

Daniel Márcio Rodrigues Silva

A PERCEPÇÃO DAS VOGAIS MÉDIAS
POSTERIORES NO PORTUGUÊS BRASILEIRO

Belo Horizonte
Faculdade de Letras da UFMG
2006

Daniel Márcio Rodrigues Silva

A PERCEPÇÃO DAS VOGAIS MÉDIAS POSTERIORES NO PORTUGUÊS BRASILEIRO

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Lingüísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Estudos Lingüísticos.

Área de concentração: Lingüística

Orientador: Prof. Rui Rothe-Neves
Universidade Federal de Minas Gerais

Belo Horizonte
Faculdade de Letras da UFMG
2006

AGRADECIMENTOS

Aos sujeitos da pesquisa, pela colaboração.

Às eternamente amadas Lucia F. Moreira e Silvia S. Pereira.

Ao Prof. Dr. Rui Rothe-Neves pela orientação.

Aos amigos Caio Lourenço, Victor Hugo Munhoz, Marcos Lourenço, Nazareth Leite, Conrado Salazar, Patrícia Lúcio e Camila Heleno, pelo apoio.

Ao CNPq pela concessão da bolsa de mestrado.

É um erro capital teorizar antes de ter os dados. Insensivelmente, começa-se a distorcer os fatos para adaptá-los às teorias, em vez de fazer as teorias se adaptarem aos fatos.

Sherlock Holmes

SUMARIO

RESUMO	6
ABSTRACT	7
LISTA DE FIGURAS	8
LISTA DE TABELAS	9
1 Introdução	10
2 Revisão Bibliográfica	13
2.2. A Categorização dos Sons da Fala	16
2.2.1. Percepção categórica	16
2.2.2 Evidências psicofisiológicas	32
2.3 Esclarecimentos adicionais	43
3 Metodologia	52
3.1 Sujeitos	52
3.3 Delineamento experimental	53
3.3.1 Tarefa de Classificação	53
3.3.2 Tarefa de discriminação 2AFC (escolha forçada com dois intervalos e duas alternativas)	54
3.3.3 Tarefa de discriminação 4I2AFC (escolha forçada com quatro intervalos e duas alternativas)	56
3.4 Procedimentos	57
3.5 Plano de Análise	60
4 Resultados	63
5 Discussão	78
6 Conclusão	85
Referências Bibliográficas	86
ANEXO A - Proporções das respostas “u”, “o” e “ɔ” na tarefa de classificação	93
ANEXO B - Tabelas de d'	94
Média	95
ANEXO C – <i>Scripts</i> criados para a síntese das vogais	96

RESUMO

No português brasileiro são atestados pares mínimos estabelecendo o contraste entre vogais médias abertas e vogais médias fechadas na sílaba tônica. Entretanto, há casos em que estas duas categorias variam neste mesmo contexto sem conseqüências semânticas. O objetivo deste trabalho foi verificar se este fenômeno se reflete nas representações armazenadas na memória de longo prazo e empregadas nos processos perceptivos. Para isto, o contraste entre as vogais médias posteriores [o] e [ɔ] e o contraste entre [o] e [u] (sendo o estatuto fonológico do último isento de dúvidas) foram comparados em relação aos resultados obtidos por doze falantes do Português Brasileiro em uma tarefa de classificação de vogais ao longo de um *continuum* [u]-[o]-[ɔ] e ao grau em que estes mesmos resultados prevêm os resultados nas tarefas de discriminação. Duas tarefas de discriminação de sons da fala foram usadas: a tarefa 2AFC, que encoraja estratégias classificatórias baseadas em categorias armazenadas na memória, e a tarefa 4I2AFC, que encoraja estratégias baseadas na informação acústica. A média dos coeficientes de inclinação da curva representando a proporção da resposta “ɔ” foi significativamente menor que a média dos valores correspondentes à resposta “u”. Numa análise entre-sujeitos, foi observada uma relação mais estreita entre os resultados da tarefa de classificação e da tarefa 2AFC no *continuum* [u-o] que no *continuum* [o-ɔ]. Os resultados são interpretados como indícios de que as representações das vogais [o] e [ɔ] são menos distintas entre si que as representações das vogais [o] e [u].

Palavras-chave: vogais médias, percepção da fala, categorias fonéticas, contraste.

ABSTRACT

Minimal pairs are attested in Brazilian Portuguese that establish the contrast between open- and close-mid vowels in the stressed syllable. However, there are cases in which this two categories show free variation in this same context without semantic consequences. The purpose of this study was to verify whether this phenomenon is reflected in the representations stored in long term memory and employed in the perceptual processes. With this aim the contrasts between the back mid vowels [o] and [ɔ] and the contrast between [o] and [u] (the fonological status of the latter is doubtless) were compared in respect to the results obtained by twelve Brazilian Portuguese native speakers in a vowel classification task and to the degree to which those same results predict the results in vowel discrimination tasks.

Two discrimination tasks were used: the 2AFC task, that encourage labeling strategies based on categories stored in memory, and the 4I2AFC task, that encourages strategies based on the acoustical information. The mean of the slopes of the curves representing the proportions of “o” responses was significantly less then the slopes observed for the “u” response. It was observed in a between subjects analysis a closer relationship between classification and discrimination results for the [u -o] continuum than for the [o-u] continuum. The results are interpreted as suggesting that the representations of the vowels [o] and [ɔ] are less distinct from each other than are the categories of [o] and [u].

Key-words: mid vowels, speech perception, phonetic categories, contrast.

LISTA DE FIGURAS

FIGURA 1 - Esquema de uma prova na tarefa de classificação, se iniciando com a apresentação do som S e terminando com um intervalo de 1s após a resposta.....	54
FIGURA 2 - Esquema de uma prova da tarefa de discriminação 2AFC.....	55
FIGURA 3 - Esquema de uma prova da tarefa de discriminação 4I2AFC em que o som diferente ocorre no terceiro intervalo.....	57
GRÁFICO 1 - Proporções das respostas “u”, “o” e “ɔ” ao longo do <i>continuum</i> em F1.....	65
GRÁFICO 2 - Valores de probito das proporções das respostas ao longo do <i>continuum</i> em F1.....	67
GRÁFICO 3 - Graus de ambigüidade dos estímulos do <i>continuum</i> em F1.....	68
GRÁFICO 4 - Discriminação obtida e discriminação prevista (classificação) no <i>continuum</i> [u-o]; tarefa 2AFC.....	72
GRÁFICO 5 - Discriminação obtida e discriminação prevista (classificação) no <i>continuum</i> [o-ɔ]; tarefa 2AFC.....	72
GRÁFICO 6 - Discriminação obtida e discriminação prevista (classificação) no <i>continuum</i> [u-o]; tarefa 4I2AFC.....	73
GRÁFICO 7 - Discriminação obtida e discriminação prevista (classificação) no <i>continuum</i> [o-ɔ] tarefa 4I2AFC.....	73

LISTA DE TABELAS

TABELA 1 - Médias dos índices de percepção categórica (PC) nas tarefas 2AFC e 4I2AFC para os <i>continua</i> [u-o] e [o-ɔ] e significância (p bicaudal) da diferença entre os valores correspondentes aos dois <i>continua</i> (análise intra-sujeitos).....	75
TABELA 2 - Índice de percepção categórica (PC), coeficiente de correlação entre discriminação obtida e prevista (r) e desvio quadrático médio (DQM) da discriminação obtida em relação à discriminação prevista, calculados a partir das médias de d' do grupo de sujeitos (análise entre-sujeitos) nas tarefas 2AFC e 4I2AFC para os <i>continua</i> [u-o] e [o-ɔ].....	76
TABELA A 1 - Proporções das respostas “u”, “o” e “ɔ” na tarefa de classificação.....	93
TABELA B 1 - Valores de d' no continuum [u-o].....	94
TABELA B 2 - Valores de d' no continuum [o-ɔ].....	95

1 Introdução

Tradicionalmente o sistema vocálico do Português Brasileiro é descrito como um sistema composto por sete categorias distintivas na sílaba tônica, os fonemas /i, e, ε, a, o, u/. Nas sílabas átonas, os contrastes ε/e e o/o seriam então neutralizados, o que reduz o número de vogais a cinco. Uma segunda neutralização nas sílabas átonas em final de palavra resulta em um sistema de três vogais pela anulação do contraste fonológico entre as vogais médias e as vogais altas.

