

Universidade de Brasília
Instituto de Psicologia
Departamento de Processos Psicológicos Básicos
Pós-Graduação em Ciências do Comportamento



DESENVOLVIMENTO DE UM CONJUNTO DE ESTIMULOS PARA TREINAMENTO
AUDITIVO DE IDOSOS EM RECONHECIMENTO DE FALA EM CONDIÇÃO DE
RUÍDO

Luciana Carelli Henriques de Andrade

Orientadora: Prof^a Dr^a Maria Ângela Guimarães Feitosa

Dissertação apresentada ao Programa de Pós-Graduação em Ciências do Comportamento, do Departamento de Processos Psicológicos Básicos, Instituto de Psicologia, Universidade de Brasília, como parte dos requisitos para obtenção do grau de Mestre em Ciências do Comportamento área de concentração: Cognição e Neurociência do Comportamento.

Brasília,

2010

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE PSICOLOGIA
DEPARTAMENTO DE PROCESSOS PSICOLÓGICOS BÁSICOS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DO COMPORTAMENTO

A Banca Examinadora foi composta por:

Prof^a Dr^a Maria Ângela Guimarães Feitosa (Presidente)
Universidade de Brasília

Prof. Dr. Francisco Dyonísio Cardoso Mendes (membro efetivo)
Universidade de Brasília

Prof^a Dr^a Helena Bolli Mota (membro efetivo)
Universidade Federal de Santa Maria

Prof^a Dr^a Wânia Cristina de Souza (membro suplente)
Universidade de Brasília

Agradecimentos

A Deus pela força e o amparo durante os anos em que estive em meio aos artigos e livros.

Ao professor Cláudio Ladys e ao meu amigo Antônio José R. Duque pelo incentivo que me deram para fazer a seleção de mestrado. Carinhosa gratidão aos amigos Gléna Salgado, Thiago Bastos, Neinielly Cristina, João R. Simczak pela amizade e companheirismo nos meus momentos de dificuldade.

Aos amigos Hidemberg Alves da Frota, Geovani Sorrentino, Marta R. D. Prestes, Magda F. G. e Andrade e Roberta L. Leonardo que me auxiliaram em várias atividades acadêmicas durante o mestrado.

À professora Wânia Cristina de Souza pela amizade e apoio durante o estágio em docência, aos funcionários do Departamento de Processos Psicológicos Básicos Joyce N. do Rêgo, Keules R. do Nascimento e Kleber A. Freitas. Agradeço ao CNPq por concessão temporária de bolsa de estudo.

À Faculdade de Comunicação da Universidade de Brasília, representada pelo professor Carlos Eduardo Esch e aos técnicos do Laboratório de Rádio José Carlos Fontes Jr., em especial, a Josevaldo Souza por gravar e editar os estímulos gravados. Ao prof. Gérson A. Janczura que permitiu o uso do Laboratório de Processos Cognitivos durante um período importante da pesquisa. Agradeço também aos voluntários que doaram sua voz para criação do banco de palavras.

À minha orientadora Maria Ângela G. Feitosa que foi mais que uma orientadora durante todo mestrado, e a minha família pelo apoio e companheirismo.

“Creio que quase sempre é preciso um golpe de loucura para se construir um destino”.

Marguerite Yourcenar

Índice de Figuras

Figura 1 – Espectrograma da palavra barro	17
Figura 2 – Espectrograma mostrando a distribuição de energia no intervalo de 0 a 20.000Hz	17
Figura 3 – Espectrogramas dos pares mínimos: Barro-Jarro e Mola-Bola	43
Figura 4 – Espectrogramas dos pares mínimos: Fila-Vila e Pó-Vó	44
Figura 5 – Espectrogramas dos pares mínimos: Pão-Chão e Xiz-Giz	45
Figura 6 – Espectrogramas dos pares mínimos: Couro-Touro e Data-Nata	46
Figura 7 – Espectrogramas dos pares mínimos: Tia-Dia e Mapa-Tapa	47
Figura 8 – Espectrogramas dos pares mínimos: Ventar-Sentar e Vinho-Ninho	48
Figura 9 – Espectrogramas dos pares mínimos: Mão-Não e Fanta-Janta	49
Figura 10 – Espectrogramas dos pares mínimos: Bote-Pote e Cinco-Zinco.....	50
Figura 11 – Espectrogramas dos pares mínimos: Tela-Zela e Conta-Ponta	51

Índice de Tabelas

Tabela 1 – Fonemas que fazem o contraste entre diferentes pares de palavras, segundo o número de ocorrências, sua posição inicial (PI) ou medial (PM) na 1ª e 2ª palavra do par e sua faixa de frequência característica.	20
Tabela 2a – Distribuição de ocorrências das combinações de faixas de frequências para o fonema discriminativo integrando a 1ª palavra do par.	22
Tabela 2b – Distribuição de ocorrências das combinações de faixas de frequência para o fonema discriminativo integrando a 2ª palavra do par.	23
Tabela 3 – Características acústicas dos pares mínimos Barro-Jarro e Mola-Bola	43
Tabela 4 – Características acústicas dos pares mínimos: Fila-Vila e Pó-Vó	44
Tabela 5 – Características acústicas dos pares mínimos: Pão-Chão e Xiz-Giz	45
Tabela 6 – Características acústicas dos pares mínimos: Couro-Touro e Data-Nata.....	46
Tabela 7 – Características acústicas dos pares mínimos: Tia-Dia e Mapa-Tapa	47
Tabela 8 – Características acústicas dos pares mínimos: Ventar-Sentar e Vinho-Ninho	48
Tabela 9 – Características acústicas dos pares mínimos: Mão-Não e Fanta-Janta	49
Tabela 10 – Características acústicas dos pares mínimos: Bote-Pote e Cinco-Zinco.....	50
Tabela 11 – Características acústicas dos pares mínimos: Tela-Zela e Conta-Ponta.....	51

Resumo

Níveis mais avançados de presbiacusia podem causar problemas de comunicação oral. Um importante meio de intervenção que minimiza este problema, e em especial a dificuldade de reconhecimento da fala, é o treinamento auditivo. Justificam seu uso o aumento do desempenho em tarefas de treinamento e a plasticidade cerebral. Poucas pesquisas, softwares e registros em áudio são encontrados em português brasileiro para fins de treinamento auditivo. Neste trabalho, foi desenvolvido um conjunto de estímulos adequados para treinamento auditivo do falante do português brasileiro, incluindo um banco de palavras escritas, na configuração de pares mínimos, e um banco de áudio contendo estes pares e o murmúrio multifalante. Foram organizadas três listas de pares contendo todos os fonemas do português brasileiro, balanceadas quanto às combinações por faixa de frequências. Os pares mínimos e o murmúrio foram gravados em ambiente acusticamente tratado, mostrando-se acusticamente adequados para uso em treinamento auditivo. Acredita-se que, acusticamente, os pares propostos propiciem graus variados de dificuldade de reconhecimento e que o murmúrio simule o ruído encontrado no dia a dia. Sugere-se que os estímulos sejam utilizados em tarefas de treinamento auditivo em idosos e adultos com perda neurossensorial.

Palavras chave: reconhecimento de fala, perda neurossensorial, treinamento auditivo, pares mínimos, murmúrio multifalante.

Abstract

Severe presbycusis provides oral communication problems. Auditory training proves to be an important means of intervention to minimize the difficulty of speech recognition. The increase in performance on tasks of training and brain plasticity justify its application. Few investigations, softwares and audio recordings of auditory training are found for Brazilian Portuguese. A set of stimuli for auditory training was developed for Brazilian Portuguese speakers, including a bank of words written in the configuration of minimal pairs. A database of the minimal pairs and the babble noise was recorded. The minimal pairs were balanced with regard to number of phoneme contrasts and frequency region. The minimal pairs and the babble noise were recorded in an acoustically treated room and they showed to be acoustically suitable for using in auditory training. It is believed that these stimuli provide different levels of difficulty in auditory recognition tasks. It is suggested these stimuli be used in auditory training tasks by the elderly and adults with sensorineural hearing loss.

Key words: speech recognition, sensorineural hearing loss, auditory training, minimal pairs, babble noise.

Lista de Anexos

Anexo A- Audiograma de fonemas de Russo e Behlau (1993)	34
Anexo B- Lista de pares mínimos para familiarização e treinamento auditivo.....	35
Anexo C- Amostra do texto original de Clarisse Lispector e do texto sem sentido correspondente.....	36

Sumário

Agradecimentos.....	iii
Índice de Figuras	iv
Índice de Tabelas.....	v
Resumo.....	vi
Abstract	vii
Lista de Anexos	viii
Introdução.....	1
<i>A Natureza da Dificuldade em Reconhecimento de Fala em Idosos</i>	2
<i>Qualidade do Som</i>	2
<i>Integridade do Ouvido e dos Sistemas Nervosos Auditivos Periféricos e Central</i>	4
<i>Habilidades Cognitivas</i>	5
<i>A Pertinência do Treinamento em Reconhecimento de Fala com Base em Evidência</i>	
<i>Biológica</i>	6
<i>A Pertinência do Treinamento com Base Comportamental e as Razões para Desenvolver</i>	
<i>Materiais para Treinamento</i>	7
Método	11
<i>Delineamento do Estudo</i>	11
<i>Participantes</i>	12
<i>Equipamentos e Ambiente de Gravação dos Estímulos</i>	12
<i>Procedimentos</i>	13
<i>Preparação dos estímulos e critérios de escolha</i>	13
<i>Gravação dos estímulos</i>	14
<i>Análise acústica</i>	15
Resultados	18
<i>Análise Fonética das Palavras</i>	18
<i>Análise Qualitativa e Critérios de Seleção Final dos Estímulos Gravados</i>	24
<i>Análise Acústica dos Pares Mínimos</i>	24
Discussão.....	30
Referências	34
Anexos.....	39

O envelhecimento constitui processo degenerativo progressivo, marcado por alterações estruturais e fisiológicas nos sistemas cardiovascular, respiratório, músculo-esquelético, nervoso e sensorial, as quais desencadeiam complicações importantes para os processos sensoriais e cognitivos (Feitosa, 2001). Com o avançar da idade, o sistema nervoso diminui sua capacidade de elaborar, integrar e reagir aos estímulos prejudicando as percepções sensoriais (Willet, 1991).

A presbiacusia, que é um tipo de perda neurossensorial bilateral, ocorre em razão das mudanças degenerativas e fisiológicas naturais do sistema auditivo associadas ao envelhecimento, gerando mudança gradual na sensibilidade auditiva para todas as frequências, sendo que as perdas para as frequências altas são as de maior magnitude (Bess, Hedley-Williams & Lichtenstein, 2001; Working Group on Speech Understanding and Aging *i. e.* CHABA, 1988; Corso, 1977; Jerger, Jerger, Oliver & Pirozzolo, 1989; Schucknecht, 1955). Isto ocorre porque as células cocleares localizadas no giro basal da cóclea, responsáveis pela transdução de ondas de alta frequência, se degeneram com o avançar da idade. A perda auditiva neurossensorial é o principal fator explicativo para a dificuldade de compreensão da fala (Humes, Watson, Christensen, Cokely, Halling & Lee, 1994; Schucknecht, 1955). Os idosos com presbiacusia tendem a ter baixo desempenho para tarefas de compreensão de fala especialmente em ambientes ruidosos ou com reverberação (Helfer & Huntley, 1991). O Ministério da Saúde por meio da Política Nacional de Saúde da Pessoa Portadora de Deficiência (2002) estima que a presbiacusia seja a principal causa de deficiência auditiva nos idosos por abranger cerca de 30% dessa população.

A Natureza da Dificuldade em Reconhecimento de Fala em Idosos

Existe uma grande dificuldade metodológica para se pesquisar as contribuições específicas dos sistemas sensoriais periféricos e centrais, e dos processos cognitivos para presbiacusia, porque esses estão intimamente relacionados (CHABA, 1988; Marshall, 1981). Marshall (1981) destaca que para interpretar as diferenças de desempenho entre ouvintes jovens e idosos deve-se notar que, mesmo quando jovens e idosos têm o mesmo grau de perda auditiva, os ouvintes idosos vão apresentar problemas adicionais em virtude de disfunções no Sistema Nervoso Auditivo Central (SNAC), diferenças cognitivas e problemas auditivos periféricos associados à presbiacusia, que não são completamente explicados pelo aumento de limiar ou a combinação de todos esses fatores. Três classes de fatores principais contribuem para o reconhecimento da fala no idoso: qualidade do som, integridade estrutural do ouvido e do sistema nervoso auditivo central e periférico, e habilidades cognitivas (CHABA, 1988; Marshall, 1981).

Qualidade do Som

Vários aspectos podem comprometer a acuidade auditiva para os sons da fala: sintetização ou digitalizações da fala que comprometam a qualidade do som; a presença de distorções e atenuações durante a fala; ambientes sonoros complexos (que incluem ruído e reverberação) e as alterações temporais da fala como, por exemplo, falas rápidas (CHABA, 1988; Marshall, 1981; Gordon-Salant, 2005).

Sintetizações ou Digitalizações da Fala. Empresas de telecomunicação, bancos, hospitais, entre outros, têm adotado programas eletrônicos que utilizam sentenças gravadas para fazerem os atendimentos aos clientes. Esse tipo de som tornou-se uma interface comum de comunicação oral. É de difícil compreensão por não reproduzir com a necessária fidelidade

as características acústicas da fala. Falantes rápidos ou pouco claros e sons degradados estão muito presentes no cotidiano (CHABA, 1988; Marshall, 1981).

Distorções e Atenuações na Fala. Um modelo quantitativo (Plomp e Duquesnoy, 1980) propõe que a redução da acuidade auditiva para fala é devido a: (1) atenuação do som processado que reduz os níveis de intensidade tanto da fala quanto do ruído e (2) distorção que resulta na redução funcional da razão sinal-ruído. O fator distorção afeta primariamente a compreensão da fala em situações de ruído.

Ambientes Sonoros Complexos. O desempenho no reconhecimento da fala em condição de ruído ou reverberação apresenta um declínio progressivo em função da idade (CHABA, 1988; Gordon-Salant, 2005) e tem sido constatado por vários autores (Calais, Borges, Baraldi & Almeida, 2008; Dubno et al., 1984; Gordon-Salant, 1986; Veras & Matos, 2007).

Rajan, James e Cainer (2008) pesquisaram a habilidade de discriminação de fala em participantes com idade entre 20 e 69 anos, utilizando dois tipos de ruído: um constituído de várias pessoas falando ao mesmo tempo (murmúrio multifalante) e outro com espectro das sentenças faladas. Encontrou-se para o murmúrio multifalante: a) correlações significativas entre o limiar de recepção de fala de cada participante e a idade, b) para ambos os tipos de sentenças (com baixo e alto limiar de recepção de fala) foi necessária, para que pessoas na faixa etária de 60 anos discriminassem corretamente 50% das sentenças, uma razão sinal-ruído mais positiva do que a necessária para outras faixas etárias e c) uma diminuição na compreensão da fala, com o avançar da idade, inclusive para os idosos sem perda auditiva. Todavia, registrou-se que para o ruído semelhante ao espectro das sentenças não havia interações significativas.

O mascaramento informacional, que ocorre ao se ouvir o murmúrio multifalante e o sinal simultaneamente, gera dificuldade para diferenciar a voz alvo e o mascarador. O

mascarador mais perturbador para adultos mais velhos sem perda auditiva, segundo Helfer e Freyman (2008), era aquele em que o interlocutor do mesmo sexo falava ao mesmo tempo em que o interlocutor alvo. Adultos mais velhos com perda auditiva tiveram maior dificuldade na compreensão da fala alvo mesmo nos casos em que os interlocutores mascaradores eram do sexo oposto ao interlocutor alvo

Alterações Temporais da Fala. Estudos psicofísicos apontam que falas rápidas são mais difíceis de serem compreendidas pelos idosos, especialmente na presença de ruído, devido a um déficit importante no processamento das características temporais de estímulos sequenciais complexos. Ouvintes idosos com perda auditiva experimentam dificuldade para compreender falas rápidas por causa de dois fatores: o declínio no processamento da informação rápida associado a uma lentificação generalizada e a incapacidade de processar informações acústicas muito breves provindas das consoantes (Gordon-Salant, 2005).

