

CHRISTIANSEN, C.; DAU, T. Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise. **The Journal of the Acoustical Society of America**, v.132, p.1655-1666, 2012.

COSTA, E. A. Audiometria tonal e testes de reconhecimento de fala: Estudo comparativo para aplicação em audiologia ocupacional. **Acta AWHO**, v.11, n.3, p.137-42, 1992.

COSTA, E. A. Estudo da correlação entre audiometria tonal e o reconhecimento de monossílabos mascarados por fala competitiva nas perdas auditivas induzidas pelo ruído. **Acta AWHO**, 1995.

COSTA, M. J.; IORIO, M. C. M.; MANGABEIRA-ALBERNAZ, P. L.; CABRAL JUNIOR, E. F.; MAGNI, A. B. Desenvolvimento de um ruído com espectro de fala. **Acta AWHO**, v.17, n.2, p.84-89, 1998.

COSTA, E. A. Desenvolvimento de Teste de Reconhecimento da Fala, com ruído, em português do Brasil, para aplicação em audiologia ocupacional. [Tese de Doutorado]. Campinas: Faculdade de Ciências Médicas/Unicamp, 1998.

COSTA, M. J. Lista de sentenças em Português: apresentação e estratégias de aplicação na audiologia. Santa Maria: Pallotti, 1998.

COSTA, M. J.; DANIEL, R. C.; SANTOS, S. N. Reconhecimento de sentenças no silêncio e no ruído em fones auriculares: valores de referência de normalidade. **Revista CEFAC**, v.13, n.4, p.685-691, 2011.

DIRKS, D. D.; BOWER, D. R. Influence of pulsed masking on spondee words. **The Journal of the Acoustical Society of America**, v.50, p.1204-1207, 1971.

DESLOGE, J. G.; REED, C. M.; BRAIDA, L. D., PEREZ, Z. D., DELHORNE, L. A. Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise. **The Journal of the Acoustical Society of America**, v.128, p.342-359, 2010.

DUBNO, J. R.; HORWITZ, A. R.; AHLSTROM, J. B. Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. **The Journal of the Acoustical Society of America**, v.113, p.2084-2094, 2003.

DUNCAN, K. R.; AARTS, N. L. A comparison of the HINT and Quick SIN Tests. **Journal of Speech-Language Pathology and Audiology**, v.30, n.2, 2006.

EISENBERG, L. S.; DIRKS, D. D.; BELL, T. S. Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. **Journal of Speech Language and Hearing Research**, v.38, p.222-233, 1995.

FRANCART, T.; VAN WIERINGEN, A.; WOUTERS, J. Comparison of fluctuating maskers for speech recognition tests. **International Journal of Audiology**, v.50, p.2-13, 2011.

FESTEN, J. M. Speech-reception threshold in fluctuating background sound and its possible relation to temporal auditory resolution, in *The Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Nijhoff, Dordrecht, The Netherlands), 1987.

FESTEN, J. M.; PLOMP, R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. **The Journal of the Acoustical Society of America**, v.88, p.1725-1736, 1990.

FREITAS, C. D.; LOPES, L. F. D.; COSTA, M. J. Confiabilidade dos limiares de reconhecimento de sentenças no silêncio e no ruído. **Revista Brasileira de Otorrinolaringologia**, v.71, n.5, p.624-630, 2005.

FREYMAN, R. L.; BALAKRISHNAN, U.; HELFER, K. S. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. **The Journal of the Acoustical Society of America**, v.115, n.5, p.2246-2256, 2004.

FÜLLGRABE, C.; BERTHOMMIER, F.; LORENZI, C. Masking release for consonant features in temporally fluctuating background noise. **Hearing Research**, v.211, p.74-84, 2006.

GEORGE, E. L.; FESTEN, J. M.; HOUTGAST, T. Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. **The Journal of the Acoustical Society of America**, v.120, p.2295-2311, 2006.

GNANSIA, D.; JOURDES, V.; LORENZI, C. Effect of masker modulation depth on speech masking release. **Hearing Research**, v.239, p.60-68, 2008.

GORDON-SALANT, S.; FITZGIBBONS, P. J. Sources of age-related recognition difficulty for time-compressed speech. **Journal of Speech Language and Hearing Research**, v.44, p.709-719, 2001.

GROSE, J. H.; MAMO, S. K.; HALL, J. W. 3RD. Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. **Ear and Hearing**, v.30, p.568-575, 2009.

GUSTAFSSON, H. A.; ARLINGER, S. D. Masking of speech by amplitude-modulated noise. **The Journal of the Acoustical Society of America**, v.95, n.1, p.518-529, 1994.

GELFAND, S. A. Hearing: An Introduction to Psychological and Physiological Acoustic. 2010.

GIGUÈRE, C.; LAROCHE, C.; SOLI, S. D.; VAILLANCOURT, V. Functionally-based screening criteria for hearing-critical jobs based on the Hearing in Noise Test. **International Journal of Audiology**, v.47, p.319-328, 2008.

GORDON-SALANT, S. Age-related differences in speech recognition performance as a function of test format and paradigm. **Ear and Hearing**, v.8, p.277-282, 1987.

GAMA, M. R. Percepção da fala: uma proposta de avaliação qualitativa. São Paulo: Pancast, 1994.