Entretanto, há variação na sílaba tônica entre vogais médias abertas e fechadas em nomes como, por exemplo, “[e]xtra/[ε]xtra” e cr[o]sta/cr[ɔ]sta. Ademais, a realização das vogais médias, no que concerne ao grau de abertura, apresentaria um grau considerável de previsibilidade. Deste modo, é possível colocar em questão o estatuto fonológico do contraste entre vogais médias abertas e fechadas.

Se adotarmos uma concepção em que o fonema não é apenas um construto da teoria fonológica, mas um conceito que se refere a uma representação, ou categoria mental, armazenada na memória de longo prazo que se manifesta no processamento lingüístico do falante, da percepção à produção, tornam-se de extremo interesse, especialmente em relação a este debate sobre o estatuto fonológico da oposição entre as vogais médias abertas e fechadas, evidências empíricas independentes da fonologia relacionadas à forma com que tais vogais são representadas na mente do falante do Português Brasileiro.

Com este estudo se busca efetuar os primeiros passos de uma investigação empírica sistemática a respeito do modo como as distinções entre as vogais médias abertas e as vogais médias fechadas estão representadas no sistema perceptivo do falante do Português Brasileiro.

Mais especificamente, foram realizados experimentos cujos resultados fornecem uma série de informações sobre a localização e os limites das categorias fonéticas correspondentes às vogais médias posteriores e sobre a relação entre a classificação de sons nestas categorias e a discriminação entre eles. O objetivo principal é comparar os padrões observados no limite entre as categorias [o] e [ɔ] com os padrões observados no limite entre as categorias [o] e [u] de maneira a verificar se há evidências de que o contraste [o/ɔ] é menos consistente (no que se refere ao às representações na memória de longo prazo) que o contraste [u/o], um contraste sobre a funcionalidade do qual não há margens para dúvidas.

O conceito de percepção categórica, muito importante e tradicional na pesquisa sobre a percepção da fala, é de particular interesse neste trabalho. Basicamente, percepção categórica significa equivalência, ou associação estreita, entre a classificação e a discriminação de estímulos. Isto é, dois estímulos são discrimináveis na medida em que compõe categorias diferentes. Na acepção mais antiga (STUDDERT-KENNEDY; LIBERMAN; HARRIS; COOPER, 1970), há percepção categórica quando dois estímulos são discrimináveis se e somente se pertencerem a categorias diferentes. Porém, após décadas de pesquisa, pode-se concluir com segurança que não é isso o que se observa, tornando-se possível falar em graus de percepção categórica (VANHESSSEN; SCHOUTEN 1999). Um conceito intimamente relacionado é o conceito de “efeito de limite de categoria”, que se refere aos picos de discriminabilidade nas regiões dos limites entre categorias. Considera-se que estes dois conceitos se refiram aos fenômenos que são manifestações das representações de categorias de sons da fala armazenadas na memória de longo prazo e empregadas na percepção.

Pretende-se observar ao longo de um *continuum* composto por sons variando entre as categorias [u] e [ɔ], passando pela categoria [o], as formas das curvas que representam a

classificação dos estímulos nas três categorias e os graus de percepção categórica (isto é, os graus em que a discriminação e a classificação se relacionam) correspondentes aos contrastes [u/o] e [o/ɔ]. Com isto tenta-se obter elementos que ajudem a responder uma pergunta específica (“há evidências de que a distinção entre [o] e [ɔ] é menos saliente que a distinção entre [o] e [u] no nível das categorias mentais armazenadas na memória de longo prazo e empregadas na percepção da fala?”) que é formulada no âmbito de uma questão mais geral (“são os contrastes entre as vogais médias abertas e as vogais médias fechadas, do ponto de vista do processamento perceptivo, menos distintos do que os demais contrastes entre as vogais do Português Brasileiro?”).

É importante salientar que não há a pretensão de, a partir dos resultados deste estudo, responder a perguntas que devem ser discutidas principalmente no âmbito da fonologia. Este trabalho é apenas um passo numa investigação bem mais ampla sobre as representações mentais das vogais médias do Português Brasileiro. Se os futuros frutos desta investigação devem ou não ser levados em consideração na formulação de modelos fonológicos, isto é algo que deve ser ponderado.

Devido à natureza dos resultados deste trabalho, é difícil a integração dos mesmos a teorias de orientação formalista. Por outro lado, nas teorias funcionalistas evidências experimentais acerca da percepção da fala ocupam uma posição de destaque e cumprem um papel fundamental (BYBEE, 2001; BOERSMA, 1998).

2 Revisão Bibliográfica

Neste capítulo apresenta-se a literatura pertinente à questão estabelecida na introdução e aos modos de abordá-la empiricamente, em três seções. Na primeira, há uma breve descrição do fenômeno a partir do qual o problema deste estudo foi definido, isto é, a variação entre as vogais médias abertas e fechadas em num ambiente em que são verificados pares mínimos. Na seção seguinte são apresentadas e discutidas formas pelas quais as representações empregadas pelos falantes na percepção dos sons da fala podem ser investigadas experimentalmente. São contemplados resultados de experimentos comportamentais e evidências psicofisiológicas. Em seguida, algumas são discutidas algumas questões importantes no que se refere à justificativa do método aqui empregado e à interpretação dos resultados.

2.1 As vogais médias do Português Brasileiro

Considera-se tradicionalmente que no inventário fonológico do Português Brasileiro há sete diferentes fonemas vogais, quais sejam, /i, e, ε, a, ɔ, o, u/ (MATTOSO CÂMARA JR., 1977; WETZELS, 1992). Trata-se então de um sistema vocálico composto por uma vogal baixa, duas altas que contrastam entre si pela anterioridade/posterioridade, e quatro médias, duas anteriores e duas posteriores. A diferença entre as duas vogais médias anteriores e entre as duas vogais médias posteriores é estabelecida pelo grau de abertura (ou altura) apresentado por cada vogal. Desta forma temos as vogais médias abertas e as vogais médias fechadas.

Considera-se que o contraste entre as vogais médias abertas e fechadas é funcional no Português Brasileiro apenas na sílaba tônica. Esta consiste em uma posição privilegiada no

que concerne ao estabelecimento de contrastes fonológicos (BECKMAN, 1998). O valor fonológico desta oposição na sílaba tônica pode ser atestado em pares mínimos como s[e]de/s[ɛ]de e t[o]ca/t[ɔ]ca. Nas sílabas átonas, entretanto, ele é anulado, uma vez que as oposições e/ɛ e o/ɔ não estabelecem diferenças de significado. Nas sílabas pós-tônicas em final de palavra se perdem ainda as distinções u/o/ɔ e i/e/ɛ, o que resulta no subsistema de três vogais /i, a, u/. Na posição pós-tônica não-final, de acordo com Mattoso Câmara Jr. (1987), ocorre a neutralização do contraste entre /o/ e /u/, mas não do contraste entre /e/ e /i/, produzindo um sistema assimétrico composto pelas vogais /a, e, i, u/ – assimétrico porque a vogal média anterior /e/ não tem uma correspondente posterior.