Integridade do Ouvido e dos Sistemas Nervosos Auditivos Periféricos e Central

As diversas mudanças anatômicas que ocorrem no sistema auditivo do idoso interferem na eficácia do processamento da compreensão da fala. O envelhecimento do sistema auditivo está correlacionado também com várias alterações morfológicas e fisiológicas. Como mudanças morfológicas constatam-se: a) desestruturação das células ciliadas internas e externas e b) perda de células ganglionares espiraladas que formam as fibras do nervo auditivo. Nas mudanças fisiológicas identificam-se: a) diminuição na amplitude dos reflexos acústicos e dos potenciais de tronco encefálico, b) o prolongamento das latências das respostas do tronco encefálico e c) a diminuição da inibição do sistema nervoso auditivo central. Encontram-se, também, evidências sobre irregularidades no processamento da informação no que tange a: a) alterações no processamento e na resolução temporal, b) alterações no processamento espacial (i.e. localização de sons), c) diminuição

progressiva e acentuada da sensibilidade para frequências altas, d) declínio na resolução de frequências, e) aumento do limiar de audibilidade para tons puros e f) lentificação na velocidade de processamento (Syka, 2002). Idosos com perda auditiva neurosensorial para tons puros, que apresentam dificuldades de ouvir frequências médias de 4, 6 e 8 kHz, apresentam prejuízos na compreensão da fala (Calais et al., 2008; Elliott et al., 1985). Nota-se, observando o audiograma de fonemas feito por Russo e Behlau (1993) (vide Anexo A), que a maioria dos sons consonantais tem dominância de energia em frequências acima de 1000Hz. As consoantes são muito úteis como pista dentro das palavras. Pessoas com perdas auditivas neurosensoriais possuem dificuldades para diferenciar principalmente as consoantes que tem dominância de energia em frequências médias e altas; sem distinguir as consoantes torna-se difícil compreender as palavras (Helfer & Huntley, 1991). Por esse motivo é válido usar fonemas consonantais em tarefas de discriminação e reconhecimento.

Habilidades Cognitivas

A percepção auditiva da fala depende tanto do processamento do sinal que alcança o ouvido (*bottom-up*), quanto do repertório cognitivo e da experiência que a pessoa tem com a língua (*top-down*). Ao ouvir, a pessoa busca, no seu arcabouço de conhecimentos, informações sobre sintaxe, semântica, regras de conversação e, com base no seu entendimento do mundo, faz inferências sobre o que vai ser dito a seguir (CHABA, 1988).

Parte da diminuição progressiva da compreensão de fala associada à idade avançada é devido a um declínio na capacidade do processamento funcional, à deterioração da memória episódica secundária e à diminuição na velocidade dos processos mentais. Se o idoso tiver dificuldade expressiva de memória, mesmo que a perda auditiva seja moderada e ele utilize o aparelho amplificador o problema de compreensão de fala não será resolvido (Gordon Salant, 2005). O mesmo infere-se para o caso de aprendizado de pistas acústicas.

A Pertinência do Treinamento em Reconhecimento de Fala com Base em Evidência Biológica

Os mecanismos de plasticidade cortical e subcortical no sistema auditivo também justificam pesquisas sobre treinamentos. Há várias evidências de plasticidade, observada em animais e humanos, induzida pela experiência: a) alterações nos campos receptivos de neurônios a partir de estímulos condicionados, b) reorganização dos mapas de frequência do córtex auditivo primário por meio de condicionamento auditivo, c) diminuição do campo receptivo com aplicação de GABA e norepinefrina, d) modulação da afinação dos campos receptivos pela ativação do sistema colinérgico, e) inibição do potencial evocado para tom puro, f) reorganização dos mapas tonotópicos corticais por meio de treinamento utilizando a estimulação do núcleo basal magnocelular emparelhado com pulsos acústicos, g) mudanças na ativação das unidades corticais induzidas por microestimulação intracortical, h) aumento da área cortical auditiva correspondente à frequência estimulada por eletrodo, i) mudanças na interação funcional entre neurônios adjacentes dependentes do contexto comportamental dos estímulos indutores (veja detalhamento em Syka, 2002).

Medidas fisiológicas têm comprovado as mudanças plásticas após treinamento auditivo para discriminação de sílabas. Em treinamento de 6 dias em sessões de 30 minutos por dia, os participantes mostraram diminuição na amplitude das respostas na maior parte dos eletrodos, sugerindo que não são necessários treinamentos prolongados para produção de alterações fisiológicas no SNAC (Schechan, McArthur & Bishop, 2005).

A evidência de plasticidade do sistema auditivo a partir de condicionamentos e treinamentos fortalece a justificativa de pesquisas sobre percepção dos sons da fala e aprendizado de discriminação de fala (Syka, 2002; Robinson & Summerfield, 1996).

A Pertinência do Treinamento com Base Comportamental e as Razões para Desenvolver

Materiais para Treinamento

Sweetow e Palmer (2005) categorizaram treinamentos auditivos em analíticos, sintéticos e em analítico-sintéticos. Treinamentos analíticos envolvem os processamentos *bottom-up* e a prática de identificação de sons da fala, sendo mais comum o treinamento em reconhecimento de consoantes. Treinamentos sintéticos envolvem processos *top-down* e ganhos obtidos com o significado da mensagem por meio das várias estratégias de comunicação incluindo aperfeiçoamento da audição, atenção, uso de contexto, estratégias de reparo, entre outros. Dos dois tipos, o treinamento auditivo sintético é o mais benéfico para a audição em situação de ruído (Sweetow & Palmer, 2005). Entretanto, este dado não foi considerado de extrema relevância para o delineamento do método desta pesquisa em razão do pequeno número de pesquisas analisadas por estes autores. Além disso, observam-se na literatura muitos treinamentos analíticos eficazes como o de Cainer, James e Rajan (2008); Rajan, James e Cainer (2008); Burk e Humes (2007); Burk e Humes (2008).

Na avaliação de Carvalho (2007) o par mínimo, que significa um conjunto de duas palavras que se diferenciam por apenas um som, é o estímulo mais adequado para a avaliação da discriminação fonêmica, pois contorna a influência da coarticulação e usa palavras que possuem aspecto dinâmico da cadeia da fala. Yavas, Hernandorena e Lamprecht (apud Carvalho, 2007) afirmam que os fonemas discriminativos em um par mínimo podem variar segundo a quantidade de traços distintivos, unidades mínimas que compõem um segmento da língua (fonema). Os traços distintivos se distinguem em traços de classes principais e não-principais. São considerados de classes principais os traços [soante], [aproximante] e [vocóide], os quais compõem o Nó de Raiz dos segmentos. Traços de classes não-principais são os que abrangem o ponto de articulação, o modo de articulação e a sonoridade; estes

abrangem o Nó Laríngeo, o Nó de Cavidade Oral, Nó de Ponto. Estes podem ser [voz], [contínuo], [labial], [coronal], [dorsal] e [anterior].

Cainer, James e Rajan (2008) treinaram 58 adultos com audição normal em tarefas de discriminação de fala com a presença de dois tipos de ruído (ruído sintetizado no espectro médio das sentenças e murmúrio multifalante) em seis sessões sucessivas com intervalos de 15 minutos entre elas. Cada sessão usava uma lista diferente de 15 sentenças, totalizando 6 listas no treinamento. Para ambas as condições de ruído após o treinamento ocorreu uma diferença sutil, mas significativa entre eles, a situação em que o ruído é mais perturbador para o reconhecimento de fala, a de murmúrio multifalante, era a que os participantes tinham o melhor desempenho. Verificou-se que as discriminações das palavras alvo treinadas em ambas as situações de ruído podem ser aprendidas (Cainer, James & Rajan, 2008).

Burk e Humes (2008) pesquisaram o reconhecimento de palavras isoladas e sentenças (condição de ruído) em oito idosos (de 58 a 78 anos, portadores de perda auditiva entre leve e moderadamente severa) e observaram: a) um aumento de 40,2% no reconhecimento de palavras lexicalmente difíceis (i.e. com baixa frequência de ocorrência na língua e fonemas parecidos) e de 35% para palavras fáceis, com retenção dessas palavras treinadas mesmo com o treinamento de mais 75 palavras novas; b) a manutenção do aprendizado das 150 palavras totais após 3 meses e meio; c) boa generalização das palavras treinadas para falantes (locutores) não familiares, ou seja, que não participaram do treinamento; d) melhora do desempenho no reconhecimento de palavras isoladas e palavras chaves imersas em sentenças na condição de ruído; e) a possibilidade de o aprendizado ser estendido e potencializado para falas corridas, principalmente, com o treinamento de palavras isoladas e comuns ao dia a dia. Tremblay (2005) afirma que os resultados do treinamento auditivo são encorajadores, reforçam a importância de exercícios focados na escuta como elemento da reabilitação auditiva e mostram que percepção e fisiologia são passíveis de alteração.

Atualmente, no Brasil, encontram-se poucas pesquisas, Cd's de estímulos e softwares na área de reconhecimento de fala em condição de ruído, que possibilitem o treinamento e a verificação da eficácia desse tipo de intervenção para o falante do português brasileiro. Os motivos prováveis são a ausência de estudos, na nossa língua, que proponham estímulos e instrumentos adequados acusticamente para uso em condição de ruído e a ausência de protocolos adaptados ao contexto lingüístico brasileiro que favoreçam as pesquisas em treinamento auditivo em presença de ruído. A complexidade do tema é outro aspecto complicante, pois esse tipo de pesquisa exige investimento multidisciplinar envolvendo as áreas de fonética, processos cognitivos, psicofísica, acústica e lingüística, além de necessitar o uso de tecnologias para registro e tratamento dos estímulos. Por meio de busca na base de dados do Scielo em março de 2010, utilizando a palavra chave “treinamento auditivo” foi encontrado **um** artigo (Kozlowski, Wiemes, Magni & Silva, 2004); utilizando a palavra chave “discriminação de fala” encontrou-se **um** artigo (Freitas, Lopes & Costa, 2005); utilizando a palavra chave “reconhecimento de fala” encontrou-se **um** artigo (Neves & Feitosa, 2003) e “reconhecimento de palavras” **dois** artigos (Lúcio, Pinheiro & Nascimento, 2009; Guimarães, 2004) abordando temas incomuns a esta pesquisa. Realizando buscas sobre treinamento e discriminação auditiva em outros bancos de dados (e.g. Portal de Domínio Público da Capes) encontraram-se **sete** dissertações de mestrado (Lewkovicz, 2008; Carvalho, 2007; Padilha, 2008; Schröder, 2008; Raimundo, 2008; Souza, 2010; Jorge, 2008), **uma** tese de doutorado (Costa, 1997) e **uma** monografia de conclusão de curso (Jorge, 2008), alguns relacionados ao tema dessa pesquisa. Observam-se nesses trabalhos que são utilizados para discriminação os estímulos: Palavras Monossilábicas, Sílabas sem Sentido, Listas de Sentenças em Português (LSP) (Costa, 1997), e os testes: Teste Dicótico de Dissílabos Alternados (SSW) (Borges, 1997), Teste monótico de fala com ruído (Pereira & Schochat, 1997), Teste de identificação de sentenças sintéticas com mensagem competitiva ipsilateral (Aquino, Almeida & Oliveira,

1993). Carvalho (2007) desenvolveu o Teste de Figuras para Discriminação Fonêmica para crianças. Dirigidos para o contexto brasileiro a autora cita o teste elaborado por Rodrigues chamado Prova para Avaliar a Discriminação Auditiva, que utiliza pares de sílabas sem sentido com diferença de apenas um fonema, e o teste, adaptado para o português, “Teste de Figuras para Discriminação Auditiva” (Mota, Keske-Soares & Vieira, 2000), o qual utiliza como estímulos 25 pares mínimos de palavras. Observa-se que os testes existentes mostram-se, usualmente, voltados para avaliação de discriminação auditiva e apresentam um número pequeno de estímulos o que dificulta sua adaptação para uso em treinamento.

O objetivo geral deste trabalho é propiciar condições para o desenvolvimento de um protocolo adequado para futuros treinamentos auditivos, visto que os materiais existentes tem objetivos diagnósticos. Os objetivos específicos são: 1) desenvolver um banco de estímulos adequados para uso em treinamento auditivo de base comportamental para idosos falantes do português brasileiro, em condições desfavoráveis e usuais de interação social, ou seja, em presença de ruído mascarador do tipo murmúrio multifalante e 2) analisar e comparar as características acústicas de uma amostra de pares mínimos com os dados disponíveis na literatura para validar acusticamente a propriedade da seleção dos pares mínimos.

Método

Delineamento do Estudo

O trabalho de Mota (2001) foi escolhido como fonte para a escolha das palavras-estímulo por apresentar uma lista ampla de 284 pares mínimos abrangendo todos os fonemas do português brasileiro e algumas especificidades técnicas, destacadas por outros autores como mais apropriadas para a atividade proposta, ou seja, as palavras que compuseram os pares tinham grande frequência de ocorrência em nossa língua e eram palavras fáceis, usadas no cotidiano, atendendo as recomendações de Burk & Humes, (2008). Foram escolhidos estímulos tanto para tarefas de familiarização, importante para o aprendizado da tarefa, quanto para atividades de treinamento.

Para as atividades de familiarização, foram selecionados pares mínimos contendo fonemas que se diferenciavam por quatro e cinco traços a fim de facilitar o aprendizado da tarefa de reconhecimento de palavras, tendo em vista que, quanto maior o número de traços distintivos, maior a facilidade de reconhecimento de palavras. Para a tarefa de treinamento, foram escolhidos um ou dois exemplos de cada contraste de pares mínimos que envolvessem fonemas com até três traços. Dessa forma, haveria uma representatividade significativamente homogênea de todos os fonemas do português brasileiro nos variados contrastes, ou seja, os fonemas seriam apresentados em quantidade parecida. Isto asseguraria que, aos participantes seriam apresentados igualmente a todos os fonemas de forma equivalente o que facilitaria a análise do desempenho na tarefa.

Foi escolhido um texto contemporâneo de Clarice Lispector como base para criar o murmúrio multifalante porque continha poucas palavras com carga afetiva e dessa maneira

seria possível controlar parcialmente a variável cognitiva durante a realização da tarefa. Este é o tipo de ruído mais prejudicial para a compreensão da fala (Burk & Humes, 2008).

Esta pesquisa não forneceu riscos à saúde de seus participantes e foi aprovada pelo Comitê de Ética da Faculdade de Ciências da Saúde da Universidade de Brasília (protocolo nº 055/2009).

Participantes

Participaram da gravação das palavras integrantes dos pares mínimos (sinal): uma mulher graduada em Letras de 29 anos, residente em Brasília há 28 anos e um homem de 20 anos, estudante de Comunicação Social e residente em Brasília há 20 anos. Para gravação do murmúrio a ser utilizados nas tarefas de familiarização e treinamento, participaram quatro pessoas: duas do sexo feminino (15 e 27 anos) e duas do sexo masculino (20 e 55 anos), estudantes ou trabalhadoras. Os critérios de seleção para doador de voz para o sinal foram: ter formação em áreas que valorizam a expressão da comunicação verbal, ser residente por longo tempo nesta unidade da federação e para o murmúrio multifalante o critério para composição do grupo foi incluir ambos os sexos e uma faixa etária ampla.

Equipamentos e Ambiente de Gravação dos Estímulos

Todos os estímulos foram gravados no Laboratório de Estúdio de Rádio da Faculdade de Comunicação Social da UNB, em ambiente acusticamente tratado. As palavras (sinal) e o murmúrio multifalante (mascarador) foram gravados na configuração Mono DCM, Wave File. Utilizaram-se dois microfones para gravação do murmúrio: Eletrovoice e Neumann P48A. Somente este último foi utilizado para gravação das palavras. O distribuidor

de áudio era o D35 Apel e os estímulos foram gravados no PC por meio do programa Sound Forge 8.0. O murmúrio multifalante foi finalizado pelo programa Sony Vegas. A placa de áudio era do modelo Delta 1010 Professional. A mesa de controle do som era da marca Tascam 2600, caixa de referência Alesis e monitor One.