HAGERMAN, B. Sentences for testing speech intelligibility in noise. **Scandinavian Audiology**, v.11, p.79-87, 1982.

HALL, J. W.; BUSS, E.; GROSE, J. H.; ROUSH, P. A. Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. **Ear and Hearing**, v.33, n.3, p.340-8, 2012.

HARRIS, R. W.; GOFFI, M. V. S.; PEDALINI, M. E. B.; GYGI, M. A. Reconhecimento de palavras dissilábicas psicometricamente equivalentes no português brasileiro faladas por

indivíduos do sexo masculino e do sexo feminino. **Pró-Fono Revista de Atualização Científica**, v.13, n.2, p.249-62, 2001.

HENRIQUES, M. O.; COSTA, M. J. Limiares de reconhecimento de sentenças em indivíduos normo-ouvintes na presença de ruído incidente de diferentes ângulos. **Revista da Sociedade Brasileira de Fonoaudiologia**, v.16, n.1, p.54-8, 2011.

HENRIQUES, M. O.; MIRANDA, E. C.; COSTA, M. J. Limiares de reconhecimento de sentenças no ruído, em campo livre: valores de referência para adultos normo-ouvintes. **Revista Brasileira de Otorrinolaringologia**, v.74, n.2, p.188-92, 2008.

HOSOI, H.; TSUTA, Y.; NISHIDA, T.; MURATA, K.; OHTA, F.; MEKATA, T.; KATO, Y. Variable-speech-rate audiometry for hearing aid evaluation. **Auris, nasus, larynx**, v.26, p.17-27, 1999.

HUARTE, A. The Castilian Spanish in noise test. **International Journal of Audiology**, v.47, n.6, p.369-70, 2008.

JACOB, R. T. S.; MONTEIRO, N. F. G.; MOLINA, S. V.; BEVILACQUA, M. C.; LAURIS, J. R. P.; MORET, A. L. M. Percepção da fala em crianças em situação de ruído. **Arquivos Internacionais de Otorrinolaringologia**, v.15, n.2, p.163-7, 2011.

JENSTAD, L. M.; SOUZA, P. E. Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults. **Journal of Speech Language and Hearing Research**, v.50, p.1123-1138, 2007.

JIN, S. H.; NELSON, P. B. Speech perception in gated noise: the effects of temporal resolution. **The Journal of the Acoustical Society of America**, v.119, p.3097-3108, 2006.

KALIKOW, D. N.; STEVENS, K. N.; ELLIOT, L. L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. **The Journal of the Acoustical Society of America**, v.61, n.5, p.1337-51, 1977.

KILLION, M. C.; NIQUETTE, P. A.; GUDMUNDSEN, G. I.; REVIT, L. J.; BANERJEE, S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. **The Journal of the Acoustical Society of America**, v.116, p.2395-2405, 2006.

KOLLMEIER, B.; GILKEY, R. H.; SIEBEN, U. K. Adaptive staircase techniques in psychoacoustics: A comparison of human data and a mathematical model. **The Journal of the Acoustical Society of America**, v.83, p.1852-1862, 1988.

KOLLMEIER, B.; WESSELKAMP, M. Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. **The Journal of the Acoustical Society of America**, v.102, p.2412-2421, 1997.

KWON, B. J.; PERRY, T. T.; WILHELM, C. L.; HEALY, E. W. Sentence recognition in noise promoting or suppressing masking release by normal-hearing and cochlear-implant listeners. **The Journal of the Acoustical Society of America**, v.131, n.4, p.3111-9, 2012.

LORENZI, C.; HUSSON, M.; ARDOINT, M.; DEBRUILLE, X. Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception. **International Journal of Audiology**, v.45, p.487-495, 2006.

LEVITT, H. Transformed up-down methods in psychoacoustics. **The Journal of the Acoustical Society of America**, v.49, p.467-477. 1971.

LEVITT, H. Adaptive testing in audiology. **Scandinavian Audiology**, Suppl 6, 241-291. 1978.

LEVITT, H.; RABINER, L.R. Use of a sequential strategy in intelligibility testing. **The Journal of the Acoustical Society of America**, v.42, p.609-612, 1967.

LORENZI, C.; HUSSON, M.; ARDOINT, M.; DEBRUILLE, X. Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception. **International Journal of Audiology**, v.45, n.9, p.487-95, 2006.

LOLOV, S. R., RAYNOV, A. M.; BOTEVA, I. B.; EDREV, G. E. The Bulgarian hearing in noise test. **International Journal of Audiology**, v.47, n.6, p.371-2, 2008.

LUTS, H; BOON, E; WABLE, J; WOUNTERS, J. FIST: a French sentence test for speech intelligibility in noise. **International Journal of Audiology**, v.47, n.6, p.373-4, 2008.

MANTELATTO, S. A. C. Percepção da Inteligibilidade de fala por Sujeitos Jovens com audição normal frente à ruídos competitivos [Dissertação de Mestrado]. Ribeirão Preto: FFCLRP/USP, 1998.

MANTELATTO, S. A. C.; DA SILVA, J. A. Inteligibilidade de fala e ruído: Um estudo com sentenças do dia-a-dia. **Pró-Fono Revista de Atualização Científica**, v.12, p.48-55, 2000a.