Nas análises de Mattoso Câmara Jr. (1987) e, mais recentemente, Wetzels (1992), são considerados três processos de neutralização, que originam quatro subsistemas vocálicos, o de sete vogais para as sílabas tônicas, o de cinco vogais para as sílabas pré-tônicas, o de quatro vogais para as sílabas pós-tônicas não-finais e o de três vogais para as sílabas átonas finais. A primeira neutralização eliminaria as vogais médias abertas ε/ɔ nas sílabas átonas, a segunda eliminaria a vogal média /o/ das sílabas pós-tônicas e a terceira eliminaria a vogal /e/ das pós-tônicas finais. Bisol (2003) sugere que o Português Brasileiro possui apenas duas regras de neutralização, uma que neutraliza os contrastes ε/e e o/ɔ e outra que neutraliza o contraste entre médias e altas. A primeira se aplicaria categoricamente às sílabas átonas e a segunda se aplicaria categoricamente às sílabas átonas finais e variavelmente às sílabas átonas pós-tônicas não-finais. A autora argumenta que como, em alguns casos, a vogal /e/ também se mostra sensível ao alçamento na posição pré-tônica não-final e a vogal /o/ é mantida nesta mesma posição (VIEIRA, 2002, *apud* BISOL, 2003), parece haver uma expansão do sistema mínimo de três vogais que cria uma variação entre ele e o subsistema de cinco vogais, sendo o

subsistema de quatro vogais proposto nos modelos tradicionais apenas um efeito de frequência.

Estas análises admitem um sistema vocálico composto por sete fonemas, já que a primeira neutralização – aquela que reduz o número de vogais a cinco anulando o contraste fonológico entre as vogais médias abertas e fechadas – obviamente supõe um sistema com esse número de componentes. Entretanto, alguns aspectos do comportamento das vogais médias na sílaba tônica suscitam algumas dúvidas a respeito do estatuto fonológico da oposição entre vogais médias abertas e fechadas e, por conseguinte, sobre se realmente o Português Brasileiro possui sete fonemas vogais. Em primeiro lugar, a oposição em questão não se apresenta entre os verbos (Lee, 2003; Magalhães, 1990), embora o timbre da vogal média estabeleça muitos pares mínimos compostos por uma forma nominal e a sua correspondente verbal como, por exemplo, alm[o]ço/alm[ɔ]ço e desesp[e]ro/desp[ɛ]ro. Em segundo lugar, apesar dos pares mínimos encontrados entre os nomes, que são em número reduzido, há variação no timbre das vogais médias tônicas no que concerne ao grau de abertura, tanto no caso das anteriores como no caso das posteriores (CUNHA, 1991; ALVES, 1999). Isto é, uma mesma palavra com uma vogal média na sílaba tônica pode ser pronunciada com esta vogal média aberta ou fechada, como nos casos de [e]xtra/[ɛ]xtra, av[e]ssas/av[ɛ]ssas, p[o]ça/p[ɔ]ça e cr[o]sta/cr[ɔ]sta. Em terceiro lugar, afirma-se que a realização das vogais médias, no que diz respeito ao grau de abertura, apresenta algum grau de previsibilidade (WETZELS, 1992; CAGLIARI, 1997). Cagliari (1997) sugere a possibilidade de se admitir um sistema vocálico com cinco vogais, não sendo [ɛ] e [ɔ] fonemas da língua portuguesa, mas ocorrendo foneticamente na sílaba tônica.

Um ponto que torna a idéia de um sistema com cinco vogais atrativa é o fato de que é este o número de vogais nasais no Português Brasileiro, considerando-se que uma descrição

em que o número de vogais nasais e orais é o mesmo, seria mais elegante. Os pares mínimos são poucos e são observados apenas na sílaba tônica, contexto este em que também há variação. Por outro lado, a simples verificação de alguns pares mínimos, juntamente com a sistemática alternância entre vogais médias fechadas e vogais médias abertas em formas nominais e verbais constitui evidência de que as representações correspondentes a cada uma se distinguem na mente do falante não apenas em níveis de processamento de informação acústica.

2.2. A Categorização dos Sons da Fala

2.2.1. Percepção categórica

Durante o processamento lingüístico, no que concerne à produção e à percepção da fala, devem ser manipuladas representações mentais correspondentes às categorias dos sons da fala, isto é, grupos de sons dentro dos quais é atribuída uma mesma identidade a todos os componentes. Por exemplo, se imaginarmos um plano estabelecido por dois eixos que representam o primeiro e o segundo formantes (F1 e F2) de um som vocálico, todos os sons em volta de um ponto e dentro de um determinado limite serão considerados como exemplares de certa vogal, como [a], sendo agrupamentos em outras regiões do plano correspondentes a outras vogais.

Nos laboratórios Haskins, o primeiro sintetizador de fala orientado para pesquisa, o “*Pattern Playback*”, foi empregado por Liberman, Harris, Hoffman e Griffith (1957) para produzir um *continuum* que se inicia em [be], passa por [de], e termina em [ge]. Isto foi feito por meio da

variação em passos iguais do valor da frequência inicial da transição do segundo formante. Foi observado que, numa tarefa de classificação, embora formassem um *continuum* físico, os estímulos eram classificados pelos sujeitos do experimento em três categorias bem definidas, ocorrendo nas regiões limítrofes entre elas mudanças abruptas nas respostas. Para testar se as diferenças físicas entre estímulos dentro de uma mesma categoria assim estabelecida podem ser detectadas pelos ouvintes, foi realizada uma tarefa de discriminação ABX, em que dois estímulos diferentes, A e B, são apresentados em sucessão e o sujeito indica qual deles é idêntico a um terceiro estímulo, X. Os sujeitos discriminaram com relativa facilidade estímulos identificados como pertencentes a categorias diferentes e apresentaram um desempenho inferior na discriminação intra-categoria. Este padrão nos resultados é denominado “efeito de limite de fonema”.

A partir da hipótese de que os sujeitos são capazes de discriminar apenas estímulos de categorias diferentes, isto é, que não há informação além daquela relativa à categoria fonêmica, Liberman *et al.* (1957) empregaram os resultados da tarefa de identificação para gerar uma equação para prever o desempenho na tarefa de discriminação. O termo “percepção categórica” foi cunhado para designar a dependência exclusiva da discriminação em relação à identificação – isto é, dois sons podem ser discriminados apenas na medida em que eles podem ser identificados como pertencentes a categorias diferentes. Embora a equação tenha predito razoavelmente bem os picos nas funções que representam o desempenho dos sujeitos na discriminação de cada par de estímulos, este foi subestimado, o que sugere que o ouvinte dispõe de informações adicionais sobre os estímulos. Todavia, a correlação significativa observada entre a discriminação prevista (classificação) e a discriminação observada foi tomada como uma importante evidência de percepção categórica.

Resultados semelhantes foram encontrados em vários estudos que se seguiram, especialmente no caso das consoantes oclusivas, e com exceção das vogais (revisão em REPP, 1984). Quase sempre foram observados um efeito de limite de fonema e uma discriminação obtida superior à discriminação prevista (classificação). Enquanto no caso das consoantes oclusivas era observado um padrão próximo do ideal de percepção categórica, nos estudos com vogais, embora tenha sido frequentemente encontrado um efeito de limite de fonema, este se mostrou menos pronunciado, e o desempenho na discriminação é sempre consideravelmente superior ao previsto pela classificação (FRY; ABRAMSON; EIMAS; LIBERMAN, 1962; EIMAS, 1963; HEALY; REPP, 1982, revisão em REPP, 1984). Os primeiros pesquisadores dos laboratórios Haskins não atribuíram grande importância à sistemática superioridade da discriminação em relação à identificação, e estabeleceram os seguintes critérios para percepção categórica (STUDDERT-KENNEDY; LIBERMAN; HARRIS; COOPER, 1970):

- a) Categorias distintas com limites bem definidos na tarefa de classificação.
- b) Regiões ou “vales” de desempenho em nível aleatório na discriminação de estímulos da mesma categoria de identificação.
- c) Um pico de desempenho no limite entre duas categorias (efeito de limite de fonema) na tarefa de discriminação.
- d) Uma correspondência estreita entre o desempenho obtido na discriminação e o desempenho previsto a partir dos resultados da tarefa de classificação.