Procedimentos

Preparação dos estímulos e critérios de escolha

Sinal. Foram criadas duas listas de pares mínimos a partir da lista de Mota (2001). Uma lista de 8 pares foi criada para tarefa de familiarização e outra de 129 pares para tarefa de treinamento. Para a tarefa de familiarização foram selecionados pares mínimos de quatro e cinco traços que não estivessem presentes na tarefa de treinamento, para assegurar que não haveria aprendizagem prévia dos fonemas que seriam treinados. Para a seleção da lista de palavras de treinamento utilizou-se os seguintes critérios: todos os fonemas do português brasileiro deveriam ser usados, os fonemas deveriam diferir entre si pelo número máximo de três traços distintivos e a distribuição de fonemas deveria ser homogênea. Os fonemas foram classificados por faixas de frequências em razão da expectativa de ocorrer diferença de desempenho associada à região de frequência demandada na tarefa de reconhecimento da seguinte forma: baixa - frequências de 250 a 1000Hz, média - frequências de 1000 a 4000Hz e alta - frequências de 4000 a 8000Hz.

Mascarador. O texto selecionado para desenvolvimento do mascarador foi a crônica “Restos de carnaval”, de Clarice Lispector, escolhida por ser um texto contemporâneo, em linguagem moderna e direta, composto de palavras de uso comum na língua e ser em sua maioria com carga emocional neutra.

O texto passou por uma série de manipulações para torná-lo sem sentido e desprovido de carga emocional, preservando sua estrutura. Para isso, foram trocados alguns verbos e substantivos por outros dentro do texto original. As palavras foram trocadas entre os parágrafos e dentro de um mesmo parágrafo. Alguns substantivos foram retirados por conterem carga afetiva. Destes, alguns foram substituídos porque sua falta prejudicava a estrutura da sentença. A regra para a escolha do novo substantivo foi a relação fonética ou relação semântica com a palavra anterior, entretanto, com carga afetiva neutra. Palavras que representavam sentimentos foram substituídas por outras com pouca ou nenhuma carga afetiva. Os nomes próprios foram retirados pelo potencial de evocar lembranças e não foram substituídos. A troca dos verbos ocorreu utilizando os verbos do próprio texto, respeitando a regência verbal e nominal e conservando o tempo verbal dos verbos do texto original. Alguns verbos foram repetidos por não haver disponibilidade no texto de verbos que preservassem a estrutura da sentença ou por serem verbos naturalmente repetidos ao longo do texto. Do texto sem sentido resultante (vide Anexo C) foram criados mais três textos sem sentido, pela reorganização da sequência de parágrafos, para a criação do murmúrio multifalante.

O mascarador foi, então, constituído de uma mistura de quatro vozes de falantes que liam os diferentes textos sem sentido em voz alta e simultaneamente. As quatro vozes foram duplicadas digitalmente aparentando oito vozes, para minimizar interferências fonéticas, semânticas e lexicais que poucas vozes poderia causar (Cainer, James e Rajan, 2008).

Gravação dos estímulos

As gravações dos estímulos ocorreram durante o dia. Pela manhã gravou-se o murmúrio multifalante. Antes da gravação do murmúrio foram dadas aos participantes orientações técnicas de como realizar a tarefa. Foi realizada uma gravação-teste e depois a gravação válida dos textos (vide Anexo C) que perdurou 20 minutos. À tarde foi feita a

gravação do sinal, com duração aproximada de uma hora e vinte minutos.

Análise acústica

Para análise acústica foi escolhida arbitrariamente, uma amostra de 18 pares mínimos da lista B, tendo em vista que as três listas eram equivalentes, porque abrangiam todos os fonemas do português brasileiro, e balanceadas quanto à quantidade dos fonemas discriminativos. O número de pares analisados não pôde ser maior porque o escopo da dissertação não permite um número muito grande de pares, entretanto este número é representativo para o objetivo proposto. O critério para selecionar os pares mínimos a serem submetidos à análise acústica foi de que o fonema discriminativo estivesse na posição inicial da palavra para minimizar efeitos de coarticulação e permitir uma análise acústica mais precisa.

Escolheu-se o software PRAAT 5.1.32 para análise acústica por ser o indicado por especialistas da área, ser utilizado em contexto de pesquisa e ter vários recursos de análise de dados acústicos. O software, desenvolvido por Paul Boersma e David Weenink do Institute of Phonetic Sciences da Universidade de Amsterdã, é acompanhado de manuais explicativos escritos pelos autores e por outros especialistas.

Utilizando o comando apropriado é possível obter quantitativamente vários dados para análise de um arquivo de áudio, como a altura tonal, a intensidade média, a frequência, o número dos formantes, dentre outros recursos. É possível também observar vários registros gráficos como o contorno dos formantes. Neste trabalho os dados acústicos foram obtidos por meio dos recursos fornecidos pelo programa e por inspeção visual. As medidas diretas obtidas com o software foram: a variação da intensidade, variação da altura tonal e o tempo de duração dos fonemas discriminativos. As medidas indiretas obtidas por meio de inspeção visual foram: o tipo de modulação da frequência (observado a partir de todos os formantes do

fonema discriminativo), a taxa de modulação de frequência (observado somente no fonema discriminativo), a região de frequência com maior energia no fonema discriminativo e o local da modulação (observado somente no fonema discriminativo). Os dados foram sistematizados em nove tabelas, dois pares mínimos por tabela. Para análise acústica de cada palavra do par foram utilizados três tipos de representação gráfica, duas podem ser vistas na Figura 1 e uma na Figura 2.

Na Figura 1 observam-se as descrições acústicas contidas nos espectrogramas criados pelo PRAAT. O painel superior dessa figura contém um espectrograma com a descrição da distribuição de energia ao longo da emissão vocal e o painel inferior contém um espectrograma com a descrição do espectro de frequência e de intensidade relativa ao longo do tempo. Os dados de frequência estão representados em vermelho tanto na função quanto na indicação da escala à esquerda. Os dados de altura tonal estão representados em azul, tanto na função quanto na escala à direita. A intensidade é representada por função em cor amarela e por números em cor verde na escala do lado direito. O contorno dos formantes é identificado em vermelho. A linha vermelha hachurada vertical no meio da figura mostra a descrição do tempo percorrido durante a emissão da palavra no ponto escolhido pelo usuário (parte inferior da imagem). Na parte inferior, está indicada também a duração total do som, por vezes incluindo um curto tempo de silêncio no início e no final da palavra.

Utilizou-se também o tipo de espectrograma visto na Figura 2. Este permitiu uma análise mais precisa da distribuição de frequências presente no fonema discriminativo. A escala do lado esquerdo mostra as frequências no intervalo entre 0 e 20.000Hz (frequência nominal), que possibilitou a identificação de formantes de ordem superior (4° e 5°) dentro do intervalo de audibilidade de frequências do ouvido humano.

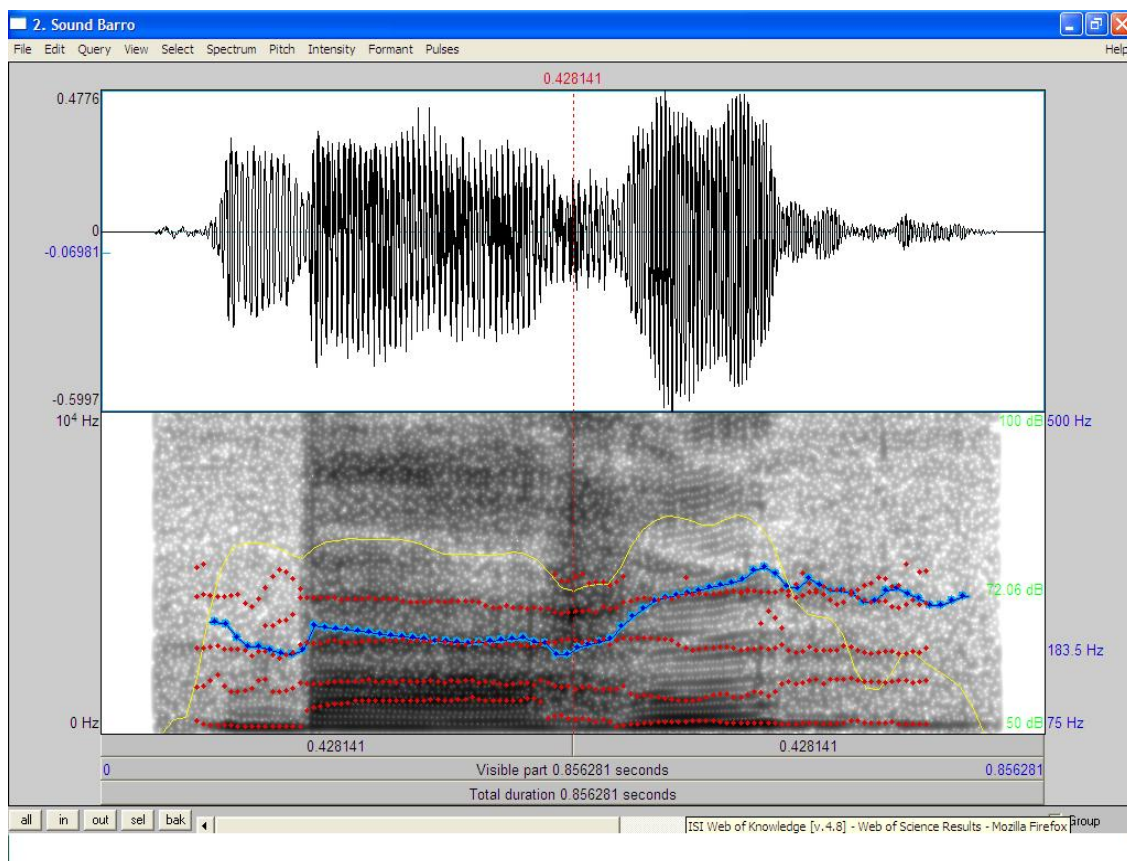


Figura 1 – Espectrograma da palavra barro

Barro

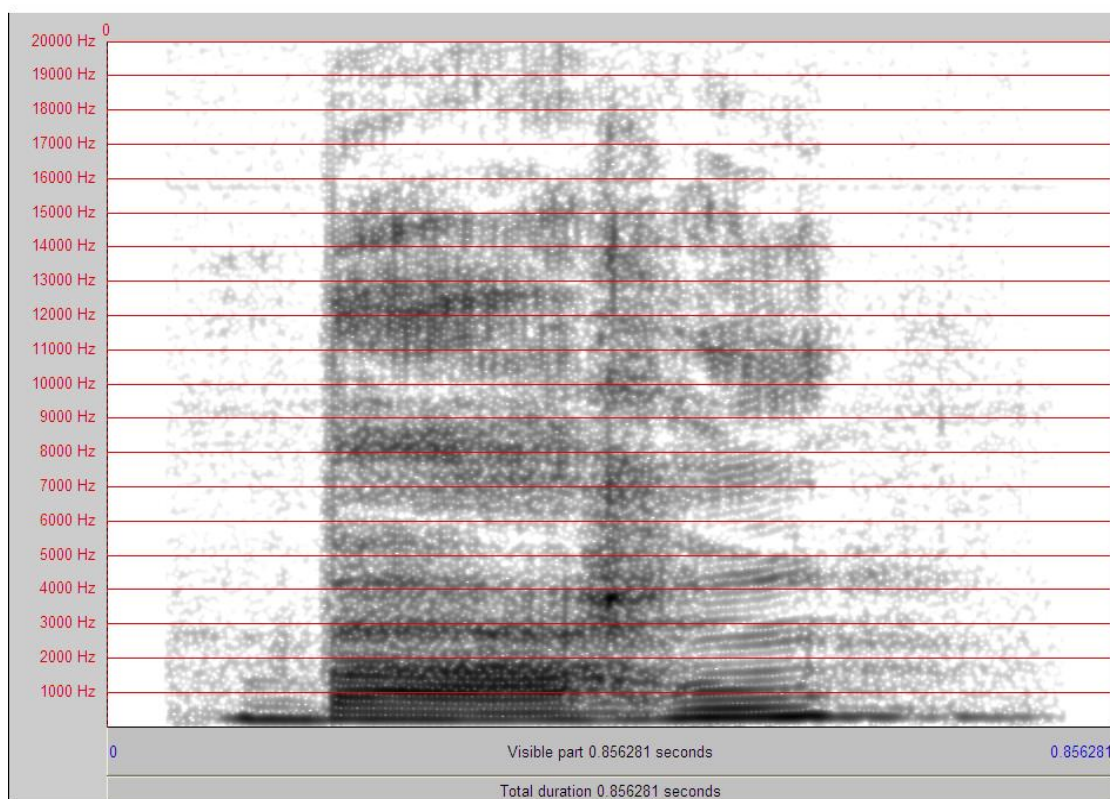


Figura 2 – Espectrograma mostrando a distribuição de energia no intervalo de 0 a 20.000Hz

Resultados

Análise Fonética das Palavras

Na Tabela 1 estão descritos todos os fonemas dos pares mínimos estudados de acordo com sua posição (início ou meio) na palavra. Na primeira coluna são apresentados os fonemas do português brasileiro segundo a lista de Mota (2001). Na segunda coluna, apresenta-se o número de primeiras palavras dos pares selecionados com os fonemas respectivos da primeira coluna. Nas duas colunas seguintes informam-se quantas primeiras palavras do par possuem os respectivos fonemas na posição inicial e na posição medial da palavra. A quinta coluna apresenta o número de segundas palavras do par que têm os fonemas referentes à primeira coluna. Na sexta e sétima coluna detalha-se quantas segundas palavras do par têm os respectivos fonemas na posição inicial e medial. Na penúltima coluna temos a soma de quantas vezes cada fonema é apresentado. Na última coluna informa-se a faixa de frequência de cada fonema segundo o audiograma de fonemas (Russo & Behlau, 1993). Nessa coluna as letras representam faixas de frequência: **b** baixas, de 250 a 1000Hz, **m** médias, de 1000 a 4000Hz, e **a** altas, de 4000 a 8000Hz. Como diferentes tipos de perda auditiva podem se caracterizar por elevação de limiar em diferentes regiões de frequência, e pessoas com presbiacusia têm dificuldades para frequências médias e altas, a organização em faixas de frequência pode auxiliar saber quais fonemas tiveram pior ou melhor desempenho no treinamento e, também, pode facilitar a análise dos dados após treinamento (análise detalhada dos valores da frequência dos fonemas discriminativos dos pares mínimos são apresentados na sessão Análise Acústica dos Pares Mínimos).

Observando o audiograma de fonemas detectam-se apenas cinco fonemas com características de alta frequência, enquanto de frequências médias e baixas totalizam-se 14

fonemas. Foi possível balancear os pares da lista pela frequência de ocorrência dos diferentes fonemas (Tabela 1), entretanto o baixo número de fonemas de alta frequência no português brasileiro não permitiu compor uma lista com efetivo balanceamento por combinações de faixas de frequências (Tabelas 2a e 2b). Ressalta-se que a maior parte das palavras tem o fonema discriminativo na posição inicial.

Tabela 1 – Fonemas que fazem o contraste entre diferentes pares de palavras, segundo o número de ocorrências, sua posição inicial (PI) ou medial (PM) na 1ª e 2ª palavra do par e sua faixa de frequência característica.

Fonemas	Total 1ª pal.	1ª palavra		Total 2ª pal.	2ª palavra		Total 1ª+2ª	Fx freq.
		PI	PM		PI	PM		
/p/	12	12	0	0	0	0	12	b
/b/	13	9	4	1	1	0	14	b
/t/	12	10	2	2	2	0	14	a
/d/	13	6	7	2	2	0	15	m
/k/	11	6	5	4	4	0	15	m
/g/	8	6	2	6	6	0	14	m
/f/	9	9	0	6	6	0	15	a
/v/	7	5	2	7	6	1	14	a
/s/	4	3	1	7	7	0	11	a
/z/	3	1	2	8	4	4	11	a
/ʃ/	6	3	3	8	3	5	14	m
/ʒ/	2	0	2	12	11	1	14	m
/m/	4	4	0	9	7	2	13	b
/n/	4	1	3	10	6	4	14	b
/ɲ/	3	0	3	9	0	9	12	b
/l/	4	0	4	9	5	4	13	b
/ʎ/	6	0	6	5	0	5	11	b
/r/	3	3	0	10	0	10	13	b
/R/	0	0	0	13	2	11	13	m

Na Tabela 2a, são apresentados, na primeira coluna, os fonemas discriminativos do português brasileiro utilizados na composição dos pares mínimos e entre parênteses o número de vezes em que o fonema apareceu na primeira palavra do par como fonema a ser discriminado. Nas demais colunas são descritas todas as combinações de pares mínimos verificadas, segundo as faixas de frequência características no par, por exemplo, o fonema /p/ ocorria 12 vezes como fonema a ser discriminado dentro do par, e em três pares mínimos a palavra com esse fonema foi combinada com outra palavra contendo um fonema com concentração de energia em baixa frequência, segundo a classificação identificada na Tabela 1. Descrição complementar encontra-se na Tabela 2b, tendo como referencial o fonema apresentado na segunda palavra do par.