MANTELATTO, S. A. C.; DA SILVA, J. A. Efeitos de diferentes tipos de ruído na Inteligibilidade de Fala. **Arquivos Brasileiros de Psicologia**, v.52, p.35-48, 2000b.

MILLER, G. A.; LICKLIDER, J. C. R. The intelligibility of interrupted speech. **The Journal of the Acoustical Society of America**, v.22, p.167-173, 1950.

MOON, S. K.; KIM, S. H.; MUN, H. A.; JUNG, H. K.; LEE, J. H.; CHOUNG, Y. H.; PARK, K. The Korean hearing in noise test. **International Journal of Audiology**, v.47, n.6, p.375-6, 2008

MYHRUM, M.; MOEN, I. The Norwegian hearing in noise test. **International Journal of Audiology**, v.47, n.6, p.377-8, 2008.

NAKAMURA, K.; GORDON-SALANT, S. Speech Perception in Quiet and Noise Using the Hearing in Noise Test and the Japanese Hearing in Noise Test by Japanese Listeners. **Ear and Hearing**, v.32, n.1, p.121-131, 2011.

NELSON, P. B.; JIN, S. H.; CARNEY, A. E.; NELSON, D. A. Understanding speech in modulated interference: cochlear implant users and normal hearing listeners. **The Journal of the Acoustical Society of America**, v.113, n.2, p.961-8, 2003

NILSSON, M. J.; SOLI, S. D.; SULLIVAN, J. Development of the hearing in noise test for the measurement of speech reception threshold in quiet and in noise. **The Journal of the Acoustical Society of America**, v.95, n.2, p.1085-99, 1994.

OXENHAM, A. J.; SIMONSON, A. M. Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. **The Journal of the Acoustical Society of America**, v.125, p.457-468, 2009.

OTERO, C. B.; BRIK, G.; FLORES, L.; ORTIZ, S.; ABDALA, C. The Latin American Spanish Hearing in Noise Test. **International Journal of Audiology**, v.47, n.6, p.362-3, 2008.

PEN, M. G.; MANGABEIRA-ALBERNAZ, P. L. Desenvolvimento de testes para logaudiometria: discriminação vocal. **Anales II Congresso Pan Americano de Otorrinolaringologia Y Broncoesofasologia**, v.2, p.223-6, Lima-Peru, 1970.

PEREIRA, L. D. Audiometria verbal: teste de discriminação vocal com ruído. [Tese de Doutorado]. São Paulo: Escola Paulista de Medicina, 1993.

PLOMP, R.; MIMPEN, A. M. Speech-reception threshold for sentences as a function of age and noise level. **The Journal of the Acoustical Society of America**, v.66, n.5, p.1333-42, 1979.

PETERS, R. W.; MOORE, B. C., BAER, T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. **The Journal of the Acoustical Society of America**, v.103, p.577-587, 1998.

QUAR, T. K.; MUKARI, S. Z.; ABDUL WAHAB, N. A.; ABDUL RAZAK, R. A.; OMAR, M.; MAAMOR, N. The malay hearing in noise test. **International Journal of Audiology**, v.47, n.6, p.379-80, 2008.

ROLL, E.; WALLENHAUPT, D.; RAMOS, A. P. F.; MENEGOTTO, I. H. Novas listas de monossílabos para avaliação do reconhecimento de fala. **Pró-Fono Revista de Atualização Científica**, v.15, n.2, p.159-68, 2003.

RHEBERGEN, K. S.; VERSFELD, N. J.; DRESCHLER, W. A. Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise. **The Journal of the Acoustical Society of America**, v.120, n.6, p.3988- 97, 2006.

SANTOS, S. N.; DANIEL, R. C.; COSTA, N. J. Estudo da equivalência entre as listas de sentenças em português. **Revista CEFAC**, v.11, n.4, p.673-680, 2009.

SANTOS, T. M. M.; RUSSO, I. C. P. A prática Audiologia clínica. 3 Ed. São Paulo: Cortez; 1986.

SCHOCHAT E. Percepção de fala: Presbiacusia e perda auditiva induzida pelo ruído. [Tese de Doutorado]. São Paulo: FFLCH/USP, 1994.

SHEJF, S.; YOST, W. A. Discrimination of start phase with sinusoidal envelope modulation. **The Journal of the Acoustical Society of America**, v.121, n.2, p.84-89, 2007.

SHIROMA, M.; IWAKI, T.; KUBO, T.; SOLI, S. The Japanese hearing in noise test. **International Journal of Audiology**, v.47, p.381-2, 2008.

SMITS, C.; FESTEN, J. M. The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: steady state noise. **The Journal of the Acoustical Society of America**, v.130, p.2987-2998, 2011.

SMITS, C.; FESTEN, J. M. The interpretation of speech reception threshold data in normal-hearing and hearing-impaired listeners: II. Fluctuating noise. **The Journal of the Acoustical Society of America**, v.133, p.3004-3015, 2013.

SMOORENBURG, G. F. Speech reception in quiet and in noisy conditions by individuals with noise - induced hearing loss in relation to their tone audiogram. **The Journal of the Acoustical Society of America**, v.91, n.1, p.421-37, 1992.

SOLI, S. D.; WONG, L. L. Assessment of speech intelligibility in noise with the Hearing in Noise Test. **International Journal of Audiology**, v.47, p.356-361, 2008.