Embora uma percepção perfeitamente categórica não tenha sido encontrada na grande maioria dos estudos, essa passou a ser a definição padrão de percepção categórica, estabelecendo uma dicotomia entre “percepção categórica” e “percepção contínua”.

É importante notar que a divisão do *continuum* em categorias, geralmente observada em tarefas de classificação de sons da fala, com limites de categoria pronunciados, não pode ser

interpretada como evidência de percepção categórica, pois é natural que ela se apresente também em casos de percepção contínua (MASSARO, 1987a). Dado um *continuum* percebido de forma contínua entre A e não-A, podemos imaginar uma reta descendente, de 1 a 0, entre A e não-A, indicando o grau $G(A)$ em que o estímulo representa a categoria A. Em uma tarefa de classificação em que o sujeito deve responder A ou não-A, uma regra ótima de decisão seria fixar o valor do critério de decisão no ponto $G(A) = 0,5$ – isto é, se $G(A) < 0,5$, responder “não-A”; se $G(A) > 0,5$, responder “A”. Admitindo que haja ruído (variabilidade) nos processos perceptivos, e que este ruído tem distribuição normal, vemos que um mesmo estímulo apresentado várias vezes não produz sempre o mesmo valor em $G(A)$, mas sim uma distribuição normal de valores em torno de uma média diretamente relacionada ao valor da variável A. Desta forma, se o ruído é normal e tem a mesma variância ao longo do *continuum*, um estímulo cujo valor médio em $G(A)$ é igual ao valor do critério de decisão, produzirá respostas “não-A” em metade das provas e respostas A na outra metade. Na medida em que o valor médio de $G(A)$ se afasta do valor do critério, o ruído tem um efeito menor nas respostas. Isso produz, a partir de informação *continua*, uma curva de classificação em que declinações abruptas são observadas representando os limites de categoria. Desta forma, resultados de tarefas de discriminação são imprescindíveis no que concerne às investigações acerca do conceito de percepção categórica.

De acordo com a visão tradicional do fenômeno, a percepção categórica é interessante por violar certas expectativas estabelecidas no âmbito da psicofísica. Em primeiro lugar, o efeito de limite do fonema estaria em desacordo com a lei de Weber, segundo a qual a “diferença minimamente perceptível” entre dois estímulos é uma razão constante do valor correspondente ao estímulo-padrão (aquele com o qual se compara um segundo estímulo) na magnitude física em relação à qual se realiza a comparação. Isto é, de acordo com a lei de

Weber, a discriminabilidade é constante para estímulos separados por uma razão fixa de magnitudes físicas, e decresce para estímulos separados por uma diferença fixa. Em segundo lugar, a percepção categórica divergiria de uma importante generalização de Miller (1956), segundo a qual, para a maioria dos *continua* de estímulos, a discriminação é mais fácil do que a identificação. Isto motivou a idéia de que os estímulos da fala são especiais, sendo tratados de forma diferente e outros estímulos acústicos. Macmillan (1987) argumenta, entretanto, que a maioria dos *continua* perceptuais podem violar a lei de Weber de formas que se generalizam naturalmente para dimensões percebidas categoricamente e que o princípio de Miller é geralmente verdadeiro, inclusive para os sons da fala. De fato, como já foi mencionado, os dados em geral evidenciam uma superioridade da discriminação em relação à identificação.

As diferenças nos resultados para vogais isoladas e consoantes oclusivas no que se refere à percepção categórica, foram tomadas como uma das principais evidências favoráveis à teoria motora da percepção da fala (STTUDERT-KENNEDY; LIBERMAN; HARRIS; COOPER, 1970; LIBERMAN; MATTINGLY, 1985), segundo a qual os mecanismos de percepção da fala estão intimamente relacionados aos mecanismos de produção, envolvendo inferências sobre os gestos articulatórios necessários para a realização da unidade a ser percebida. Descontinuidades na articulação seriam responsáveis pelas descontinuidades na percepção. Desta forma, categorias fonéticas que, no que concerne à sua produção, diferem umas das outras por gestos articulatórios discretos serão percebidas categoricamente, ao passo que as categorias que permitem variações articulatórias contínuas (como as vogais) serão percebidas continuamente.

Para explicar as diferenças entre a discriminação e a identificação, Fujisaki e Kawashima (1969; 1970; 1971 *apud* PISONI, 1973; 1975) formularam um modelo de duplo processo para

a discriminação de sons da fala, em que dois componentes funcionam simultaneamente ou em rápida sucessão, sendo o primeiro estritamente categórico, representando a classificação fonética, e o segundo contínuo, representando processos gerais da percepção auditiva. A discriminação observada seria superior porque, quando o ouvinte tenta discriminar dois estímulos classificados de forma idêntica, ele pode recorrer a informações de natureza psicoacústica. O modelo permite estabelecer graus de percepção categórica, que refletem a previsibilidade da discriminação pela classificação.

A partir de demonstrações de percepção categórica para *continua* de estímulos não relacionados à fala (por exemplo, CUTTING; ROSNER, 1974; PASTORE; AHROON; BAFFUTO; FRIEDMAN; PULEO; FINK, 1977), psicofísicos sugeriram que a percepção categórica poderia ser um fenômeno da percepção auditiva. Algumas dimensões auditivas da fala podem não ser contínuas, e pode ser que existam limiares psicoacústicos que coincidem com limites entre categorias de sons da fala (PASTORE *et al.*, 1977; PASTORE, 1987). Isto poderia explicar de maneira simples resultados de pesquisas realizadas com bebês (EIMAS; MILLER; JUSCZYK, 1987) e animais (KUHL, 1987) nas quais foram encontradas evidências de percepção categórica dos sons da fala, inclusive limites de categorias que coincidem com os apresentados pelos falantes adultos. Uma descontinuidade perceptiva determinada por características do sistema sensorial foi observada em relação ao contraste entre oclusivas vozeadas e surdas. Sinex e McDonald (1989) verificaram que fibras do nervo auditivo de chinchilas respondem a um *continuum* no parâmetro “tempo de início de vozeamento” (VOT, do inglês *voice onset time*) de forma a não discriminar sílabas com o valores de VOT de 0 a 20 ms. Kewley-Port e Watson (1994) observaram aumento na diferença minimamente perceptível entre vogais quando a frequência central do formante que

é variado coincide com um harmônico, mas neste caso não foram apresentadas explicações para o fenômeno.

Ades (1977) sugere que o modelo de Durlach e Braida de discriminação de intensidade sonora (DURLACH; BRAIDA, 1969; BRAIDA; DURLACH, 1972) pode ser muito útil na formulação de um modelo adequado para a percepção categórica. Este modelo admite dois componentes de memória, um “modo de traço sensorial”, por meio do qual o sujeito compara os traços dos estímulos a serem discriminados, e um “modo de contexto”. Neste, o sujeito tenta relacionar os dados sensoriais ao contexto geral dos estímulos apresentados no experimento. O modelo especifica três fontes de variância que limitam o desempenho na discriminação e na classificação. A primeira é a variância sensorial, oriunda de processos sensoriais como a transdução. A variância do traço limita a habilidade do sujeito em comparar os traços sensoriais de dois sons consecutivos. Esta aumenta com o intervalo de tempo entre os dois estímulos a serem comparados. A variância de contexto é relacionada com a habilidade do sujeito em rotular os estímulos e aumenta com a extensão do *continuum* de sons (extensão significando a diferença entre os pontos extremos do *continuum*). A variância sensorial limita o desempenho em todos os tipos de tarefa. Na tarefa de classificação, como apenas um som é apresentado em cada prova, apenas a variância de contexto e a variância sensorial impõem restrições ao desempenho. Em tarefas de discriminação de nível fixo, em que os mesmos dois sons são apresentados ao longo de cada bloco experimental, a variância de contexto é muito pequena, e considera-se que a variância do traço não seja importante, de modo que o desempenho é limitado apenas pela variância sensorial. Nas tarefas de discriminação de nível variável, em que os dois sons a serem comparados variam entre as provas num mesmo bloco experimental, os dois componentes de memória, isto é, o modo de contexto e o modo de traço sensorial, podem influenciar o desempenho, sendo a relação entre

eles análoga à relação entre duas resistências em paralelo. Quando a variância de um tipo de memória é muito menor que o outro, o processo que envolve maior variância tem um efeito muito pequeno sobre o desempenho. Assim, por exemplo, se a variância de contexto é muito maior do que a variância do traço, o modo de traço sensorial, mais eficiente, predomina.