O objetivo da criação dessas tabelas foi procurar balancear as apresentações das palavras por combinação de faixa de frequências. Observa-se nas Tabelas 2a e 2b que o maior número de combinações de faixa de frequências eram de pares mínimos contendo fonema discriminativo com grande energia em baixas frequências em ambas as palavras do par (combinação b/b). Não foi possível balancear o número de apresentações da combinação de faixa de frequências b/b porque o número de fonemas em português brasileiro com concentração de energia em frequências baixas é bem maior do que o número de fonemas com concentração de energia em altas frequências. Entretanto, observa-se que em ambas as tabelas as combinações de faixa de frequências variam entre 10 e 16 apresentações, mostrando que foi possível homogeneizar a quantidade de apresentações por combinações de faixa de frequências.

Tabela 2a – Distribuição de ocorrências das combinações de faixas de frequências para o fonema discriminativo integrando a 1ª palavra do par.

Fonemas	Faixas de frequências								
	b-b	b-m	b-a	m-b	m-m	m-a	a-b	a-m	a-a
/p/ (12)	3	4	5						
/b/ (13)	4	4	5						
/t/ (12)							3	5	4
/d/ (13)				7	3	3			
/k/ (11)				2	5	4			
/g/ (8)				4	1	3			
/f/ (9)							1	4	4
/v/ (7)							3	2	2
/s/ (4)							1	2	1
/z/ (3)							2	1	
/ʃ/ (6)				1	5				
/ʒ/ (2)				2					
/m/ (4)	4								
/n/ (4)	4								
/ŋ/ (3)	3								
/l/ (4)	3	1							
/ʎ/ (6)	4	2							
/r/ (3)		3							
/R/ (0)									
Total	25	14	10	16	14	10	10	14	11

Tabela 2b – Distribuição de ocorrências das combinações de faixas de frequência para o fonema discriminativo integrando a 2ª palavra do par.

Fonemas	Faixas de Frequências								
	b-b	b-m	b-a	m-b	m-m	m-a	a-b	a-m	a-a
/p/ (0)									
/b/ (1)	1								
/t/ (2)			2						
/d/ (2)		1						1	
/k/ (4)		2			1			1	
/g/ (6)		2			3			1	
/f/ (6)			2			3			1
/v/ (7)			2			2			3
/s/ (7)			2			2			3
/z/ (8)			2			3			3
/ʃ/ (8)		1			2			5	
/ʒ/ (12)		2			4			6	
/m/ (9)	2			3			4		
/n/ (10)	3			4			3		
/ŋ/ (9)	5			3			1		
/l/ (9)	5			2			2		
/ʎ/ (5)	4			1					
/r/ (10)	5			2			3		
/R/ (13)		7			5			1	
Total	25	15	10	15	15	10	13	15	10

Análise Qualitativa e Critérios de Seleção Final dos Estímulos Gravados

Após a audição de todos os pares mínimos gravados selecionaram-se aqueles que apresentaram todos os elementos acústicos preservados e a consistência da pronúncia no par. Foram excluídos oito pares mínimos dos 129 pares por conterem uma das palavras do par com baixa qualidade técnica de gravação ou mais de um fonema de distinção entre as palavras do par. Foi excluída a gravação do sinal feita pelo participante masculino em razão da elevada variabilidade na característica de locução de uma palavra para outra do par descaracterizando a prosódia usual da comunicação.

Os 121 pares mínimos foram divididos em três listas: B, C e D (vide Anexo B) para propiciar intervalos de descanso e evitar que o participante ficasse exposto a alguns fonemas apenas na última lista quando já estivesse cansado e com atenção rebaixada. Procurou-se assegurar que cada lista contivesse todos os fonemas do português brasileiro e fosse homogênea na quantidade de contrastes por faixa de frequência.

Análise Acústica dos Pares Mínimos

Nesta seção estão descritas as análises acústicas realizadas com o auxílio dos espectrogramas das palavras contidas na amostra de pares mínimos. Foram utilizadas as seguintes siglas nestas descrições: TMF1 - Tipo de Modulação de Frequência, TMF2 - Taxa de modulação de frequência, $\approx I$ - Variação de intensidade, $\approx AT$ - Variação de Altura Tonal, LM - Local da Modulação, TD - Tempo de Duração, FD - fonema discriminativo. As figuras e tabelas correspondentes às análises acústicas encontram-se após os Anexos.

Par barro-jarro. Observando a Figura 3 e a Tabela 3 verifica-se que os fonemas discriminativos /b/ e /ʒ/ (Figura 3 e Tabela 3) têm parecidos os seguintes aspectos: o TMF1, o

número de formantes visíveis, e $\approx I$. Além disso, podem-se observar as seguintes diferenças nos fonemas /b/ e /z/: o TD do fonema /z/ é maior que no /b/, o LM no fonema /b/ é no final e no /z/ é no início e no final deste fonema, o fonema /z/ tem AT um pouco mais aguda que o fonema /b/.

Par bola-mola. Na Figura 3 e Tabela 3 observa-se que os fonemas discriminativos /m/ e /b/ têm parecidos os seguintes aspectos: TD, TMF1, $\approx I$, $\approx AT$ e região de frequência com maior energia no fonema discriminativo. Além disso, podem-se observar as seguintes diferenças nos fonemas /m/ e /b/: TMF1 do primeiro formante é ascendente no fonema /b/ e descendente no /m/, o LM no fonema /b/ está no final e no /m/ no início e observam-se três fonemas visíveis no fonema /m/.

Par fila-vila. Observando a Figura 4 e a Tabela 4 verifica-se que os fonemas discriminativos /f/ e /v/ têm parecida a região de frequência com maior energia (altas). Além disso, podem-se observar as seguintes diferenças nos fonemas /f/ e /v/: o TD é maior no fonema /f/, o TMF1 do fonema /f/ não é específica e o do /v/ é descendente, a modulação no fonema /f/ compreende todo o fonema e o LM no /v/ é apenas no início, o espectro do fonema /f/ é amplo, o fonema /v/ tem um formante visível e tem uma $\approx I$ menor que o fonema /f/, a AT do fonema /f/ é mais aguda que a do fonema /v/.

Par pó-vó. Observando a Figura 4 e Tabela 4 verifica-se que os fonemas discriminativos /p/ e /v/ não têm semelhanças. Observam-se as seguintes diferenças nos fonemas /p/ e /v/: o TD do fonema /p/ é menor que do fonema /v/, o TMF1 deste fonema não é específico, a TMF2 é alta no fonema /p/ e baixa no /v/, a modulação do fonema /p/ ocorre em todo fonema, e, em /v/ o LM é no início e no final, no fonema /p/ não há formantes visíveis e o fonema /v/ têm dois formantes visíveis (com bastante energia), a $\approx I$ e a $\approx AT$ é mais aguda no fonema /v/, o fonema /p/ contém maior energia em baixas frequências e o fonema /v/ em altas frequências.

Par pão-chão. Observando a Figura 5 e a Tabela 5 verifica-se que os fonemas discriminativos /p/ e /ʃ/ têm parecidos a TMF2 e $\approx I$. Todos os outros itens como TD, TMF1, LM, número de formantes visíveis, $\approx AT$ e região de frequência com maior energia no fonema discriminativo aparentam diferenças. O fonema /p/ tem o TD bem mais curto que /ʃ/, /p/ tem modulação múltipla e /ʃ/ a modulação não é específica, o LM no fonema/p/ é no início e em /ʃ/ é em todo fonema, são 4 o número de formantes visíveis no fonema /p/, o fonema /ʃ/ tem espectro amplo, a altura tonal é mais aguda em /ʃ/ e este abrange um intervalo de frequências maior que o outro fonema.

Par xiz-giz. Observando a Figura 5 e a Tabela 5 verifica-se que os fonemas discriminativos /ʃ/ e /z/ têm parecidos a região de frequência com maior energia. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, a TMF2, o LM, o número de formantes com bastante energia e a $\approx AT$. O fonema /ʃ/ tem TD mais longo que /z/, não há modulação de frequência no fonema /z/, todo o fonema /ʃ/ apresenta modulação de frequência, a TMF2 /ʃ/ é alta, há um formante visível no fonema /z/ e /ʃ/ tem o espectro amplo. A $\approx I$ e a $\approx AT$ são maiores no fonema /z/.

Par couro-touro. Observando a Figura 6 e a Tabela 6 verifica-se que os fonemas discriminativos /k/ e /t/ tem parecido a TMF2. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, o LM, o número de formantes com bastante energia e a $\approx AT$. O fonema /k/ tem TD mais longo que o /t/, não é possível identificar o primeiro formante do fonema /k/, a $\approx I$ é maior no fonema /k/, este fonema tem $\approx AT$ mais aguda e contém maior energia em regiões de frequências altas.

Par data-nata. Observando a Figura 6 e Tabela 6 verifica-se que os fonemas discriminativos /d/ e /n/ têm parecidos o LM, a TMF2 e a $\approx AT$. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, a TMF2 e o número de formantes com

bastante energia. O TD do fonema /n/ é mais longo que o /d/, o fonema /d/ tem um formante visível e o /n/ três formantes visíveis, o fonema /n/ tem uma AT um pouco mais aguda.

Par tia-dia. Observando a Figura 7 e a Tabela 7 verifica-se que os fonemas discriminativos /t/ e /d/ têm parecidos a região de frequência com maior energia. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, a TMF2, o LM, o número de formantes com bastante energia e a \approx AT. O TD do fonema /t/ é mais curto do que em /d/, a modulação no fonema /t/ no primeiro formante não é específica e no fonema /d/ é descendente, a TMF2 é alta no fonema /t/ e baixa no /d/, a modulação compreende todo fonema /t/ e ocorre apenas no início do /d/, o espectro é amplo no fonema /t/, no fonema /d/ é visível apenas o primeiro formante, o fonema /d/ tem a \approx I e a \approx AT maior que o fonema /t/ sendo que o fonema /t/ tem AT mais aguda que o /d/.

Par mapa-tapa. Observando a Figura 7 e a Tabela 7 verifica-se que os fonemas discriminativos /m/ e /t/ têm parecidos os seguintes aspectos: o TMF1, a TMF2 e a \approx I. Além disso, observam-se as seguintes diferenças nos fonemas /m/ e /t/: o TD no fonema /m/ é maior, o LM no fonema /m/ é no início e no meio do fonema discriminativo e no /t/ há modulação em todo fonema, o fonema o /m/ aparenta ter três formantes visíveis, o fonema /t/ tem espectro amplo, e, a AT do fonema /t/ é mais aguda que a de /m/, no fonema /t/ há bastante energia nas frequências altas.

Par ventar-sentar. Observando a Figura 8 e a Tabela 8 verifica-se que os fonemas discriminativos /v/ e /s/ têm parecida a região de frequência com maior energia. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, a TMF2, o LM, o número de formantes com bastante energia e \approx a AT. O TD do fonema /v/ é mais curto que o do fonema /s/, no fonema /v/ não se observa modulação de frequência, no fonema /s/ ocorre modulação múltipla em todo fonema, há dois formantes visíveis no fonema /v/ e nenhum no

/s/, pois, ele abrange um espectro amplo, o fonema /s/ tem $\approx I$ maior e a altura tonal deste fonema é mais aguda.

Par vinho-ninho. Observando a Figura 8 e a Tabela 8 verifica-se que os fonemas discriminativos /v/ e /n/ têm parecidos os seguintes aspectos: o TD, o TMF1, o número de formantes e a $\approx I$. Além disso, podem-se observar as seguintes diferenças nos fonemas /v/ e /n/: a modulação da frequência é alta no fonema /v/ e baixa no fonema /n/, o LM compreende o início e o meio do fonema /v/ e o início do fonema /n/, a AT é mais aguda no fonema /n/ e o /v/ contém mais energia em regiões de baixas frequências diferentemente do outro fonema.

Par mão-não. Observando a Figura 9 e a Tabela 9 verifica-se que os fonemas discriminativos /m/ e /n/ têm parecidos os seguintes aspectos: o TD, o número de formantes, $\approx I$, a $\approx AT$ e a região de frequência com bastante energia. Podem-se observar as seguintes diferenças nos fonemas /m/ e /n/: o fonema /m/ não aparenta ter modulação e o fonema /n/ tem modulação múltipla, o LM no fonema /n/ é no início e ambos os fonemas contém muita energia em frequências baixas.

Par fanta-janta. Observando a Figura 9 e a Tabela 9 verifica-se que os fonemas discriminativos /f/ e /z/ têm parecidos os seguintes aspectos: o TD, o TMF1, o LM, o número de formantes e a região de frequências com maior energia. Pode ser observado as seguintes diferenças nos fonemas /f/ e /z/: a AT do fonema /z/ varia bastante e parece mais aguda e a $\approx I$ parece ser pequena neste fonema.

Par bote-pote. Observando a Figura 10 e a Tabela 10 verifica-se que os fonemas discriminativos /b/ e /p/ têm parecidos a quantidade de energia em regiões de frequências baixas. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, a TMF2, o LM, o número de formantes com bastante energia e a $\approx AT$. O fonema /b/ tem duração maior que o /p/, não observa-se modulação no fonema /b/, entretanto, o fonema /p/ tem modulação ascendente localizada no final do fonema e um formante visível, o fonema /b/

não tem nenhum formante visível e o fonema /p/ tem um, a $\approx I$ é maior no fonema /b/ e a $\approx AT$ aparenta ser mais aguda no fonema /b/.

Par cinco-zinco. Observando a Figura 10 e a Tabela 10 verifica-se que os fonemas discriminativos /s/ e /z/ têm parecidos os seguintes aspectos: o LM, o número de formantes visíveis e a área de frequência com maior energia. Podem-se observar as seguintes diferenças nos fonemas /s/ e /z/: o TD do fonema /s/ é maior, a modulação no fonema /s/ é múltipla e no fonema /z/ é descendente, a $\approx I$ é maior no fonema /s/. O PRAAT não foi detalhou a altura tonal no fonema /s/.

Par tela-zela. Observa-se que os fonemas discriminativos /t/ e /z/ (Figura 11 e Tabela 11) têm parecidos os seguintes aspectos: a $\approx I$ a e região de frequência com maior energia. Em todos os outros itens observam-se aspectos diferentes como o TD, o TMF1, a TMF2, o LM, o número de formantes com bastante energia, a $\approx AT$. O TD no fonema /z/ é maior que no fonema /t/; a modulação é múltipla no fonema /t/ e ascendente no fonema /z/, no fonema /t/ a modulação compreende todo o fonema e no fonema /z/ é apenas no final; não é possível visualizar nenhum formante no fonema /t/, o fonema /z/ tem um formante visível; ocorre maior $\approx AT$ no fonema /z/ e AT parece ser mais aguda no fonema /t/.

Par conta-ponta. Observa-se que os fonemas discriminativos /k/ e /p/ (Figura 11 e Tabela 11) têm parecidos os seguintes aspectos: a TMF2, o número de formantes visíveis e a $\approx I$. Além disso, pode-se observar as seguintes diferenças nos fonemas /k/ e /p/: o TD no fonema /k/ é maior do que no fonema /p/, a modulação compreende todo o fonema /k/ e ocorre apenas no final do fonema /p/, a AT parece ser mais aguda no fonema /k/ e varia mais no fonema /p/.

Discussão

Pessoas com perda auditiva neurossensorial, e em especial idosos com presbiacusia, apresentam dificuldade para ouvir sons com concentração de energia em frequências altas. Usualmente, observam-se fonemas consonantais com estas características. Como estes fonemas servem como pista acústica para compreensão de palavras a dificuldade para distinguir as consoantes compromete a boa compreensão da fala. Por esta razão, é adequada a escolha de pares mínimos com fonema discriminativo consonantal como estímulos para treinamento em reconhecimento de fala. Acredita-se que esta lista tenha a quantidade de pares mínimos adequada para treinamento visto que Burk e Humes (2008) utilizam 150 palavras isoladas e o banco desenvolvido no presente trabalho apresenta 222 palavras.