STILLMAN, J. A. A comparison of three adaptive psychophysical procedures using inexperienced listeners. **Percepttion and Psychophysics**, v.46, p.345-50, 1989.

STUART, A.; PHILLIPS, D. P. Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. **Ear and Hearing**, v.17, p.478-489, 1996.

SUMMERS, V.; MOLIS, M. R. Speech recognition in fluctuating and continuous maskers: effects of hearing loss and presentation level. **Journal of Speech Language and Hearing Research**, v.47, n.2, p.245-56, 2004.

TAYLOR, M. M.; FORBES, S. M.; CREELMAN, C. D. PEST reduces bias in forced choice psychophysics. **The Journal of the Acoustical Society of America**, v.74, p.1367-1374, 1983.

TERBAND, H.; DRULLMAN R. Study of an automated procedure for a Dutch sentence test for the measurement of the speech reception threshold in noise. **The Journal of the Acoustical Society of America**, v.124, p.3225-3234, 2008.

VAILLANCOURT, V.; LAROCHE, C.; MAYER, C.; NALI, M.; ERIKS-BROPHY, A.; GIGUERE, C. The Canadian French hearing in noise test. **International Journal of Audiology**, v.47, p.383-5, 2008.

VERMIGLIO, A. J. The American English hearing in noise test. **International Journal of Audiology**, v.47, n.386-7, 2008.

VERSFELD, N. J.; DAALDER, L.; FESTEN, J. M.; HOUTGAST, T. Method for the selection of sentence materials for efficient measurements of the speech reception threshold. **The Journal of the Acoustical Society of America**, v.107, p.1671-1684, 2000.

WETHERILL, G. B. Sequential estimation of quantal responses. **Journal of the Royal Statistical Society**, v.25, p.1-48. 1963.

WETHERILL, G. B.; LEVITT, H. Sequential estimation of points on a psychometric function. **British Journal of Mathematical and Statistical Psychology**, v.18, p.1-10, 1965.

WILSON, R. H.; CARNELL, C. S.; CLEGHORN, A. L. The words-in-noise (WIN) test with multitalker babble and speech-spectrum noise maskers. **Journal of the American Academy of Audiology**, v.18, p.522-529, 2007.

WILSON, R. H.; MCARDLE, R.; ROBERTS, H. A comparison of recognition performances in speech-spectrum noise by listeners with normal hearing on PB-50, CID W-22, NU-6, W-1 spondaic words, and monosyllabic digits spoken by the same speaker. **Journal of the American Academy of Audiology**, v.19, p.496-506, 2008

WONG, L. L. N. The Cantonese Hearing in Noise Test. **International Journal of Audiology**, v.47, n.6, p.388-390, 2008.

WONG, L. L. N.; HUANG, V. The Taiwanese Mandarin Hearing in Noise Test. **International Journal of Audiology**, v.47, n.6, p.391-392, 2008.

WONG, L. L. N.; LIU, S.; HAN, N. The Mainland Mandarin Hearing in Noise Test. **International Journal of Audiology**, v.47, n.6, p.393-5, 2008.

WONG, L. L. N.; SOLI, S. D.; LIU, S.; HAN, N.; HUANG, M. W. Development of the Mandarin Hearing In Noise Test (MHINT). **Ear and Hearing**, v.28, n.2, p.70S-74S, 2007.



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS DA SAÚDE
DEPARTAMENTO DE FONOAUDIOLOGIA
CURSO DE FONOAUDIOLOGIA



APÊNDICE A - TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Titulo do Projeto: Teste de reconhecimento de fala HINT – Brasil em adultos do nordeste brasileiro com audição normal

Pesquisador Responsável: Fernando Augusto Pacífico, Fone para contato: 21268927 / 8527-5246 - Av. Artur de Sá, sn / Cidade Universitária, Recife-PE, CEP: 50670-901.

Instituição Responsável: Universidade Federal de Pernambuco

O Sr.(a) está sendo convidado(a) para participar de uma pesquisa que tem por objetivo - Investigar teste de reconhecimento de fala HINT - Brasil em adultos do nordeste brasileiro com audição normal. Se concordar em fazer parte deste estudo, serão realizados exames auditivos, no qual você irá escutar alguns sons provindos de um fone de ouvido.

Caso você tenha alguma dúvida durante a explicação e/ou leitura desse termo de consentimento, pergunte ao entrevistador, para que você seja bem esclarecido (a) sobre tudo que está respondendo. Você irá receber uma cópia deste termo de consentimento, que está em duas vias, uma ficará com o pesquisador e outra com você.

A realização deste estudo oferece o risco de haver algum cansaço e/ou desconforto físico causado pela apresentação dos sons aos ouvidos e/ou constrangimento por solicitar a repetição de sentenças. Asseguro a você, voluntário, que sua autorização na pesquisa pode ser retirada em qualquer momento da mesma sem perda de benefícios ou qualquer outra penalidade. As informações conseguidas por meio do estudo possibilitarão um maior conhecimento sobre a efetividade dos procedimentos de diagnóstico audiológico. Como benefício direto, os participantes receberão orientações por escrito sobre os procedimentos de diagnóstico audiológico, seus resultados de sua avaliação auditiva, bem como irão verificar suas performances de reconhecimento da fala no ruído.