De acordo com Ades (1977) o modelo poderia explicar as diferenças encontradas entre vogais e consoantes e sons da fala e outros estímulos sonoros. Macmillan (1987) e Macmillan, Goldberg e Braida (1988) apresentam uma generalização para o estudo de sons da fala de uma versão revisada do modelo de Durlach e Braida em que é incluído o conceito de “âncora perceptiva” para explicar como a variância de contexto pode depender da extensão do *continuum* (BRAIDA; LIM; BERLINER; DURLACH; RABINOWITZ; PURKS; 1984). No modo de contexto, o sujeito compara as sensações com âncoras perceptivas, isto é, representações na memória usadas como referência. Quanto maior a distância entre o evento perceptivo e a âncora pior a resolução. Na formulação do modelo para a discriminação de intensidade, as âncoras perceptivas geralmente correspondem aos pontos extremos do *continuum*, sendo assim não permanentes, embora se admita a possibilidade de âncoras permanentes como “o som mais intenso possível”. Na versão do modelo adaptada para percepção dos sons da fala, as âncoras se relacionam às categorias de sons representadas na memória de longo prazo, podendo corresponder a limites de categoria ou a protótipos (o que os autores denotam por “protótipo” é a região central da curva que representa uma categoria no resultado da tarefa de classificação).

Schouten e van Hessen (1992) e van Hessen e Schouten (1992) propõe uma teoria que combina elementos desse modelo com o modelo de duplo processo de Fujisaki e Kawashima. Estes autores, com base em resultados em diversas tarefas de discriminação, variando os

intervalos inter-estímulos, em comparação com os resultados de uma tarefa de classificação, argumentam que os resultados da discriminação de vogais são mais favoráveis ao modelo de Durlach e Braida, ao passo que a discriminação de consoantes oclusivas ocorre de modo mais coerente com o modelo de duplo processo de Fujisaki e Kawashima, pois nos resultados referentes a estas consoantes foram observadas evidências de percepção categórica e não houve indícios de que a discriminação envolva o modo de contexto. Eles formulam então um modelo em que é incluída uma “memória fonêmica de longo prazo”, além dos componentes do modelo de Durlach e Braida, de forma a obter um ajuste melhor em relação aos resultados obtidos com as consoantes oclusivas.

Considerando os propósitos do estudo aqui proposto, este paradigma experimental (a comparação dos resultados observados na tarefa de discriminação com os resultados previstos para a mesma a partir dos resultados da tarefa de classificação) proporciona medidas interessantes. Trata-se de um método empírico que torna possível observar sistemática e objetivamente, por meio do desempenho do falante em certas tarefas, a forma como os falantes categorizam os sons da fala que percebem. Entretanto, há alguns problemas. Em primeiro lugar, por serem baseados em medidas comportamentais, os experimentos não permitem separar o componente perceptivo, que é o que interessa, da atenção, vontade e processos de decisão do sujeito.

Outro aspecto problemático dos resultados dos estudos sobre percepção categórica é a grande variabilidade no grau em que a identificação prediz a discriminação. As fontes desta variabilidade são diversas. Uma delas é a natureza dos estímulos empregados. Entre os sons da fala, as consoantes oclusivas são as que apresentam o maior grau de percepção categórica. As nasais são percebidas menos categoricamente que as oclusivas; as líquidas, semivogais e

vogais ainda menos (revisão em REPP, 1984). A “naturalidade” dos estímulos parece ser também um fator importante. Van Hesson e Schouten (1999) fornecem evidências de que a maior complexidade dos estímulos de fala mais naturais dificulta o uso pelo ouvinte de parâmetros particulares dos estímulos como pistas psicoacústicas, o que aumenta o grau de percepção categórica. Eles mostraram que, quanto mais natural soa o estímulo, maior o índice de percepção categórica.

Embora tenha sido inicialmente sugerido que a percepção categórica seria específica para os sons da fala, *continua* em outros domínios se mostraram percebidos de forma categórica. Cutting e Rosner (1974) obtiveram evidências de percepção categórica para um *continuum* entre sons percebidos como produzidos por “palheta” e “arco” estabelecido pela variação do tempo de elevação da amplitude. Entretanto, foi observado posteriormente que a diferença acústica era na realidade maior na região em que se observaram os picos nas curvas de discriminação prevista e observada. Pastore, Ahroon, Baffuto, Friedman, Puleo e Fink (1977) empregaram como estímulo uma luz tremulante em diferentes frequências e observaram um limite bem definido entre as categorias “tremulante” e “contínua”. Foi também realizado um experimento em que sons eram classificados e discriminados em relação à intensidade. Quando os sons de teste eram apresentados no silêncio não se verificou relação entre discriminação e classificação, mas quando as provas eram apresentadas com um fundo que consistia em um som de referência, a função de classificação apresentou limites precisos entre as categorias “mais intenso” e “mais fraco” e um pico correspondente na tarefa de discriminação. Esta observação é particularmente interessante, pois mostra como uma descontinuidade perceptiva (estabelecida pela discriminação de intensidades relativas em relação a um som de referência) pode determinar resultados que corroboram a hipótese da percepção categórica num *continuum* que na verdade é percebido de forma *continua*.

Outros fatores importantes na determinação do grau de percepção categórica obtido nos resultados são o intervalo entre estímulos, os sujeitos e a tarefa de discriminação empregada. Van Hessen e Schouten (1992) observaram que o intervalo entre estímulos afeta o grau de percepção categórica e sugerem que intervalos mais longos aumentam a probabilidade de que os sujeitos usem uma estratégia menos acústica e mais classificatória em consequência do rápido decaimento dos traços auditivos na memória sensorial.

Gerrits (2001) notou uma significativa variabilidade entre os sujeitos em vários experimentos com vogais e consoantes oclusivas. Ela mostra, além a variabilidade no desempenho na tarefa de discriminação, que, em situações idênticas, alguns sujeitos empregam estratégias mais baseadas nas categorias fonêmicas, ao passo que outros utilizam mais as informações relacionadas aos traços auditivos dos estímulos.

Os resultados da comparação entre as tarefas de identificação e discriminação têm se mostrado extremamente sensíveis às tarefas empregadas no experimento, especialmente no que concerne à tarefa de discriminação. A tarefa de discriminação padrão entre os estudos de percepção categórica é a tarefa ABX. Nesta tarefa, o sujeito ouve em cada tentativa uma série de três estímulos consecutivos em que o terceiro é sempre idêntico ao primeiro ou ao segundo. O sujeito deve responder qual entre os dois primeiros é idêntico ao terceiro. Massaro e Cohen (1983) argumentam que esta tarefa apresenta um viés em relação aos resultados dos estudos da percepção categórica. Em virtude da duração relativamente pequena dos traços acústicos na memória auditiva, quando o estímulo X é apresentado e o sujeito tenta comparar os traços auditivos deste com os de A e B, estes traços podem já ter decaído, de forma que o sujeito é obrigado a recorrer apenas às identificações atribuídas a A e B

(estímulos cujas codificações se encontram em períodos diferentes do processamento no momento da comparação, em virtude do lapso temporal entre os dois). Isto, obviamente, produziria os resultados geralmente atribuídos à percepção categórica.