Ao se comparar as faixas de frequências com maior energia nos fonemas discriminativos que integram as palavras selecionadas, verificam-se diferenças entre o audiograma de Russo e Behlau (1993) e os fonemas utilizados nas análises acústicas feitas neste trabalho. Detectaram-se diferenças quanto ao intervalo de frequências nos fonemas discriminativos das seguintes palavras: barro, jarro, pão, chão, xiz, giz, couro, touro, data, tia, mapa, tapa, ventar, ninho, mão, não, fanta, pote, cinco, zinco, tela, zela e conta. Estas diferenças incluem a abrangência de uma maior faixa de frequências dos fonemas, tipicamente sinalizando uma participação maior das frequências médias e altas nos fonemas discriminativos dos pares analisados neste trabalho. Em razão dos fonemas discriminativos abrangerem uma região de frequência maior que a sugerida no audiograma de Russo e Behlau (1993), acredita-se que isto possa contribuir para o entendimento da dificuldade de reconhecimento em alguns fonemas dos pares mínimos para presbiacúsicos. Provavelmente, estas diferenças observadas nos espectros de frequências ocorrem porque a metodologia utilizada nos dois estudos é diferente quanto ao intervalo de frequências observado e os

estímulos analisados (e.g. fonema analisado isoladamente ou integrando uma estrutura Consoante Vogal). Os fonemas discriminativos dos pares mínimos restantes se relacionam satisfatoriamente com os dados observados no audiograma de Russo e Behlau, (1993).

Comparando os dados deste trabalho com os de Helfer e Huntley (1991) verifica-se que, em condição de ruído, jovens e idosos falantes do inglês com e sem perda auditiva tiveram mais dificuldade de discriminação dos seguintes fonemas em posição inicial da palavra: /p/, /b/, /t/, /d/, /k/, /g/ (classificados por eles como plosivos segundo o modo de articulação (MA) e /f/, /v/, /s/, /z/, /θ/, /ð/, /ʃ/, /h/ (fricativos segundo MA); /f/, /v/, /θ/, /ð/ (dentais segundo ponto de articulação PA) e /p/, /m/, /b/, /w/ (bilabiais segundo o PA). Os fonemas em posição inicial, que seriam difíceis de serem reconhecidos apenas por idosos (com e sem perda auditiva) são: /t/, /d/, /s/, /z/, /n/, /l/ (alveolares segundo PA) e /p/, /m/, /b/, /w/ (bilabiais segundo PA). Considerando os dados dos autores e os pares mínimos selecionados neste trabalho, são potencialmente difíceis de serem reconhecidos os seguintes pares mínimos: barro-jarro, bola-mola, fila-vila, pó-vó, couro-touro, tia-dia, mapa-tapa, ventar-sentar, fanta-janta, bote-pote, cinco-zinco, tela-zela, ponta-conta. Pares mínimos com discriminação de fonemas consonantais, possivelmente difíceis apenas para o grupo de idosos, são: data-nata, vinho-ninho e mão-não. Não haveria par de fácil discriminação, comparando os pares mínimos selecionados neste trabalho com os dados de Helfer e Huntley (1991). Os pares pão-chão e xiz-giz não foram analisados por que o fonema ʃ não foi analisado pelos autores.

De acordo com Silva (2003) os piores desempenhos para falantes do português brasileiro ocorreram nos fonemas: /r/, /l/, /ʎ/ (classificados por ela como líquido segundo MA) e /p/, /b/, /t/, /d/, /k/, /g/ (plosivo segundo MA); /f/, /v/ (labiodental segundo PA) e /R/, /k/, /g/ (velar segundo PA). Considerando os dados da autora, constata-se que os seguintes pares mínimos poderiam ser de difícil reconhecimento: bola-mola, fila-vila, pó-vó, couro-touro, tia-

dia, mapa-tapa, bote-pote e conta-ponta. Possivelmente seriam de grau médio de dificuldade os seguintes pares: barro-jarro, pão-chão, xiz-giz, data-nata, ventar-sentar, vinho-ninho, mão-não, fanta-janta e tela-zela. Tomando como referência os resultados relativos às análises acústicas e os dados da literatura, acredita-se que os pares mínimos da lista de estímulos têm vários níveis de dificuldade.

Encontra-se dificuldade para comparar os dados nas diversas pesquisas brasileiras porque o sistema internacional de classificação dos fonemas não é ainda universalmente adotado no Brasil. Além disto, há uma limitação natural na comparação dos dados linguísticos entre diferentes línguas, em razão das peculiaridades no conjunto de fonemas utilizados em cada língua. Por exemplo, os fonemas /θ/ e /ð/, integram a língua inglesa, mas não o português. Não obstante, a análise acústica aqui apresentada pode ser útil para a interpretação de dados de desempenho em procedimentos de treinamento auditivo para o falante do português brasileiro.

Diante das análises e estudos realizados sugere-se para futuras pesquisas que seja feita uma pequena avaliação cognitiva dos participantes, pois nota-se na literatura (Marshall, 1981; CHABA, 1989) que estes aspectos influenciam o desempenho em treinamento auditivo, principalmente em idosos. Além disso, sugere-se que a intensidade do sinal seja fixa e o mascarador variável, (conforme descrito por Cainer, James e Rajan, 2008) e que ambos os sons não passem de 90dB SPL (Elliott et. al, 1985) em pacientes sem perda auditiva, para evitar a exposição do participante a sons muito intensos. Indica-se que sejam usadas as três listas (vide Anexo B) separadamente com intervalos de descanso entre elas para evitar fadiga. Estimula-se o uso dos estímulos desenvolvidos neste trabalho para treinamentos auditivos em pacientes idosos com perdas auditivas leves, pois em casos de perdas auditivas graves outras intervenções parecem mais aconselháveis. Destaca-se a importância de analisar pelo menos o segundo formante nos próximos trabalhos, visto que o primeiro e o segundo formantes são

pistas acústicas muito importantes. É importante verificar com metodologias comportamentais a eficácia desses estímulos para procedimento de treinamento. Por fim, recomenda-se que procedimentos de treinamento incluam testes de generalização, para palavras novas e falantes que não participaram do treinamento, como parte da avaliação de sua eficácia.

Referências

- Aquino, A. M. C. M. A., Almeida, C. I. R. & Oliveira, J. A. A (1993). Teste de identificação de sentenças sintéticas (SSI) em português com mensagem competitiva: uma padronização. *Revista Brasileira de Otorrinolaringologia*, 59, 160-169.
- Baraldi, G. S., Almeida, L. C. & Borges, A. C. C. (2007). Evolução da perda auditiva no decorrer do envelhecimento. *Revista Brasileira de Otorrinolaringologia*, 73 (1), 64-70.
- Bess, F. H., Hedley-Williams, A. & Lichtenstein, M. J. (2001). Avaliação audiológica em idosos. Em Musiek, F.E, Rintellman, W. F (Orgs.), *Perspectivas atuais em avaliação auditiva*. Barueri, S.P.: Manole.
- Boersma, P. & Weenink, D. (2009): *Praat: doing phonetics by computer (version 5.1.35)* [Computer program]. Recuperado em: 10 de junho de 2010. Disponível em: <http://www.fon.hum.uva.nl/praat/>
- Borges, A.C.L.C. (1997). Dissílabos alternados: SSW. In: Pereira, L. D. & Schochat, E. Processamento auditivo central: manual de avaliação. São Paulo: Lovise.
- Burk, M. H. & Humes, L. E. (2008). Effects of long term training on aided speech-recognition performance in noise in older adults. *Journal of Speech, Language and Hearing Research*, 51, 759-771.
- Cainer, K. E., James, C. & Rajan, R. (2008). Learning speech-in-noise discrimination in adults humans. *Hearing Research*, 238, 155-164.
- Calais, L. L., Borges, A. C. L. C., Baraldi, G. S. & Almeida, L. C. A. (2008). Queixas e preocupações otológicas e as dificuldades de comunicação de indivíduos idosos. *Revista da Sociedade Brasileira de Fonoaudiologia*, 13 (1), 12-19.
- Carvalho, B. S. (2007). Teste de figuras para discriminação fonêmica: proposta e aplicação. Universidade Federal de Santa Maria, RGS. Disponível em: http://cascavel.cpd.ufsm.br/tede/tde_busca/arquivo.php?codArquivo=1383 Acessado em: 20 mar 2008.
- Corso, J. F. (1977). Presbycusis, hearing aids and aging. *Audiology*, 16 (2), 146-163.
- Costa, M. J. (1997). Desenvolvimento de listas de sentenças em português. (Tese de doutorado). Universidade Federal de São Paulo, SP. Disponível em: <http://search.bvsalud.org/unifesp/resources/epm-015144> Acessado: 15 mar 2010.

- Dubno, J. R., Dirks, D. D. & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America*, 76 (1), 87- 96.
- Elliott, L. L., Busse, L. A. & Bailet, L. L. (1985). Identification and discrimination of consonant-vowel syllables by younger and older adults. *Perception and Psychophysics*, 37, 307-314.
- Feitosa, M. A. G. (2001). Envelhecimento sensorial: pesquisa básica e implicações para a qualidade de vida. *Psychologica*, 28 (1), 159-175.
- Gordon-Salant, S. (1986). Effects of aging on response criteria in speech-recognition tasks. *Journal of Speech and Hearing Research*, 29, 155-162.
- Gordon-Salant, S. (2005). Hearing loss and aging: New research findings and clinical implications. *Journal of Rehabilitation Research and Development*, 42 (4), 9-24.
- Guimarães, S. R. K. (2004). O papel das pistas do contexto verbal no reconhecimento de palavras. *Psicologia em estudo*, 9 (2), 279-289.
- Helfer, K. S. & Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear and Hearing*, 29 (1), 87-98.
- Helfer, K. S. & Huntley, R. A. (1991). Aging and consonant errors in reverberation and noise. *The Journal of the Acoustical Society of America*, 90 (4), 1786-1796.
- Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C. & Lee, L. (1994). Factors associated with individual differences in clinical measures of speech recognition among the elderly. *Journal of Speech and Hearing Research*, 37, 465-474.
- Instituto Brasileiro de Geografia e Estatística (2002). Perfil dos idosos responsáveis por domicílios no Brasil (Estudos e Pesquisas Informação Demográfica e Socioeconômica Número 9). Rio de Janeiro, RJ: Autor.
- Jerger, J., Jerger, S., Oliver, T. & Pirozzolo, F. (1989). Speech understanding in the elderly. *Ear and Hearing*, 2, 79-89.
- Jorge, V.O. (2008). Efeitos do treinamento auditivo em idosos com envelhecimento normal. (Monografia de conclusão de curso). Universidade Federal de Minas Gerais, MG. Disponível em: <http://www.medicina.ufmg.br/fon/monografias/08-VivianJorge.pdf> Acessado: 15 mar 2010.
- Kozlowski, L., Wiemes, G. M. R, Magni, C. & Silva, A. L. G. (2004). A efetividade do treinamento auditivo na desordem do processamento auditivo central: estudo de caso. *Revista Brasileira de Otorrinolaringologia*, 70 (3), 427-432.

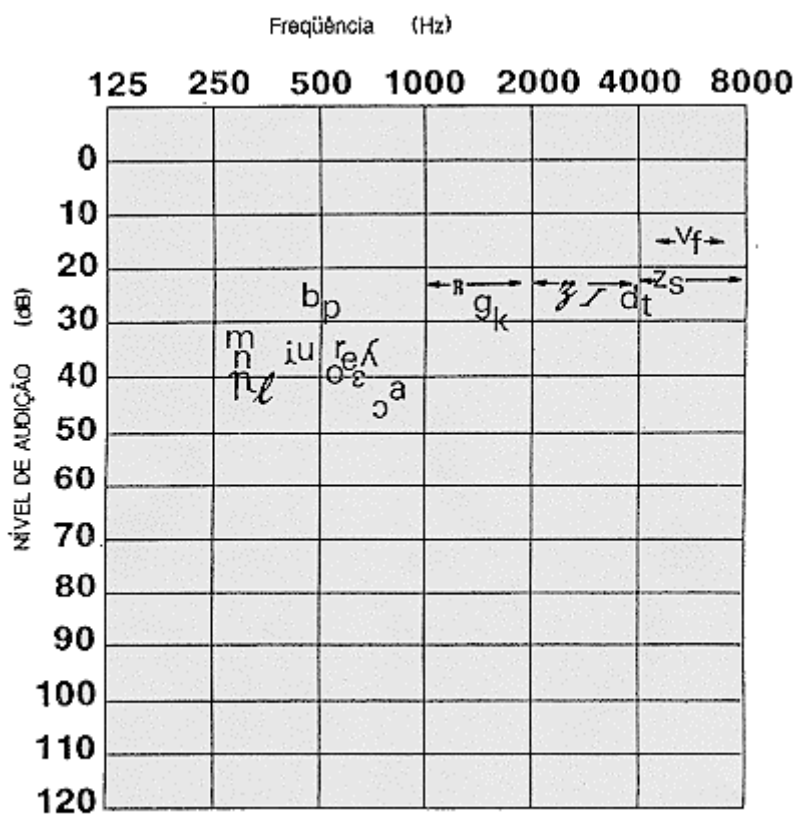
- Lewkovicz, A. A. (2008). Reconhecimento de fala no silêncio e no ruído com fones em indivíduos com perda auditiva neurossensorial de diferentes configurações. (Dissertação de mestrado). Universidade Federal de Santa Maria, RGS. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=112145 Acessado em: 15 mar 2010.
- Lispector, C. (1998). *Felicidade Clandestina - contos*. 1ª ed., Rio de Janeiro: Editora Rocco.
- Lúcio, P. S., Pinheiro, A. M. V. & Nascimento, E. (2009). O impacto da mudança no critério de acerto na distribuição dos escores do subteste de leitura do teste de desempenho escolar. *Psicologia em estudo*, 14 (3), 593-601.
- Marshall, L. (1981). Auditory processing in aging listeners. *Journal of Speech and Hearing Disorders*, 46 (3), 226- 240.
- McCroskey, R. L. & Kasten, R. N. (1982). Temporal factors and the aging auditory system. *Ear and Hearing*, 3, 124-127.
- Ministério da Saúde (2002). *Portaria n. 1060/GM, de 5 de junho de 2002*. Aprova a Política Nacional de Saúde da Pessoa Portadora de Deficiência. Recuperado em 20 de outubro de 2008, dtr2001.saude.gov.br/sas/PORTARIAS/Port2002/Gm/GM-1060.htm
- Mota, H. B. (2001). Pares mínimos: os contrastes fonológicos do português brasileiro. *Pró-Fono Revista de Atualização Científica*, 13 (1), 98-106.
- Mota, H. B.; Keske-Soares, M. & Costamilan, C. M. (2001). Discriminação auditiva em crianças com desvios fonológicos evolutivos. *Anais do IX Congresso Brasileiro de Fonoaudiologia*. Guarapari, SBFa.
- Neves, V. T. & Feitosa, M. A. G. (2003). Controvérsias ou complexidade na relação entre processamento temporal e envelhecimento? *Revista Brasileira de Otorrinolaringologia*, 69 (2), 242-249.
- Padilha, C. B. (2008). Reconhecimento de sentenças no silêncio e no ruído, em campo livre, em indivíduos portadores de perda auditiva neurossensorial de grau moderado. (Dissertação de mestrado). Universidade Federal de Santa Maria, RGS. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=111682. Acessado em: 15 mar 2010.
- Plomp, R. (1993). Comments on “Evaluating a speech-reception threshold model for hearing-impaired listeners”. *The Journal of the Acoustical Society of America*, 93 (1), 2879-2885.

- Plomp, R. & Duquesnoy, A. J. (1980). Room Acoustics for the aged. *The Journal of the Acoustical Society of America*, 68, 1616-1621.
- Raimundo, J. C. (2008). Desempenho em tarefas de percepção de fala de crianças com deficiência auditiva: familiaridade da lista de palavras. (Dissertação de mestrado). Pontifícia Universidade Católica de São Paulo, SP. Disponível em: http://www.sapientia.pucsp.br/tde_busca/arquivo.php?codArquivo=7497 Acessado em: 15 mar 2010.
- Rajan, R., James, C. & Cainer, K. E. (2008). Aging without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. *Neuroscience*, 154, 784-795.
- Robinson, K. & Summerfield, A. Q. (1996). Adult auditory learning and training. *Ear and Hearing*, 17 (3), 51S-65S.
- Russo, I. & Behlau, M. (1993). *Percepção da Fala: Análise Acústica*. São Paulo, SP: Editora Lovise.
- Schechan, K. A., McArthur, G. M., Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research*, 25, 547-553.
- Schochat, E. & Pereira, L. D. (1997). *Processamento auditivo central: manual de avaliação*. São Paulo: Lovise.
- Schröder, P. L. L. (2008). A relação entre o teste SSW e a latência da onda P3. (Dissertação de mestrado). Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=93460 Acessado: 15 mar 2010.
- Schuknecht, H. F. (1955). Presbycusis. *Laryngoscope*, 65, 402-410.
- Silva, I. M. C. (2003). Sensibilidade a tons de alta frequência e reconhecimento de fala em adultos jovens e mais velhos. Dissertação de mestrado em Ciências do Comportamento, Instituto de Psicologia da Universidade de Brasília, Brasília. DF, Brasil.
- Souza, R. R. (2010). Treinamento auditivo em idosos com comprometimento cognitivo leve. (Dissertação de mestrado). Universidade de São Paulo, SP. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=169453 Acessado em 15 mar 2010.