As informações obtidas a partir deste estudo serão rigorosamente confidenciais. Os resultados serão divulgados publicamente em congressos e publicações científicas, entretanto, sua identidade e de quem mais esteja envolvido serão mantidas sob sigilo. Os dados ficarão sob a responsabilidade do pesquisador principal pelo período mínimo de 05 anos em computador pessoal.

A sua participação neste estudo é totalmente voluntária, não sendo proporcionado dinheiro ou outras formas de bens materiais. Em caso de dúvidas adicionais sobre a sua participação retire-as com o pesquisador. Não assine o termo se não concordar em participar, ou se as dúvidas não forem esclarecidas satisfatoriamente. Caso tenha dúvida quanto aos aspectos éticos da pesquisa você poderá contatar o Comitê de Ética em Pesquisa com Seres Humanos: Av. da engenharia, s/n - 1º andar, Cidade Universitária CEP: 50670-901 Recife - PE, Brasil. Telefone/Fax do CEP: (81) 2126-8588.

Pesquisadora Responsável: Fernando Augusto Pacífico

Li e entendi as informações precedentes, descritas neste estudo e, todas as minhas dúvidas foram respondidas satisfatoriamente. Dou livre o meu esclarecimento em participar deste estudo e não abro mão de nenhum direito legal que eu tenha.

Recife, ____/____/____.

Assinatura do voluntário

Testemunha

Testemunha

APÊNDICE B - ENTREVISTA AUDIOLÓGICA

I. DADOS DE IDENTIFICAÇÃO

Nome: _____

Idade: _____ Data de Nascimento: _____

Naturalidade: _____ Estado Civil: _____

Endereço: _____

Telefone: _____ Profissão: _____

Ocupação Atual: _____ Carga horária / dia: _____

Outras Atividades: _____ Tempo na Atividade: _____

Data da Entrevista: _____ Entrevistador: _____

II. HISTÓRIA CLÍNICA

Queixa auditiva: () Sim () Não Qual(is)/Quando? _____

Antecedentes familiares: () Sim () Não Quem? _____

Dificuldade no entendimento / compreensão da fala: () Sim () Não

Qual(is) situação(ões)? _____

Hábitos de natureza externa: () sim () não Qual(is)/quando? _____

Queixa otológica: () Sim () Não Qual(is)/Quando? _____

Queixa neurológica: () Sim () Não Qual(is)/Quando? _____

Queixa psiquiátrica: () Sim () Não Qual(is)/Quando? _____

Cirurgias / Estados mórbidos: () Sim () Não Qual(is)/Quando? _____

Outras doenças: () Sim () Não Qual(is)/Quando? _____

Tratamento: () Sim () Não Qual(is)/Quando? _____

Dados Adicionais:

APÊNDICE C - ARTIGO ORIGINAL 1: EFFECT OF MODULATION RATE ON MASKING RELEASE FOR SPEECH

Artigo Original

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; *SMG* 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, *Institutional 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.

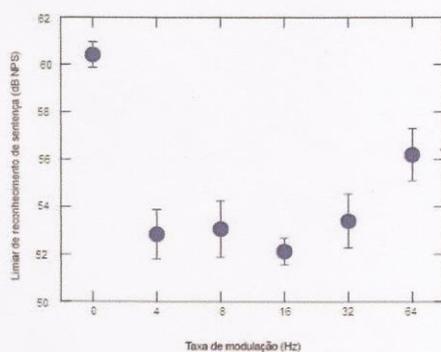


Figura 1. Média dos limiares de reconhecimento de sentença em função da frequência de modulação do ruído

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
	64	4,22*	0,548	0,000
4	0	-7,58*	0,398	0,000
	8	-0,22	0,708	0,764
	16	0,72	0,529	0,221
	32	-0,56	0,689	0,440
	64	-3,36*	0,676	0,001
8	0	-7,36*	0,587	0,000
	4	0,22	0,708	0,764
	16	0,94	0,649	0,221
	32	-0,34	0,738	0,657
	64	-3,14*	0,726	0,003
16	0	-8,30*	0,350	0,000
	4	-0,72	0,529	0,221
	8	-0,94	0,649	0,221
	32	-1,28	0,569	0,066
	64	-4,08*	0,554	0,000
32	0	-7,02*	0,563	0,000
	4	0,56	0,689	0,440
	8	0,34	0,738	0,657
	16	1,28	0,569	0,066
	64	-2,80*	0,742	0,020
64	0	-4,22*	0,548	0,000
	4	3,36*	0,676	0,001
	8	3,14*	0,726	0,003
	16	4,08*	0,554	0,000
	32	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, Delhome 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, Debrulle 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. Shejif 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, Banbara 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 Audio Electroacoustic*. 1973;21:196-201.
30. Miller GA, Licklider JCR. The intelligibility of interrupted speech. *J Acoust Soc Am*. 1950;22(2):167-73.