Uma tarefa que reduz a carga na memória auditiva e poderia encorajar uma comparação auditiva direta é a tarefa AX (ou “igual-diferente”). Nesta tarefa o sujeito simplesmente ouve dois estímulos e indica se eles são iguais ou diferentes. Entretanto, esta tarefa não é livre de vieses, uma vez que os sujeitos podem usar diferentes critérios de decisão para responder “igual” ou “diferente”. Um destes critérios pode ser justamente a classificação do estímulo em uma categoria fonêmica. Healy e Repp (1982), por exemplo, usando uma tarefa de discriminação AX e uma tarefa de identificação em que exatamente os mesmos pares de estímulos usados na tarefa de discriminação eram classificados (com este procedimento os autores pretendiam eliminar efeitos de contexto), observaram um desempenho na discriminação completamente determinado pela identificação para estímulos num *continuum* /ba/ - /da/.

Uma tarefa muito usada em experimentos psicofísicos é a tarefa 2AFC (*two alternatives forced choice*). Esta tarefa facilita análises em termos da Teoria da Detecção do Sinal (GREEN; SWEETS, 1966; MACMILLAN; CREELMAN, 2005). Em experimentos sobre percepção de intensidade, por exemplo, o sujeito ouve dois estímulos que variam em intensidade e responde indicando a ordem em que eles foram apresentados, intenso-fraco ou fraco-intenso. No caso dos sons da fala é necessário explicar ao sujeito que o termo “ordem” se refere às categorias fonêmicas – por exemplo “i-u” ou “u-i”. É evidente o viés dessa tarefa, causado pelo encorajamento do comportamento classificatório.

A tarefa 4IAX (*four intervals AX*) é uma tarefa considerada menos enviesada. Há oito combinações possíveis, quais sejam, ABAA, BAAA, AAAB, AABA, BABB, ABBA, BBBA e BBAB. O sujeito deve indicar qual dos dois pares apresenta dois estímulos idênticos. Assume-se que os sujeitos determinam as diferenças entre os estímulos de cada par e em seguida qual das duas diferenças é a menor. Desta forma, a decisão é livre de critérios subjetivos. Uma tarefa similar e menos trabalhosa é a 4I2AFC (*four intervals two alternatives forced choice*), em que apenas as duas seqüências AABA e ABAA são possíveis e o sujeito deve responder indicando se a o estímulo díspar – isto é, B – está na segunda ou na terceira posição. Neste caso o sujeito poderia ignorar o primeiro e o quarto estímulo, desempenhando uma tarefa 2AFC, mas esses podem ser usados como referência de modo a tornar possível uma decisão baseada apenas nos traços sensoriais.

Gerrits (2001) comparou os resultados de experimentos com vogais e consoantes oclusivas empregando as tarefas 2AFC, AX, AXB, 4IAX e 4I2AFC e observou consideráveis diferenças no grau de percepção categórica. Para as tarefas 4I2AFC e 4IAX não houve correlação significativa com a tarefa de classificação (embora alguns sujeitos tenham apresentado um pico de desempenho na região do limite fonêmico). A tarefa 2AFC apresentou a relação mais estreita com a classificação, o que indica que os rótulos atribuídos aos estímulos têm um papel importante durante a mesma. Estes resultados são confirmados por Schouten, Gerrits e van Hessen (2003) e Gerrits e Schouten (2004) em estudos com vogais.

O contexto em que os sons ocorrem na tarefa – isto é, os estímulos imediatamente anterior e posterior – tem também um efeito nos resultados. Healy e Repp (1982), usando uma tarefa de discriminação AX e uma tarefa de identificação em que exatamente os mesmos pares de

estímulos usados na tarefa de discriminação eram usados na tarefa de classificação, observaram um desempenho na discriminação quase que completamente determinado pela classificação para estímulos num *continuum* /ba/ - /da/. Num teste semelhante em que um *continuum* entre as vogais /i/ e /I/ foi empregado, foi observado um efeito de limite de categoria, mas o desempenho na discriminação foi superior ao que foi previsto a partir dos resultados na tarefa de classificação. Para um *continuum* entre as fricativas /j/ e /s/ não foi observada relação entre discriminação e classificação.

Massaro (1998) argumenta que apenas o fato de não encontrarmos indiscriminabilidade intra-categoria já é motivo suficiente para o abandono do termo “percepção categórica”; os processos sensoriais são contínuos e o caráter categórico das respostas é devido a processos de decisão sobre informação contínua. Massaro (1983) empregou uma tarefa de julgamento contínuo em que o sujeito atribui um valor a cada estímulo representando o grau em que ele acha que o estímulo em questão pertence a uma ou a outra categoria. A tarefa foi realizada para os *continua* /bæ/-/dæ/, /bæ/-/pæ/ e /i/-/I/. As distribuições das respostas em cada tarefa foram usadas para testar dois modelos, um categórico e um contínuo, que faziam previsões bem diversas entre si a respeito daquelas. O modelo contínuo se ajustou melhor aos dados, mostrando que havia informação contínua disponível.

Hary e Massaro (1982) observaram que quando um *continuum* em que o tempo de elevação da amplitude do som é variado entre -60 e 60 ms são obtidos resultados categóricos, isto é, a classificação prediz relativamente bem a discriminação, sendo o limite de categorias próximo ao valor zero. Entretanto, quando apenas a metade positiva do *continuum* é apresentada nas tarefas, a discriminação é mais bem descrita pela lei de Weber. Com base nestas evidências e outras, como o experimento realizado por Pastore *et al* (1977) em que foram classificados e

discriminados sons variando em intensidade na presença e na ausência de um tom de referência, os autores argumentam que os tradicionais resultados categóricos referentes à previsibilidade da discriminação em relação à classificação não podem ser considerados evidências de percepção categórica.

Barclay (1972) testou um *continuum* de oclusivas vozeadas em duas tarefas de classificação. Na primeira os sujeitos deviam identificar cada estímulo como uma das três alternativas, “b”, “d” ou “g”. Na segunda, as alternativas possíveis eram apenas /b/ e /g/. Barclay entendeu que, se a percepção fosse categórica, as respostas para os estímulos identificados na primeira tarefa como “d”, seriam aleatórias. Não foi o que ocorreu. As respostas aos estímulos foram realizadas de acordo com a categoria da qual estes se encontravam mais próximos.

Massaro (1987b) apresenta um modelo baseado na lógica difusa (*fuzzy*) que descreve um processo de reconhecimento de padrões que se realiza nos estágios de avaliação, integração e decisão, sendo o comportamento de categorização estabelecido no último. A lógica difusa é uma generalização da lógica clássica que admite valores contínuos de verdade. Cotidianamente ocorrem situações em que classificações como “completamente verdadeiro” ou “completamente falso” não se aplicam. Uma pessoa pode ser considerada como “não tão velha”, a cerveja como “bem gelada”, a gestão de um presidente como “quase catastrófica”. De acordo com o modelo de Massaro, à informação *continua* é atribuído um valor de verdade “*fuzzy*” que expressa o grau em que uma entrada representa uma categoria. O primeiro estágio, a avaliação dos traços, resulta em um valor de verdade associado à presença de cada traço relevante. No segundo estágio, a integração dos traços, a informação proveniente do estágio anterior é comparada com protótipos. A saída é composta por valores de verdade que expressam o grau em que o padrão se ajusta a cada protótipo. Uma consequência desta

formulação é que os traços mais informativos são aqueles que apresentam menos ambigüidade. No estágio em que ocorre a decisão (classificação do padrão), o “mérito” de cada protótipo é avaliado em relação à soma dos méritos de todos os protótipos, sendo selecionado aquele que apresenta o maior valor relativo.