- Sweetow, R. W. & Palmer, C. V. (2005). Efficacy of individual auditory training in adults: a systematic review of the evidence. *Journal of American Academy of Audiology*, 16 (7), 494-504.
- Syka, J. (2002). Plastic changes in the central auditory system after hearing loss, restoration of function, and during learning. *Physiological Reviews*, 82, 601-636.
- Tremblay, K. L. (2005). Beyond the ear: physiological perspectives on auditory rehabilitation. *Seminars in Hearing*, 26 (3), 127-136.
- Tremblay, K.; Kraus, N.; McGee, T. (1998). The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. *Neuroreport*, 9 (16), 3557-3560.
- Willot, J. F. (1991). *Aging and the auditory system: anatomy, physiology, and psychophysics*. San Diego: Singular.
- Working Group on Speech Understanding and Aging (CHABA, 1988). Speech understanding and aging. *The Journal of the Acoustical Society of America*, 83 (3), 859-895.

Anexos

Anexo A - Audiograma de fonemas de Russo e Behlau (1993)



Valores acústicos médios de frequência e intensidade dos sons da fala do português brasileiro, dispostos no registro gráfico do audiograma (Russo & Behlau, 1993). Acompanhando o alfabeto fonético internacional adotado nesta dissertação é importante observar que o fonema /ʎ/ equivale ao fonema /j/ e o fonema /ʒ/ equivale ao fonema /ʒ/.

Anexo B - Listas de pares mínimos para familiarização e treinamento auditivo

Lista A	bolo-rolô	disco-risco	louco-louro	osso-olho
	capa-cara	febre-lebre	maca-malha	puxa-puma

Lista B	brinco-trinco	soda-soja	colher-mulher	pão-chão	tela-zela
	rabo-raso	dedo-medo	mão-não	carinho-carrinho	tapa-mapa
	bola-mola	dama-cama	fina-filha	cinco-zinco	ventar-sentar
	bicho-lixo	fila-vila	caneca-careca	casal-canal	vinho-ninho
	bato-gato	fanta-janta	punho-pulo	lisa-lixá	loja-lona
	barro-jarro	gruta-fruta	feio cheio	xiz-giz	gatilho-gatinho
	data-nata	gato-jato	ponta-conta	cacho-carro	colhe-corre
	cadeira-caveira	caixa-faixa	pó-vó	tia-dia	touro-couro

Lista C	bala-fala	face-vaca	pala-palha	pêra-cera	teia-veia
	bota-nota	forro-zorro	bolacha-borracha	passa-massa	fita-ficha
	bola-cola	filho-milho	fila-filha	careta-carreta	carteiro-carneiro
	pote-bote	galinha-salinha	uma-unha	ceia-cheia	cavar-casar
	moda-molha	gorro-morro	sono-sonho	camisa-Camila	vaca-jaca
	roda-rosa	calo-galo	sonho-soro	casa-cama	bandeja-bandeira
	marido-marinho	carro-varro	porta-torta	chato-jato	pastilha-pastinha
	dado-gado	cão-não	posto-gosto	chato-rato	telha-terra

Lista D	bela-vela	fio-rio	cola-gola	pavio-navio	tira-gira
	carimbo-carinho	pegada-pesada	mala-malha	caro-carro	trio-frio
	bóia-jóia	figo-fino	moça-louça	selo-gelo	veia-cheia
	dente-lente	gata-lata	nua-lua	passo-paro	aveia-areia
	bolada-bolacha	roqueira-roseira	canhão-carão	lixo-linho	zangada-jangada
	vida-vira	canto-santo	pen-te-dente	chico-rico	alho-aro
	figura-segura	cama-chama	ponte-fonte	Terra-guerra	molho-morro
	face-jaca	secar-serrar	pé-Zé	tinta-cinta	vela-Vera
	timão-limão				

*Anexo C - Amostra do texto original de Clarisse Lispector e do texto sem sentido
correspondente*

Redação original de parágrafo representativo

Não me fantasiavam: no meio das preocupações com minha mãe doente, ninguém em casa tinha cabeça para carnaval de criança. Mas eu pedia a uma de minhas irmãs para enrolar aqueles meus cabelos lisos que me causavam tanto desgosto e tinha então a vaidade de possuir cabelos frisados pelo menos durante três dias por ano. Nesses três dias, ainda, minha irmã acedia ao meu sonho intenso de ser uma moça – eu mal podia esperar pela saída de uma infância vulnerável – e pintava minha boca de batom bem forte, passando também ruge nas minhas faces. Então eu me sentia bonita e feminina, eu escapava da meninice.

Redação do mesmo parágrafo com alterações realizadas

Não me causavam: no meio das conjecturas com minha mãe, ninguém em casa ficava cabeça para carnaval de criança. Mas eu enrolava a uma de minhas irmãs para podia aqueles meus cabelos lisos que me pintavam tanto gosto e explicava então a vaidade de ganhar cabelos frisados pelo menos durante três dias por ano. Nesses três dias, ainda, minha irmã vinha ao meu sonho intenso de ser uma moça - eu mal economizava constatar pela saída de uma infância exposta - e durava minha boca de batom bem forte, enrolando também ruge nas minhas faces. Então eu me criava bonita e feminina, eu restava da meninice.

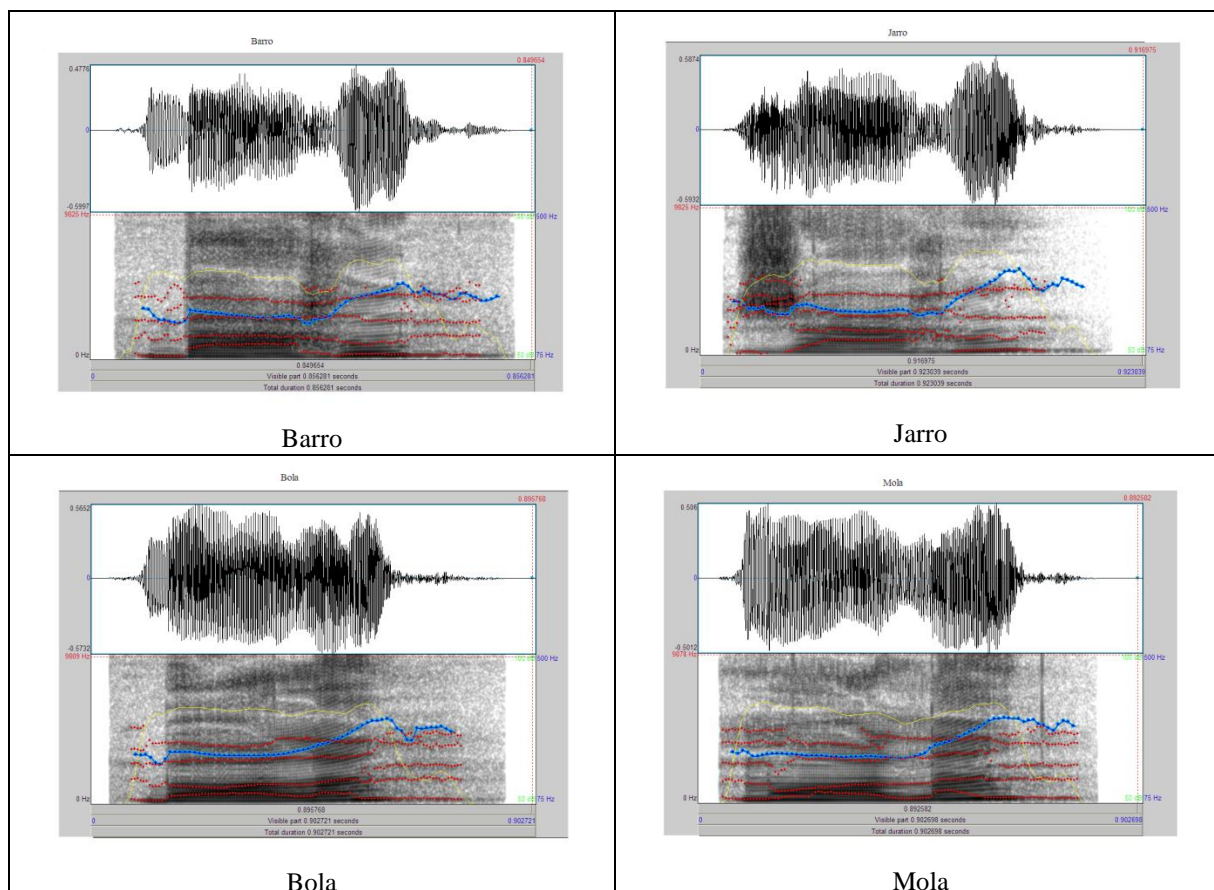


Figura 3 – Espectrogramas dos pares mínimos: Barro-Jarro e Mola-Bola

Tabela 3 – Características acústicas dos pares mínimos Barro-Jarro e Mola-Bola

Par	barro-jarro	barro-jarro	bola-mola	bola-mola
Palavra analisada	barro	Jarro	bola	mola
Duração em seg. (FD)	0,143	0,165	0,120	0,123
Modulação de frequência do 1º formante (FD)	Sim. Ascendente.	Sim. Ascendente.	Sim. Ascendente.	Sim. Descendente.
Taxa de modulação	Alta.	Alta.	Baixa.	Baixa.
Localização da modulação no 1º formante (FD)	Final.	Início e final.	Final	Início.
Número de formantes com bastante energia (FD)	1	1	1	3
Intensidade mínima em dB (FD)	50	43	46	46
Intensidade máxima em dB (FD)	80	80	80	84
Pitch mínimo em Hz (FD)	178	187	±186	±207
Pitch máximo em Hz (FD)	221	226	222	±223
Área de frequência com > energia em Hz (FD)	≤2.000	≥2.500	≤1.000	≤1.800
Área de frequência com maior energia em Hz(FD)(audiograma de Russo & Behlau)	<1.000	2.000 a 4.000	<1.000	1.000 a 2.000

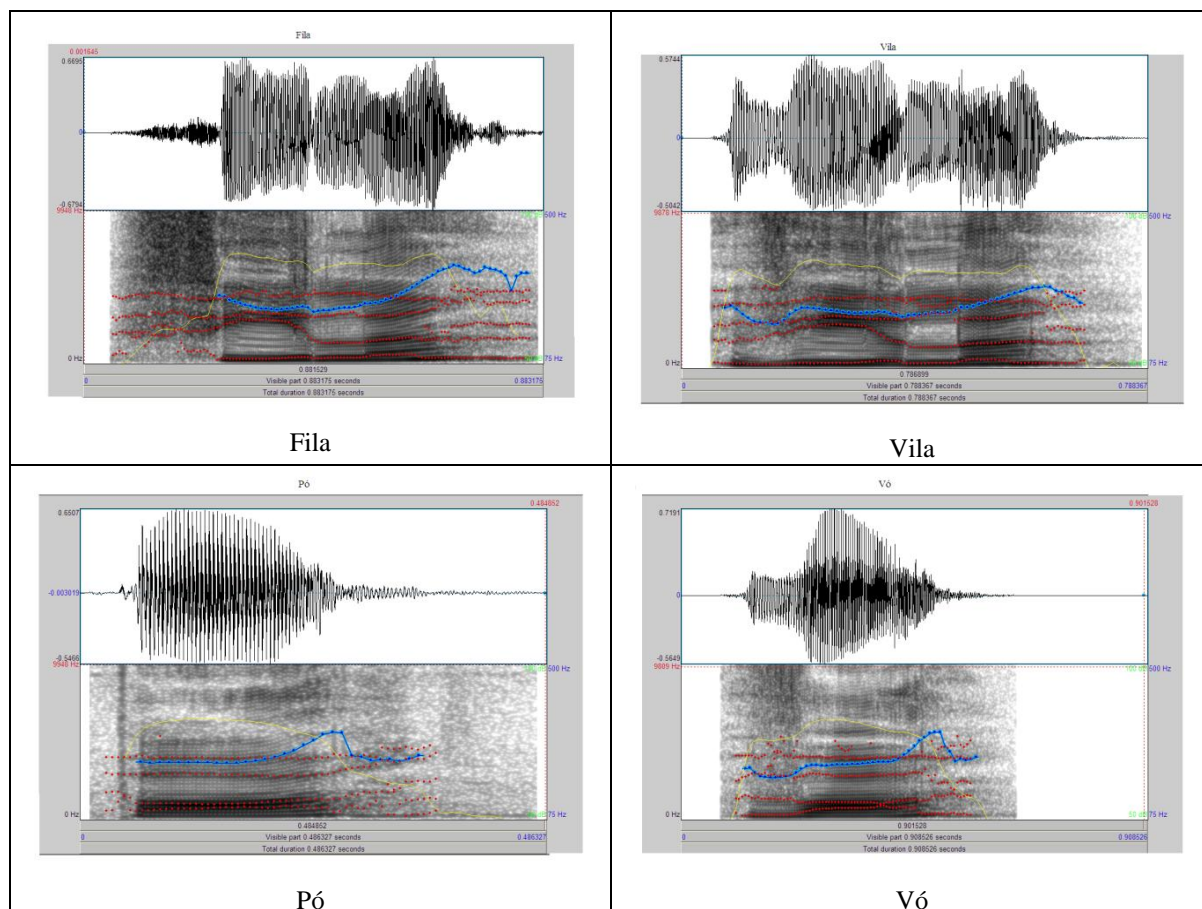


Figura 4 – Espectrogramas dos pares mínimos: Fila-Vila e Pó-Vó

Tabela 4 – Características acústicas dos pares mínimos: Fila-Vila e Pó-Vó

Par	fila-vila	fila-vila	pó-vó	pó-vó
Palavra analisada	fila	vila	pó	vó
Duração em seg. (FD)	0,219	0,125	0,065	0,174
Modulação de frequência 1º formante (FD)	Sim.	Sim. Descendente.	Sim. Não específica.	Sim.
Taxa de modulação	Alta.	Baixa.	Alta.	Baixa.
Localização da modulação no 1º formante (FD)	Em todo fonema.	Início.	Em todo fonema.	Início e Final.
Número de formantes com bastante energia (FD)	Nenhum. Espectro amplo.	1	Nenhum.	2
Intensidade mínima em dB (FD)	44	55	69	42
Intensidade máxima em dB (FD)	79	81	76	79
Pitch mínimo em Hz (FD)	216	193	231	189
Pitch máximo em Hz (FD)	347	295	231	218
Área de frequência com > energia em Hz (FD)	5.000 a 17.000	6.000 a 15.000	≤1.500	6.000 a 14.000
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	4.000 a 8.000	4.000 a 8.000	<1.000	4.000 a 8.000