ANEXO A - APROVAÇÃO DO COMITÊ DE ÉTICA AMERICANO



THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

OFFICE OF HUMAN RESEARCH ETHICS
Medical School Building 52
Mason Farm Road
CB #7097
Chapel Hill, NC 27599-7097
(919) 966-3113
Web site: ohre.unc.edu
<https://my.research.unc.edu> for IRB status
Federalwide Assurance (FWA) #4801

To: John Grose
Otolaryngology (Ent)
CB: 7070

From: Biomedical IRB

Approval Date: 6/13/2011
Expiration Date of Approval: 6/11/2012

RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Submission Type: Initial
Expedited Category: 4. Noninvasive clinical data
Study #: 11-1113

Study Title: Temporal Masking and Speech Recognition in the Aging Auditory System: US-Brazil
Sponsors: National Institutes of Health (NIH)

This submission has been approved by the above IRB for the period indicated. It has been determined that the risk involved in this research is no more than minimal.

Study Description:

Purpose: This study seeks to investigate age-related changes in speech perception abilities in complex backgrounds, with a particular interest in temporal processing abilities. **Participants:** Normal hearing listeners, ages 18 - 85 years of age, will undergo testing in a sound proof booth using headphones. **Procedures (methods):** There are three tasks associated with this research study: 1) press a button when a target stimulus is heard; 2) repeat sentences aloud as they are heard; 3) listen passively while auditory evoked potentials are recorded from the scalp.

Regulatory and other findings:

Social security numbers (SSN) may be collected for this study for tax identification purposes. The subject is required to disclose his/her SSN in order to receive the incentive payment(s).

Investigator's Responsibilities:

Federal regulations require that all research be reviewed at least annually. It is the Principal Investigator's responsibility to submit for renewal and obtain approval before the expiration date. You may not continue any research activity beyond the expiration date without IRB approval. Failure to receive approval for continuation before the expiration date will result in automatic termination of the approval for this study on the expiration date.

IF YOU SUBMITTED ON PAPER, enclosed are stamped copies of approved consent documents and other recruitment materials (when applicable). You must copy the stamped consent forms for

ANEXO B - APROVAÇÃO DO COMITÊ DE ÉTICA NACIONAL 1



SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS DA SAÚDE

Comitê de Ética em Pesquisa
Av. da Engenharia, s/n – 1º andar – Sala 4 – Cidade Universitária
50.670-901 Recife – PE, Tel/fax: 81. 2126.8588 – cepecs@ufpe.br

Ofício nº. 277/2012 - CEP/CCS/UFPE

Recife, 24 de maio de 2012

À

Dra. Denise Costa Menezes

Departamento de Fonoaudiologia/CCS/UFPE

Registro CONEP: 16.750

Registro do SISNEP FR – 457996

CAAE – 0401.0.172.000–11

Registro CEP/CCS/UFPE Nº 415/11

Título: “Mascaramento temporal e reconhecimento de fala no sistema auditivo senescente”.

Pesquisador Responsável: Denise Costa Menezes

Senhor (a) Pesquisador (a):

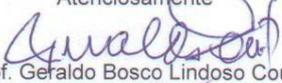
Informamos que a Comissão Nacional de Ética em Pesquisa – CONEP, de acordo com as atribuições definidas na Resolução CNS 196/96, analisou o protocolo em epígrafe e emitiu o seguinte parecer: **PROCOLO APROVADO - PARECER Nº 233/2012**. Dessa forma, este Comitê **libera** o protocolo para início da coleta de dados em **16 de maio de 2012**.

Ressaltamos que a aprovação definitiva do projeto será dada após a entrega do **relatório final**, conforme as seguintes orientações:

- a) Projetos com, no máximo, 06 (seis) meses para conclusão: o pesquisador deverá enviar apenas **um relatório final**;
- b) Projetos com períodos maiores de 06 (seis) meses: o pesquisador deverá enviar **relatórios semestrais**.

Dessa forma, o ofício de aprovação somente será entregue após a análise do relatório final.

Atenciosamente


Prof. Gefaldo Bosco Lindoso Couto
Coordenador do CEP/CCS / UFPE

ANEXO C – APROVAÇÃO DO COMITÊ DE ÉTICA NACIONAL 2

Dados do Projeto de Pesquisa

Título da Pesquisa: TESTE DE RECONHECIMENTO DE FALA HINT - BRASIL EM ADULTOS DO NORDESTE BRASILEIRO COM A AUDIÇÃO NORMAL

Pesquisador: FERNANDO AUGUSTO PACÍFICO

Área Temática:

Versão: 2

CAAE: 02466612.2.0000.5208

Submetido em: 06/10/2012

Instituição Proponente: Universidade Federal de Pernambuco - UFPE

Situação: Aprovado

Localização atual do Projeto: Pesquisador Responsável

Patrocinador Principal: Financiamento Próprio

ANEXO D – NORMAS DO PERÍODICO INTERNATIONAL JOURNAL OF AUDIOLOGY

International Journal of Audiology Instructions for Authors

The *International Journal of Audiology* invites (in English only) as follows: **Original Articles** embracing all aspects of hearing, hearing loss, and related disorders. In addition, **Clinical Notes** and **Technical Reports** presenting preliminary results from pilot studies, descriptions of new technology, or case studies are welcomed. **Discussion Papers** with editorial comments following the paper as well as **Letters to the Editor** are also encouraged. It is assumed that the submitted material will provide significant contributions/advances in scientific and clinical knowledge.



Submitted manuscripts are subject to editorial review and are received with the explicit understanding that they are not under simultaneous consideration by any other publication. Submission of a manuscript to the *International Journal of Audiology* is taken as evidence that no portion of the text or figures has been copyrighted, published, or submitted for publication elsewhere unless information regarding previous publication is explicitly cited and permission obtained. A copy of such permission must accompany the submitted manuscript.