Harnad (1987) argumenta que as conclusões de Massaro, embora consistentes com o fato de que a discriminação não é “tudo-ou-nada”, não explicam o padrão não homogêneo do *continuum* de discriminação em que diferenças físicas com a mesma magnitude são comprimidas intra- e amplificadas entre-categorias.

Schouten, Gerrits e van Hessen (2003) argumentam que, apesar de não ser desejável o abandono do conceito de percepção categórica – importante para explicações sobre como é possível para o falante segmentar em unidades fonêmicas o variável e contínuo sinal da fala – a forma como ele vem sendo investigado, isto é, por meio de experimentos envolvendo identificação e discriminação em um *continuum* de estímulos, não diz muito a respeito de como é possível tal categorização.

Kuhl (1991) empregou um método diferente para estudar a estrutura interna das categorias perceptivas da fala. Primeiramente ela demonstrou, numa tarefa de “julgamento de qualidade”, que sujeitos adultos consideravam alguns membros da categoria /i/ como melhores exemplares. Foi demonstrado então que, quando o melhor exemplar (o protótipo) é comparado com outro estímulo, a discriminação é mais difícil do que quando a comparação é feita entre dois estímulos não-prototípicos. Este efeito, denominado pela autora “efeito do magneto perceptivo”, foi encontrado em adultos e crianças de seis meses, mas em macacos. Isto indica que as categorias da fala são estruturadas de forma a que os membros mais

próximos do protótipo sejam percebidos como muito semelhantes. Lotto, Kluender e Holt (1998) observam que este efeito do magneto perceptivo poderia não ser nada além de uma manifestação do efeito de limite de fonema, resultado da falta de controle sobre um possível deslocamento do limite de fonema que ocorre quando dois sons são apresentados consecutivamente (há uma tendência no sentido de fones consecutivos serem ouvidos como fonemas diferentes). Em uma replicação em que este efeito de contexto foi controlado, os autores não encontram evidências de que as diferenças na discriminabilidade tinham outra fonte que não o efeito de limite de fonema. Entretanto, Iverson e Kuhl (2000) fornecem evidências de que o efeito do magneto perceptivo é menos influenciado por manipulações experimentais relacionadas à atenção e à memória – como o número de estímulos diferentes na tarefa de discriminação – do que o efeito de limite do fonema, o que sugere que os dois efeitos são manifestações de dois processos diferentes. Novamente, trata-se de um tipo de estudo em que as inferências são baseadas em respostas que ocorrem em momentos em que vários processos posteriores ao que se pretende investigar se sucederam. Como se viu, são contraditórias as evidências psicofísicas sobre percepção categórica.

2.2.2 Evidências psicofisiológicas

Uma forma de estudar a percepção das categorias da fala sem a necessidade de se recorrer a medidas indiretas, como os experimentos comportamentais de discriminação, é o registro da atividade cerebral relacionada ao processamento das representações destas categorias.

Há um componente de potenciais evocados, denominado *Mismatch Negativity* (MMN) que proporciona uma medida mais direta da representação mental de objetos auditivos (NÄÄTÄNEN; GAILLARD; MÄNTYSALO, 1978; NÄÄTÄNEN; WINKLER, 1999;

ROTHE-NEVES, 2001). Os potenciais evocados consistem em variações no eletroencefalograma que expressam alterações na atividade elétrica cerebral em resposta a estímulos externos. A MMN ocorre em resposta a estímulos raros (estímulos desviantes) em meio a uma série de estímulos idênticos (estímulos Padrão) e normalmente atinge um ponto máximo entre 100 e 200 ms após o início do estímulo e apresenta amplitude máxima nos registros feitos nas áreas frontais e centrais do crânio. Este componente foi observado para diversos tipos de diferenças entre estímulos-padrão e desviantes, como diferenças em frequência, intensidade, duração, localização e diferenças em níveis mais abstratos. Um ponto interessante é a observação de que a MMN ocorre sem que o sujeito preste atenção aos estímulos (revisão em NÄÄTÄNEN; WINKLER, 1999). A resposta MMN geralmente é interpretada como uma manifestação de um processo cortical automático (independente da atenção) de detecção de mudanças em que uma diferença é detectada entre a entrada auditiva atual e a representação dos aspectos regulares das entradas precedentes. De acordo com Näätänen, Jacobsen, e Winkler, (2005) e Näätänen, Tervaniemi, Sussman, Paavilainen e Winkler (2001), a atividade de populações de neurônios no córtex cerebral gera a negatividade observada. Acredita-se que as principais áreas envolvidas sejam o córtex auditivo – a localização exata parece depender da natureza dos estímulos e das diferenças investigadas – e o lobo frontal (revisão em ALHO, 1995). Desta forma, o processo que gera a MMN em um dado momento parece ser baseado em traços de memória formados pelos estímulos auditivos apresentados anteriormente. Näätänen e Winkler (1999) e Näätänen *et al.* (2001) concluem, com base em resultados de vários estudos, que os traços neurais subjacentes às respostas de MMN no córtex auditivo codificam a informação sensorial específica que aparece na percepção auditiva e na memória sensorial, ou seja, o que eles chamam de Representação Central do Som, uma representação integrada e unitária do evento auditivo, disponível cerca de 200 ms após o início do estímulo, combinando as várias características do

estímulo em um objeto perceptivo único e abstraindo invariâncias da variabilidade acústica da estimulação, demonstrando assim um tipo de “inteligência primitiva” no nível do processamento sensorial pré-atencional (ver, por exemplo, AULANKO, HARI, LOUNASMAA, NÄÄTÄNEN; SAMS, 1993; SUSSMAN; RITTER; VAUGHAN, 1998; PAAVILLAINEN; JARAMILLO; NÄÄTÄNEN; WINKLER, 1999; PAAVILAINEN; SIMOLA; JARAMILLO; NÄÄTÄNEN; WINKLER; 2001; JACOBSEN; SCHRÖGER; ALTER, 2004).

Em investigações acerca da categorização dos sons da fala é possível, por meio da observação da MMN, tentar verificar algo como o efeito de limite de fonema comparando as respostas para condições em que os estímulos desviante e padrão pertencem à mesma categoria com as respostas para condições em que, embora a diferença seja fisicamente idêntica à da primeira condição, os estímulos desviante e padrão pertencem a categorias diferentes. Há, entretanto, circunstâncias de natureza prática que reduzem as possibilidades. Por exemplo, enquanto numa tarefa comum de discriminação podemos registrar o desempenho dos sujeitos relativo a um grande número de estímulos diferentes, no paradigma experimental empregado para se obter a MMN apenas um par (ou talvez muito poucos) pode ser observado em cada bloco, que consiste em cerca de mil apresentações.

Aaltonen, Niemi, Nyrke e Tuhkanen (1987) realizaram um estudo para observar a MMN em resposta a vogais. Os estímulos eram as vogais finlandesas /i/ e /y/ e um som intermediário. O par padrão-desviante em cada bloco era formado por dois entre estes três sons. Foram obtidas MMN's quando o par era formado pelas duas vogais puras e quando o par era formado por uma vogal pura e a vogal intermediária, independente da atenção dos sujeitos em relação à estimulação. Mas a MMN observada no primeiro caso era maior em amplitude e apresentava

uma latência menor. Não há como decidir com base nestes resultados se a diferença resulta da diferença acústica maior entre as duas vogais puras, de algum efeito relacionado às categorias fonêmicas análogo ao efeito de limite de fonema ou ao efeito do magneto perceptual, ou de uma combinação de ambos.