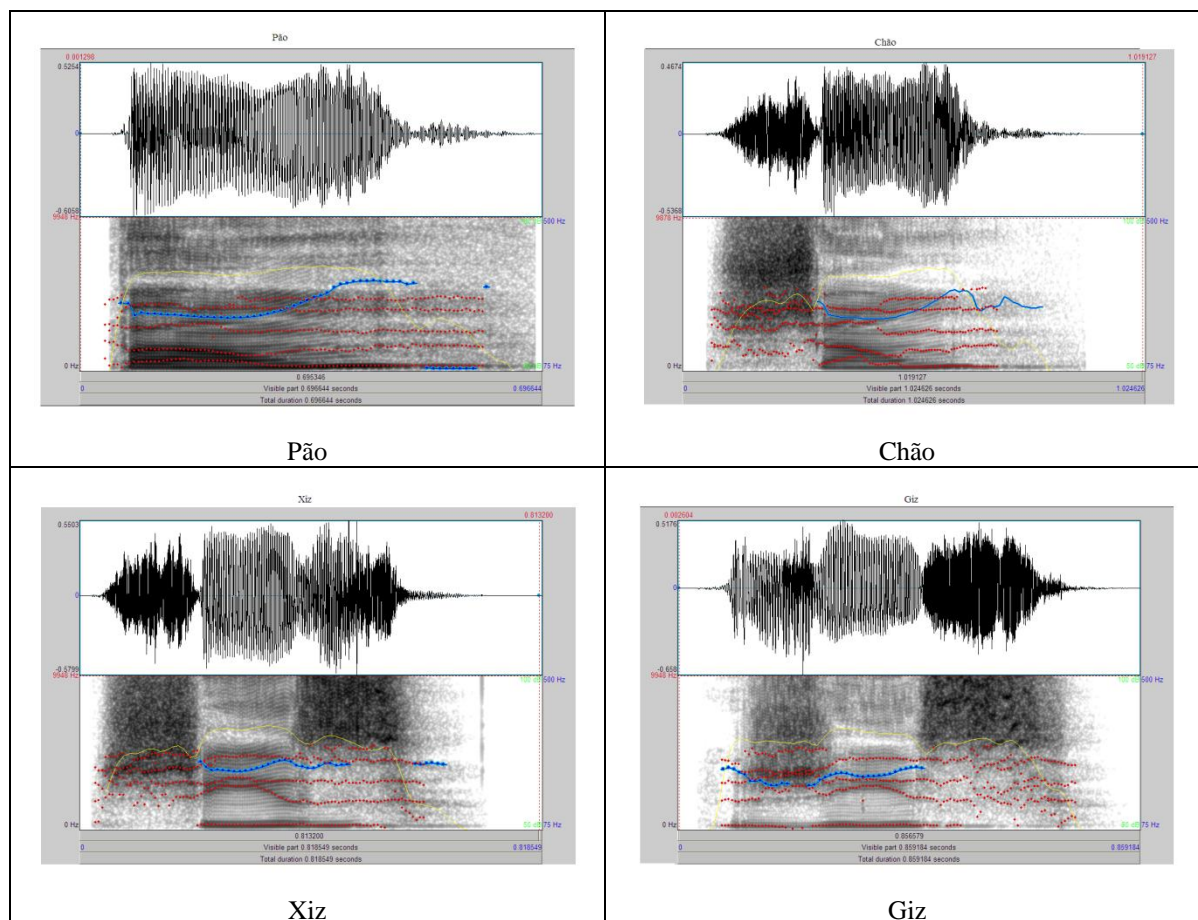


Figura 5 – Espectrogramas dos pares mínimos: Pão-Chão e Xiz-Giz

Tabela 5 – Características acústicas dos pares mínimos: Pão-Chão e Xiz-Giz

Par	pão-chão	pão-chão	xiz-giz	xiz-giz
Palavra analisada	pão	chão	xiz	giz
Duração em seg. (FD)	0,039	0,309	0,206	0,243
Modulação de frequência do 1º formante (FD)	Sim. Múltipla.	Sim. Não específica.	Sim. Não específica.	Não.
Taxa de modulação	Alta.	Alta.	Alta.	Nenhuma.
Localização da modulação no 1º formante (FD)	Início.	Todo fonema.	Em todo fonema.	-----
Número de formantes com bastante energia (FD)	4	Nenhum. Espectro amplo.	Nenhum. Espectro amplo.	1
Intensidade mínima em dB (FD)	49	40	62	42
Intensidade máxima em dB (FD)	82	79	79	79
Pitch mínimo em Hz (FD)	226	258	245	196
Pitch máximo em Hz (FD)	262	266	262	245
Área de frequência > energia em Hz (FD)	≤5.000	2.500 a 15.000	3.000 a 14.000	3.000 a 12.000
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	<1.000	2.000 a 4.000	2.000 a 4.000	2.000 a 4.000

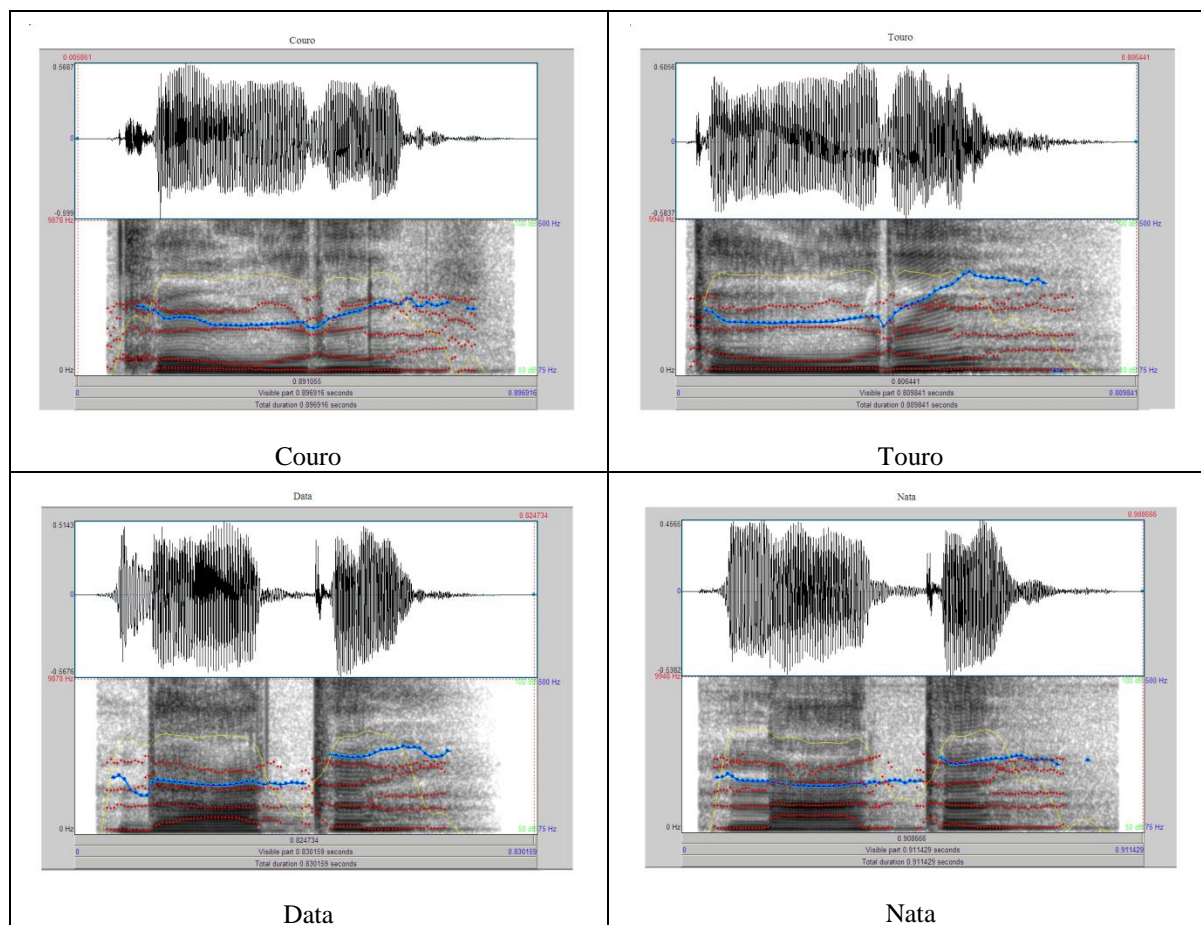


Figura 6 – Espectrogramas dos pares mínimos: Couro-Touro e Data-Nata

Tabela 6 – Características acústicas dos pares mínimos: Couro-Touro e Data-Nata

Par	couro- touro	couro-touro	data-nata	data-nata
Palavra analisada	couro	touro	data	nata
Duração em seg. (FD)	0,102	0,050	0,105	0,154
Modulação de frequência do 1º formante (FD)	Não identificável.	Sim.	Sim.	Sim.
Taxa de modulação	-----	Alta.	Baixa.	Alta.
Localização da modulação no 1º formante (FD)	-----	Início.	Início e final.	Início e final.
Número de formantes com bastante energia	1	2	1	3
Intensidade mínima em dB (FD)	46	72	49	49
Intensidade máxima em dB (FD)	79	82	81	83
Pitch mínimo em Hz (FD)	243	240	178	208
Pitch máximo em Hz (FD)	261	249	234	233
Área de frequência > energia em Hz (FD)	6.000 a 15.000	1.500 a 15.000	≤1.000	≤3.400
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	1.000 a 2.000	4.000 a 8.000	2.000 a 4.000	2.000 a 4.000

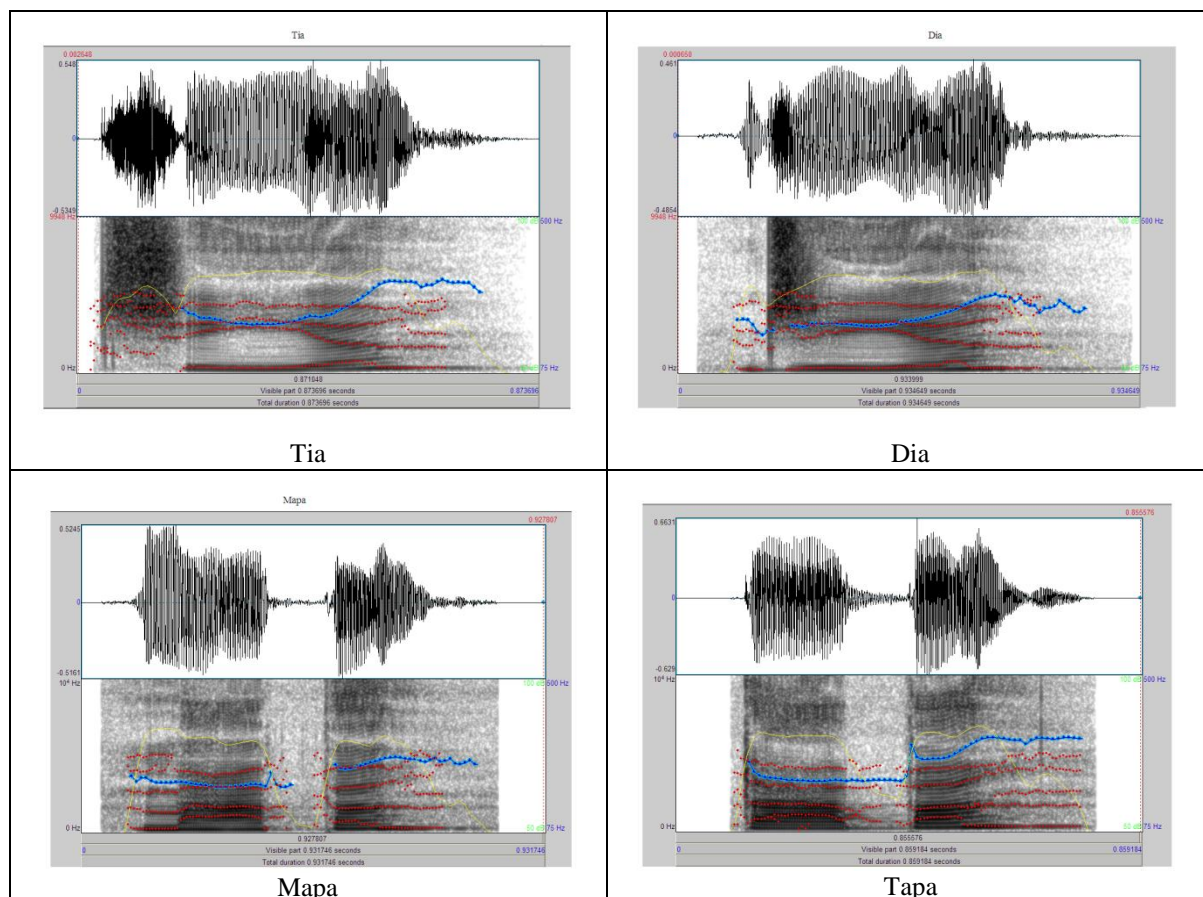


Figura 7 – Espectrogramas dos pares mínimos: Tia-Dia e Mapa-Tapa

Tabela 7 – Características acústicas dos pares mínimos: Tia-Dia e Mapa-Tapa

Par	tia-dia	tia-dia	mapa-tapa	mapa-tapa
Palavra analisada	tia	dia	mapa	tapa
Duração (FD)	0,168	0,229	0,160	0,053
Modulação de frequência do 1º formante (FD)	Sim. Não específica.	Sim. Descendente.	Sim. Múltipla.	Sim. Múltipla.
Taxa de modulação	Alta.	Baixa.	Alta.	Alta.
Localização da modulação no 1º formante (FD)	Em todo fonema.	Início.	Início e meio.	Em todo fonema.
Número de formantes com bastante energia (FD)	Nenhum. Espectro amplo.	1	3	Nenhum. Espectro amplo.
Intensidade mínima em dB (FD)	64	47	44	45
Intensidade máxima em dB (FD)	78	78	84	80
Pitch mínimo em Hz (FD)	250	177	212	236
Pitch máximo em Hz (FD)	250	220	230	262
Área de frequência com maior energia em Hz (FD)	2.000 a 18.000	2.800 a 14.000	≤3.000	2.000 a 17.000
Área de frequência > energia em Hz (FD) (audiograma de Russo & Behlau)	4.000 a 8.000	2.000 a 4.000	<1.000	4.000 a 8.000

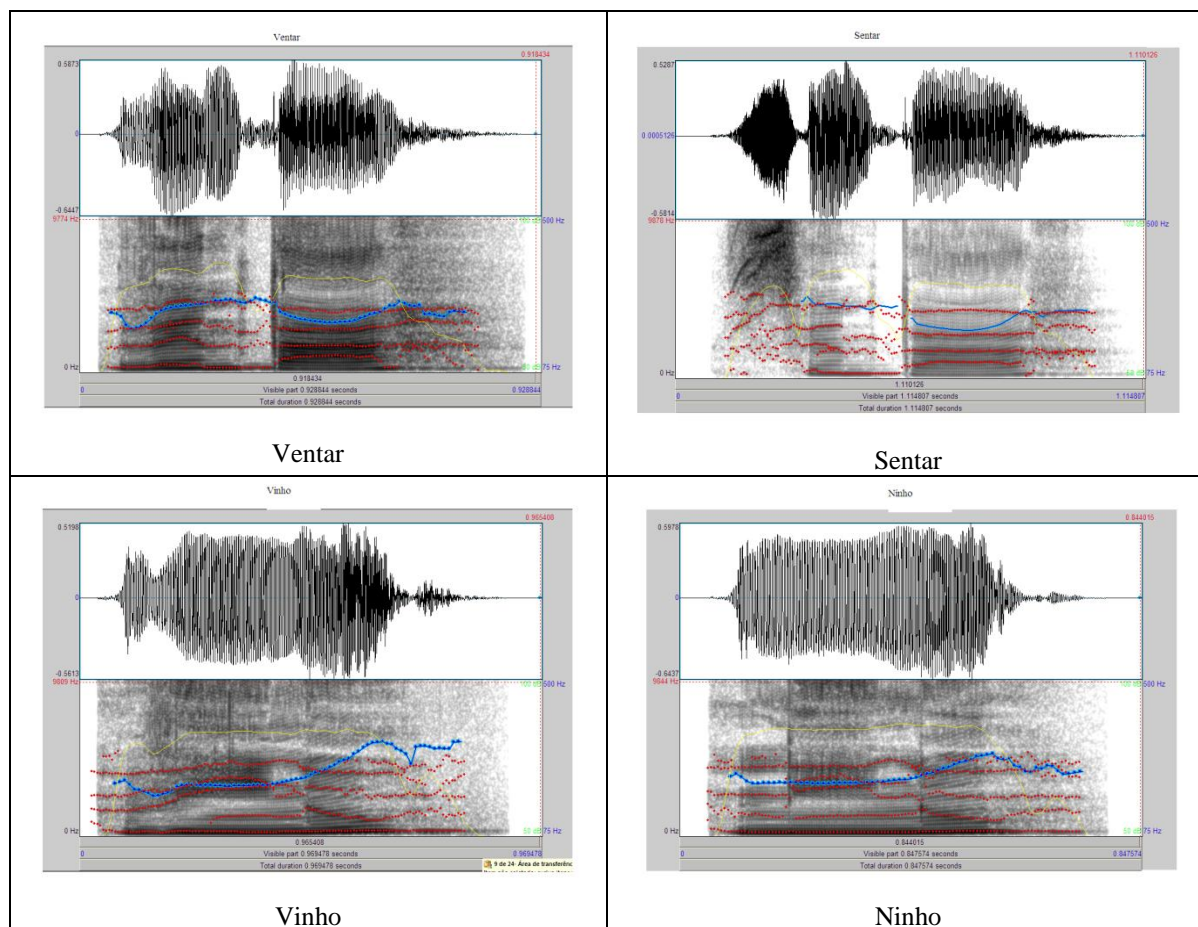


Figura 8 – Espectrogramas dos pares mínimos: Ventar-Sentar e Vinho-Ninho

Tabela 8 – Características acústicas dos pares mínimos: Ventar-Sentar e Vinho-Ninho