Reports concerning studies involving human or animal test subjects should include a statement regarding approval of an ethical committee. Authors should identify any concerns about possible conflicts of interest concerning the study being reported.

Authors are expected to have consulted statistical expertise as required in order to apply suitable statistical methods in assessing results of studies reported.

Manuscript Preparation

MS Word or Word Perfect should be used for the text. MS Word or MS PowerPoint, EPS, or JPG should be used for figures, and MS Excel for tables. PDF files will not be accepted for the main document. The review process will make use of electronic transmission of manuscripts to reviewers.

Manuscripts should be typed double-spaced with 2.5-cm (1-inch) margins on only one side of each A4 (or 8½ × 11 inch) page. Headings, including up to three levels of subheadings, should be used to designate the major sections of the article. Pages should be numbered consecutively in the upper right-hand corner. The surname of the first author should appear on the upper left-hand corner, followed by a brief running title. *Authors are encouraged to propose up to five names and email addresses of possible referees for their paper.*

Either British or American English may be used, but the version chosen must be consistent throughout the manuscript.

Title Page. The first page of each manuscript should include the following: title of the article; names of all authors in full without academic degrees; institutional affiliations of each author (superscript letters should be used to link authors to affiliations); a short list of key words reflecting the content; a list of acronyms and abbreviations with their meanings spelled out; and the full postal address as well as the E-mail address of the corresponding author.

Abstract. An abstract of **NO MORE THAN** 200 words should precede each article, including review articles, using the following headings: **Objective, Design, Study Sample, Results, Conclusions.**

Text. The text of the manuscript should be double spaced and left justified, with the first line of each paragraph flush with the margin. Paragraphs should be separated by one double line space.

Acknowledgements and Declaration of Interest. If the article was presented at a professional meeting, include the full title, date, and location of the meeting. Acknowledgments, grant numbers, and supporting agencies should also be listed. This information should be included in a separate section at the end of the text, before the references. Please see below for more information.

Tables. Tables should be numbered consecutively using arabic numerals and typed on separate pages. Each table should have a brief but sufficiently detailed explanatory title. All tables must be cited in the text in numerical order.

Figures. All figures should be numbered consecutively in the order in which they appear in the manuscript. Legends should be prepared on a separate page following the main text and tables, and explain each figure in detail. If color figures are submitted they will appear in the online version at no charge to the authors. Color charges for the print version will be borne by the authors. Original artwork should be in the form of high-quality electronic format. Any identifiable photographs of patients must be accompanied by a release form signed by the patient. All figures must be cited in the text in numerical order.

Figure files should be kept as separate files, in TIF, EPS, PDF or JPG format. Providing these formats will guarantee that the quality of the graphics is good throughout the publishing process, if provided with sufficient resolution. Photographic illustrations should be rendered with at least 300 DPI; please use CMYK colour conversion if possible. Graphs made with Office software such as Microsoft Excel, can be provided in their original format to facilitate conversion into printable format with preserved quality. Any other line graphs/illustrations should preferably be provided in EPS format with a resolution of at least 600 DPI to prevent ragged lines when printed.

References. The maximum number of references is 40. The reference system shall essentially follow the Harvard style but with some exceptions, the main one being the use of abbreviated journal names according to MEDLINE (see <http://www.nlm.nih.gov/tsd/serials/lji.html>). Authors are encouraged to look at a current version of the journal for examples.

References must be cited in the text using name(s) of author(s) and year of publication. Examples: Anderson (1995); (Anderson & Jones, 1998). When a work has three or more authors, cite only the first author followed by et al and the year. The list of references should be in alphabetical order and written double-spaced on a separate page at the end of the main text.

References to journals should include (in this order): all authors' names with initials for up to five authors or, for six or more, the first five authors followed by: et al; year of publication; article title; journal name in italics, abbreviated according to Index Medicus; volume number; and inclusive page numbers. An example is given below:

Litovsky, R.Y., Johnstone, P.M., Godar, S.P. 2006. Benefits of bilateral cochlear implants and/or hearing aids in children. *Int J Audiol*, 45, 3-11.

References to books should include (in this order): authors' names as above; year of publication; chapter title; editors' names with initials (as applicable); book title and

edition; city; publisher; year of publication; and inclusive page numbers. Examples are given below.

Valente, M. & Valente, L.M. (2008). Earhooks, tubing, earmolds, and shells. Valente, Hosford-Dunn, Roeser (eds.) *Audiology Treatment: Second Edition*. New York: Thieme, pp. 36 - 71.

Luxon, LM. & Prasher, D. (eds.) 2007. *Noise and it's effects*. Chichester, England: John Wiley & Sons.

The corresponding author is responsible for ensuring that the references are complete and correct. When a revised manuscript is returned, authors shall certify that all references cited in the text are included and quoted correctly in the list of references. Manuscripts submitted to another publication, but not yet accepted, should be cited in the text as unpublished data (in parentheses) and not included in the reference list. References to manuscripts accepted but not yet published are discouraged, but when they are required should be included in the reference list, followed by "in press".

References should be left justified, with one line space between consecutive references.

Abbreviations. Abbreviations, such as Hz, kHz, etc should be consistent throughout the text of the manuscript. Exceptions can be made for tables and figures if necessary.

Supplementary Materials. Supplementary materials submitted with the manuscript, such as appendices, copies of questionnaires and test materials, or sound files of stimuli used in the study can be posted on the IJA website at www.informahealthcare.com/ija if the paper is accepted for publication. Any supplemental materials should be submitted as separate files. Please make sure to insert proper references to any material posted online, as this will be the responsibility of the author. Please include the following information where appropriate "supplementary figure/table/material available in the online version of the journal. Please find this material with the direct link to the article at: [http://www.informaworld.com/\(DOI number\)](http://www.informaworld.com/(DOI number)).

Manuscript Types: Descriptions and Typical Sizes. While the size of submissions will vary, the typical paper is 8-10 printed pages. Based on this, the following are guidelines for submissions:

- a. **Original papers**—present novel and unique data, new findings or creative concepts on important audiology issues (4500-5000 words with 3-5 figures/tables).
- b. **Discussion papers**—present topics that are controversial or challenging. Reviewer comments follow the published paper and author responses to the review comments follow the reviewer comments (6000-6500 words with 3-5 figures/tables).
- c. **Clinical Notes**—present unique or exceptional clinical cases with important/significant findings in audiology (2500-3000 words with 2-4 figures/tables).
- d. **Technical Reports**—present clinically relevant technical information about new hardware or processing strategies (2500-3000 words with 2-4 figures).
- d. **Letters to the editor**—responses to previously published papers, or comments on important topics/issues (1200 words).

It is important that authors make all efforts to keep their submissions succinct and prepare their work with the minimal size submission. Please note that authors will be charged an excessive page charge of \$200 USD for each printed page in excess of 8 - 9 pages. Authors who are unable to pay these charges, can contact the Editor-in-Chief to request an exemption.

Page proofs: 4 - 8 weeks after a manuscript has been accepted, page proofs are sent by the publisher to the corresponding author and should be returned with the least possible delay (usually within 48 hours of receipt).

Authors' responsibility. The *International Journal of Audiology* is not responsible for the statements made or the views put forward in the various papers, nor does it endorse any of the products or materials herein advertised.

Acknowledgments and Declaration of Interest sections

Acknowledgments and Declaration of interest sections are different, and each has a specific purpose. The Acknowledgments section details special thanks, personal assistance, and dedications. Contributions from individuals who do not qualify for authorship should also be acknowledged here. Declarations of interest, however, refer to statements of financial support and/or statements of potential conflict of interest. Within this section also belongs disclosure of scientific writing assistance (use of an agency or agency/ freelance writer), grant support and numbers, and statements of employment, if applicable.

Acknowledgments section

Any acknowledgments authors wish to make should be included in a separate headed section at the end of the manuscript preceding any appendices, and before the references section. Please do not incorporate acknowledgments into notes or biographical notes.

Declaration of Interest section

All declarations of interest must be outlined under the subheading "Declaration of interest". If authors have no declarations of interest to report, this must be explicitly stated. The suggested, but not mandatory, wording in such an instance is: *The authors report no declarations of interest.* When submitting a paper via ScholarOne Manuscripts, the "Declaration of interest" field is compulsory (authors must either state the disclosures or report that there are none). If this section is left empty authors will not be able to progress with the submission.

Please note: for NIH/Wellcome-funded papers, the grant number(s) must be included in the Declaration of Interest statement.

Click here to view our full [Declaration of Interest Policy](#).

Submitting a Manuscript

The International Journal of Audiology reserves the right to retain any manuscripts submitted and cannot assume responsibility for loss of manuscripts. Therefore, authors should retain a complete copy of the manuscript, including figures and tables.

All manuscripts should be submitted electronically through <http://mc.manuscriptcentral.com/tija>. The review process will make use of electronic transmission of manuscripts to reviewers.

Please contact the Editor-in-Chief for all author queries:

Ross J. Roeser, Editor-in-Chief

UTD/Callier Center for Communication Disorders 1966 Inwood Road,

Dallas, TX 75235 USA Tel: + 1 214 905 3001 Fax: +1 214 905 3022

E-mail: editor-ija@utdallas.edu

ANEXO E - CERTIFICADO DE APRESENTAÇÃO DE TRABALHO NO 21º CONGRESSO BRASILEIRO E 2º IBERO-AMERICANO DE FONOAUDIOLOGIA



CERTIFICADO

Certificamos que o trabalho **EFFECT OF MODULATION RATE ON MASKING RELEASE FOR SPEECH** de autoria de **Advíncula, K.P.; Menezes, D.C.; Pacífico, F.A.; Griz, S.M.S.**, foi apresentado na forma de Apresentação Oral no **21º Congresso Brasileiro e 2º Ibero-Americano de Fonoaudiologia**, realizado de 22 de setembro a 25 de setembro de 2013, no ENOTEL Resort - Porto de Galinhas-PE.

Porto de Galinhas, 25 de Setembro de 2013.


DR. MARIA MARCONDES
PRESIDENTE DO SBFA


DR. MARILENE CAROLINA TOSTA
DIRETORA GERAL


DR. HELTON JUNIOR
DIRETOR GERAL

REALIZAÇÃO


PATROCÍNIO ORO


PATROCÍNIO PRATA


APOIO