Par	ventar-sentar	ventar-sentar	vinho-ninho	vinho-ninho
Palavra analisada	ventar	sentar	vinho	ninho
Duração em seg.	0,105	0,233	0,136	0,127
Modulação de frequência do 1º formante (FD)	Não	Sim. Múltipla.	Sim. Descendente.	Sim. Descendente.
Taxa de modulação	-----	Alta.	Alta.	Baixo.
Localização da modulação no 1º formante (FD)	-----	Início, meio e final.	Início e meio.	Início.
Número de formantes com bastante energia (FD)	2	Nenhum. Espectro amplo.	3	3
Intensidade mínima em dB (FD)	48	38	47	40
Intensidade máxima em dB (FD)	79	79	80	84
Pitch mínimo em Hz (FD)	192	280	183	216
Pitch máximo em Hz (FD)	237	288	226	244
Área de frequência com > energia em Hz (FD)	≤13.000	5.000 a 15.000	5.000 a 16.000	≤5.500
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	4.000 a 8.000	4.000 a 8.000	4.000 a 8.000	<1.000

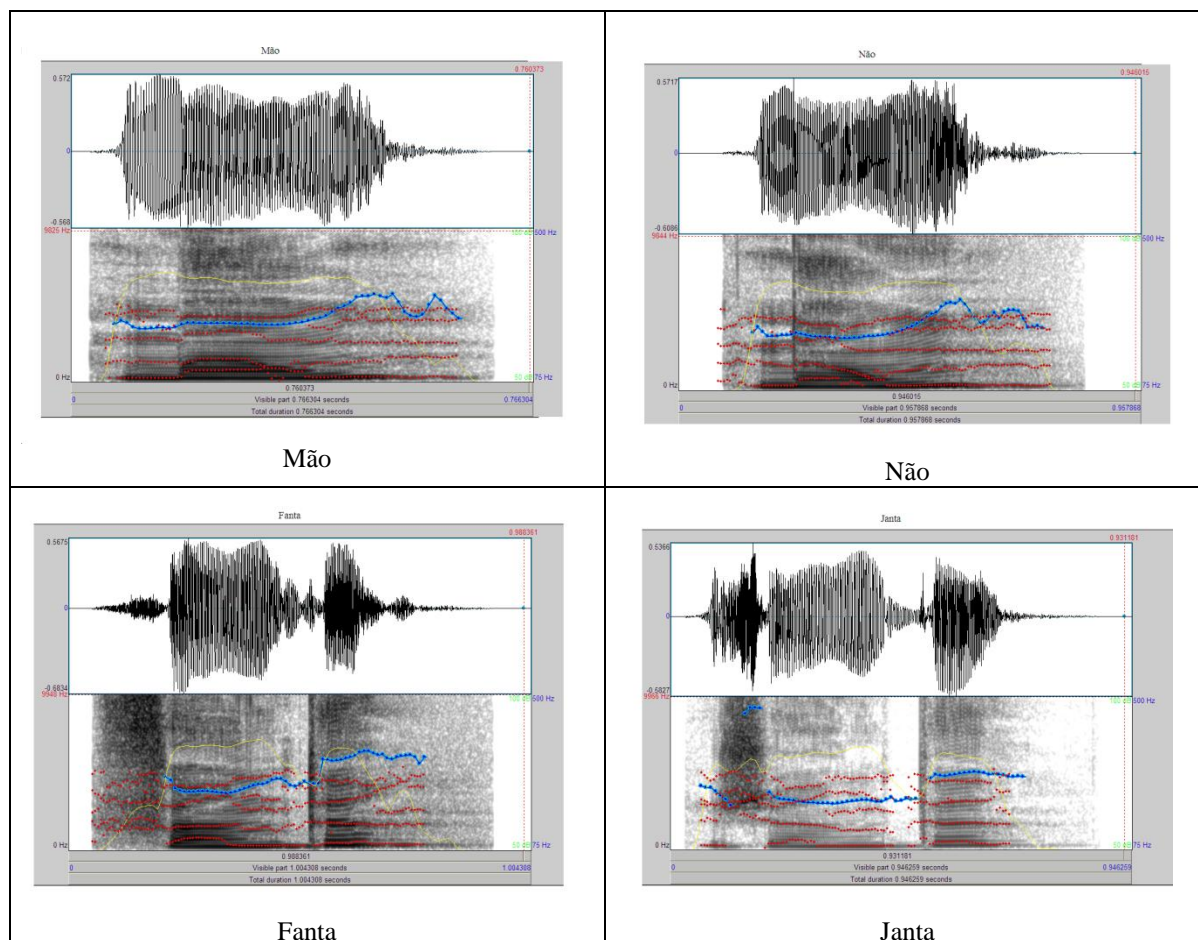


Figura 9 – Espectrogramas dos pares mínimos: Mão-Não e Fanta-Janta

Tabela 9 – Características acústicas dos pares mínimos: Mão-Não e Fanta-Janta

Par	mão-não	mão-não	fanta-janta	fanta-janta
Palavra analisada	mão	não	fanta	janta
Duração em seg. (FD)	0,146	0,144	0,182	0,171
Modulação de frequência do 1º formante (FD)	Não.	Sim. Múltipla.	Sim. Múltipla.	Sim. Múltipla.
Taxa de modulação	-----	Alta.	Alta.	Média.
Localização da modulação no 1º formante (FD)	-----	Início.	Em todo fonema.	Em todo fonema.
Número de formantes com bastante energia (FD)	3	3	Nenhum. Espectro amplo.	Nenhum. Espectro amplo.
Intensidade mínima em dB (FD)	48	42	39	48
Intensidade máxima em dB (FD)	85	85	76	80
Pitch mínimo em Hz (FD)	221	221	266	197
Pitch máximo em Hz (FD)	244	246	273	469
Área de frequência com > energia em Hz (FD)	≤3.100	≤3.100	1.000 a 20.000	2.500 a 15.000
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	<1000	<1000	4.000 a 8.000	2.000 a 4.000

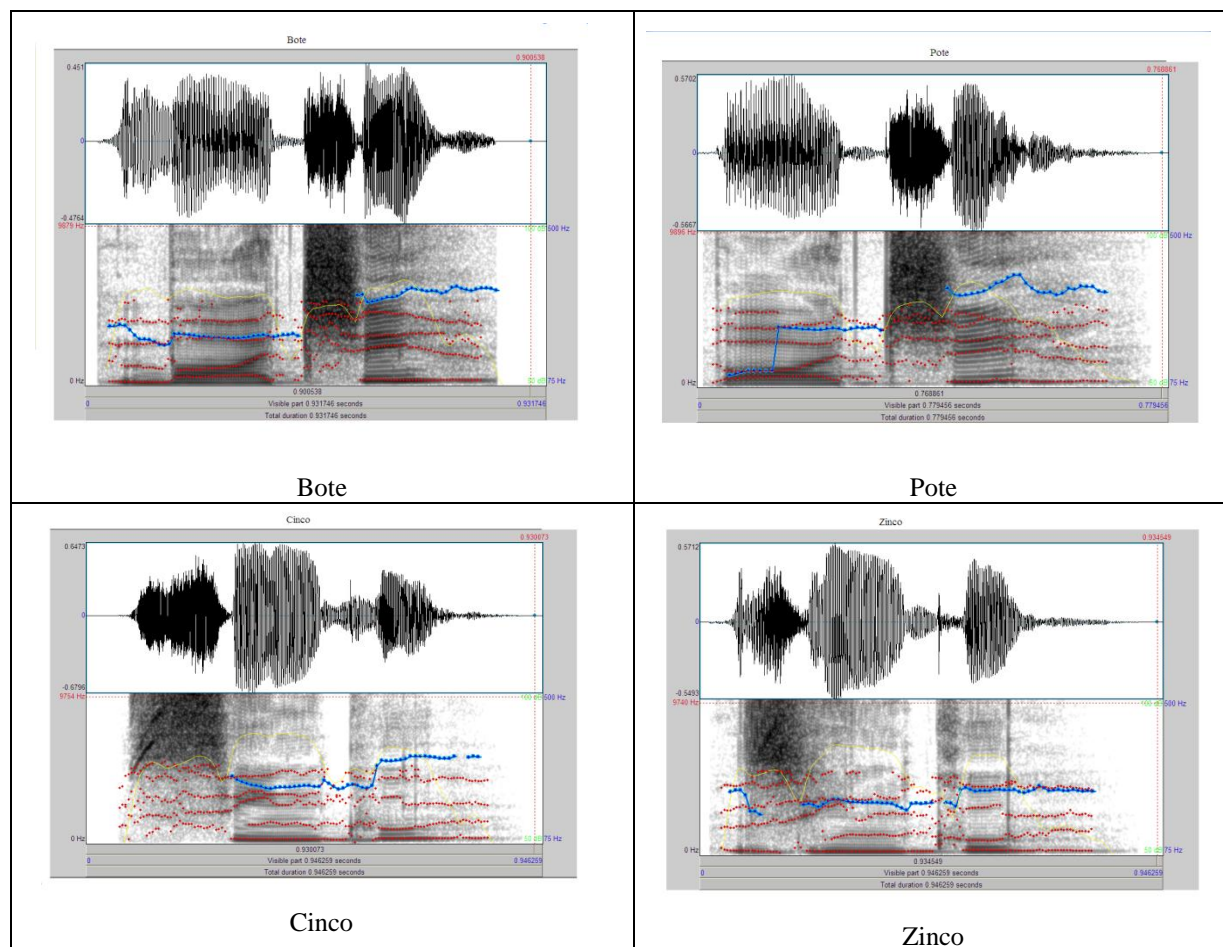


Figura 10 – Espectrogramas dos pares mínimos: Bote-Pote e Cinco-Zinco

Tabela 10 – Características acústicas dos pares mínimos: Bote-Pote e Cinco-Zinco

Par	bote-pote	bote-pote	cinco-zinco	cinco-zinco
Palavra analisada	bote	pote	cinco	zinco
Duração em seg. (FD)	0,160	0,050	0,237	0,199
Modulação de frequência do 1º formante (FD)	Não.	Sim. Ascendente	Sim. Múltipla.	Sim. Descendente.
Taxa de modulação	-----	Média.	Alta.	Baixa.
Localização da modulação no 1º formante (FD)	-----	Final.	Em todo fonema.	Em todo fonema.
Número de formantes com bastante energia (FD)	Nenhum.	1	Nenhum. Espectro amplo.	Nenhum. Espectro amplo.
Intensidade mínima em dB (FD)	55	75	41	52
Intensidade máxima em dB (FD)	80	78	79	78
Pitch mínimo em Hz (FD)	181	107	Não definido.	185
Pitch máximo em Hz (FD)	235	107	Não definido.	253
Área de frequência com > energia em Hz (FD)	<1.000	<2.000	5.000 a 17.000	4.000 a 16.000
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	<1.000	<1.000	4.000 a 8.000	4.000 a 8.000

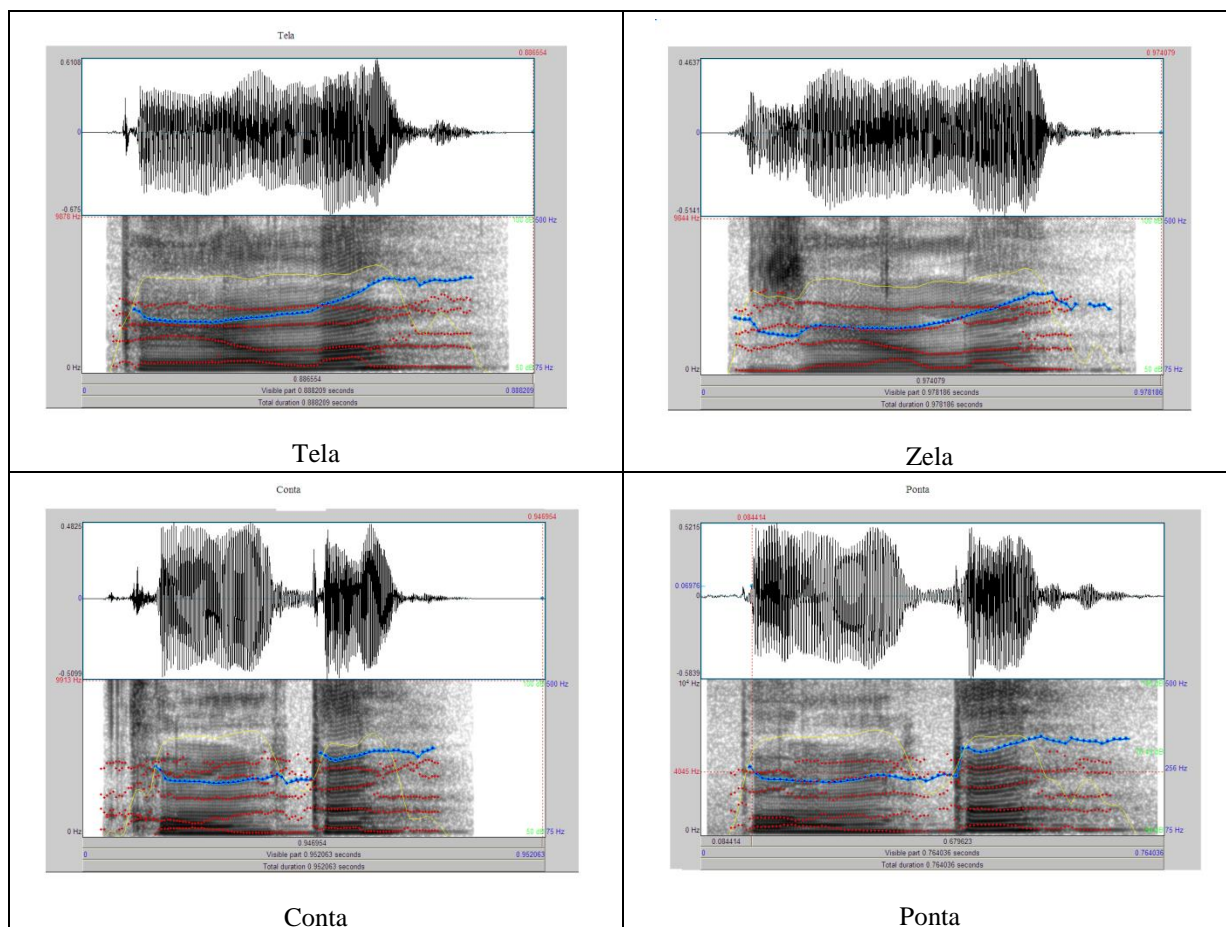


Figura 11 – Espectrogramas dos pares mínimos: Tela-Zela e Conta-Ponta

Tabela 11 – Características acústicas dos pares mínimos: Tela-Zela e Conta-Ponta

Par	tela-zela	tela-zela	conta-ponta	conta-ponta
Palavra analisada	tela	zela	conta	ponta
Duração em seg. (FD)	0,066	0,169	0,118	0,098
Modulação de frequência (FD)	Sim. Múltipla.	Sim. Ascendente.	Sim. Múltipla.	Sim.
Taxa de modulação	Alta.	Baixa.	Alta.	Alta.
Localização da modulação no 1º formante (FD)	Todo fonema.	Final.	Em todo fonema.	Final.
Número de formantes com bastante energia (FD)	Nenhum.	1	4	4
Intensidade mínima em dB (FD)	47	49	48	49
Intensidade máxima em dB (FD)	76	76	74	78
Pitch mínimo em Hz (FD)	240	177	257	233
Pitch máximo em Hz (FD)	245	227	265	261
Área de frequência com > energia em Hz (FD)	4.000 a 14.000	4.500 a 14.000	≥2.000	≤5.000
Área de frequência com maior energia em Hz (FD) (audiograma de Russo & Behlau)	4.000 a 8.000	4.000 a 8.000	1.000 a 2.000	<1.000