Sharma, Kraus, McGee, Carrel e Nicol (1993) conduziram um estudo para testar a hipótese de que o mecanismo gerador da MMN é sensível a contrastes entre categorias fonéticas. Foi realizada uma tarefa de identificação em que os sujeitos deveriam classificar sílabas consoante-vogal (doravante CV) em um *continuum* /da/-/ga/. A partir dos resultados desta tarefa foram concebidas duas condições experimentais. Na primeira um estímulo identificado como /da/ foi apresentado repetidamente como estímulo-padrão e um estímulo diferente também identificado como /da/ ocorria ocasionalmente como estímulo desviante. Na segunda condição o estímulo-padrão foi mantido e o estímulo desviante era uma sílaba CV identificada como /ga/. A magnitude da diferença acústica entre os estímulos padrão e desviante era a mesma nas duas condições. Os sujeitos ignoravam a estimulação durante as sessões. Se a MMN depende de uma representação acústica dos estímulos, espera-se que uma MMN seja encontrada nas duas condições. Se, por outro lado, no caso dos sons da fala, a MMN depende apenas de representações categóricas, uma MMN seria esperada apenas na segunda condição. Uma MMN maior e/ou com menor latência indicaria a presença tanto de representações acústicas como de representações fonéticas. Nas duas condições foram observadas MMN's que não diferiam significativamente entre si, o que levou os autores à conclusão de que, pelo menos no caso de contrastes de ponto de articulação de oclusivas, a MMN reflete apenas representações acústicas. Entretanto, esta conclusão não é inevitável, pois poderia ser que os efeitos dos contrastes em vários níveis de representação não fossem aditivos ou que, ainda que o fossem, a resposta relacionada às representações categóricas seja muito pequena se

comparada à resposta relacionada às representações acústicas, o que ocultaria o efeito das primeiras em razão de uma diferença real mas que não chega a atingir significância estatística.

Outros estudos demonstraram efeitos de categoria fonética nas respostas de MMN. Dehaene-Lambertz (1997) usou estímulos de um *continuum* de /ba/ a /Da/ (uma retroflexa do Hindu) passando por /da/. Foi comparada a MMN em resposta ao contraste inter-categorias /ba/-/da/ com a MMN em resposta aos contrastes intra-categoria /ba/-/ba/ e /da/-/Da/ (um contraste intra-categoria para os sujeitos do experimento, falantes nativos do Francês, mas inter-categoria para falantes nativos do Hindu). Uma MMN maior foi encontrada para os contrastes inter-categoria. Estas observações contrastam com os resultados de Sharma *et al.* (1993), o que pode ter sido causado por diferenças na metodologia. Dehaene-lambertz (1997) utilizou um número maior de eletrodos para realizar as medidas, um paradigma experimental diferente do usual em que são apresentados grupos de três estímulos-padrão seguidos de um desviante, os estímulos variam de grupo para grupo e o sujeito presta atenção à estimulação a fim de detectar mudanças. Seus resultados fornecem evidências de que a MMN reflete o processamento tanto de representações categóricas quanto de representações acústicas.

Sharma e Dorman (1999) obtiveram MMN's em resposta a contrastes entre estímulos num *continuum* de tempo de início de vozeamento (VOT; *voice onset time*). Em uma condição o estímulo-padrão era uma sílaba CV com um VOT de 30 ms e o desviante uma sílaba semelhante com um VOT de 50 ms. A consoante da primeira sílaba foi identificada pelos sujeitos como vozeada (/d/) e a segunda como surda (/t/). Numa segunda condição o estímulo-padrão apresentava um VOT de 60 ms e o desviante um VOT de 80 ms, ambos sendo identificados como sílabas com consoantes surdas. Foram obtidas MMN's nas duas condições, mas a MMN em resposta aos estímulos desviantes da condição inter-categoria

apresentou uma área maior. Adicionalmente, foi observado outro componente dos potenciais evocados, o N1 – componente negativo que ocorre a cerca de 100 ms após o início do estímulo, considerado como uma medida de codificação sensorial. Os sons identificados como /da/ eliciaram um componente N1 único e os sons identificados como /ta/ produziram dois subcomponentes, denominados N1' e N1. A latência destes componentes sugere que o N1' ocorreu em resposta à explosão da consoante oclusiva e que o N1, que é positivamente correlacionado ao VOT, foi produzido pelo início do vozeamento. Os autores concluem então que a categorização de consoantes no parâmetro VOT ocorre em um nível inicial de processamento, no nível sensorial. Coerente com esta afirmação é o estudo de Sinex e McDonald (1989), em que foi registrada a atividade de fibras neurais individuais do nervo auditivo de chinchilas anestesiadas em resposta a sílabas de um *continuum* de VOT's de 0 a 80 ms. Estímulos com VOT's de 0 a 20 ms provocaram um aumento na taxa de descarga 20 ms após a apresentação do estímulo. Estímulos com VOT's de 30 a 80 ms provocaram um aumento na taxa de descarga que coincidia no tempo com o início do vozeamento. Estes resultados proporcionam uma interessante ilustração sobre como descontinuidades características da organização do sistema nervoso já em níveis muito baixos de processamento da informação sensorial podem ser utilizadas pelo ser humano no estabelecimento de categorias fonéticas.

Ainda sobre categorizações fonêmicas no *continuum* de VOT, além dessas descontinuidades inatas evidenciadas nos resultados citados no parágrafo anterior, parece haver também descontinuidades perceptivas manifestas nas respostas de MMN que são estabelecidas com a experiência lingüística. Sharma e Dorman (2000) produziram, a partir de sons produzidos naturalmente, um *continuum* de sílabas CV (com consoantes oclusivas bilabiais) com VOT's que variavam entre - 90 e 0 ms. Para os falantes do inglês trata-se de um *continuum* de sons

de uma mesma categoria, mas os falantes do Hindu o dividem em duas categorias, sendo o limite entre elas localizada perto do valor de VOT de - 30 ms. Falantes nativos do inglês apresentaram um desempenho no nível aleatório na discriminação entre os estímulos, ao passo que falantes do Hindu apresentaram um desempenho muito bom quando os estímulos comparados pertenciam a categorias diferentes. Num experimento eletrofisiológico, apenas para os falantes do Hindu foi observada uma MMN robusta e estatisticamente significativa em resposta ao estímulo desviante – que era uma sílaba identificada como não pertencente à categoria do estímulo padrão. As mudanças no componente N1, que refletiram a duração do pré-vozeamento, não diferiram entre os dois grupos de falantes.

Näätänen *et al.* (1997) testaram falantes nativos do Finlandês e falantes nativos do Estoniano em um estudo de MMN em que todas as condições tinham como estímulo padrão a vogal /e/. Variando os valores do segundo formante foram produzidos os estímulos desviantes, as vogais /ö/, /õ/ e /o/. Os sons /e/, /ö/ e /o/ são vogais de ambas as línguas, e /õ/ é uma vogal do Estoniano. O valor de F2 desta se localiza entre os valores de F2 das vogais /ö/ e /o/. No grupo de falantes do Finlandês embora tenham sido observadas MMN's para todos os desviantes, amplitudes menores foram observadas para o desviante /õ/, mesmo sendo o desvio acústico desta vogal em relação à vogal /e/ maior do que o da vogal /ö/. No grupo dos falantes do Estoniano todos os desviantes produziram MMN's com amplitudes semelhantes. Por meio de medidas de MMNm – a versão da MMN obtida por magnetoencefalografia (MEG) – foi demonstrado que o componente da MMN relacionado ao contraste fonêmico originou-se no córtex auditivo esquerdo e o componente relacionado ao contraste acústico tem sua fonte nos córtices auditivos dos dois hemisférios. Isto sugere que um processo de detecção de mudanças acústicas que ocorre bilateralmente e um processo relacionado às representações de categorias fonêmicas lateralizado à esquerda contribuem para a MMN obtida com vogais.

Adicionalmente, os resultados fornecem evidências de que traços de memória representando fonemas são formados com a experiência lingüística no córtex auditivo esquerdo. Tervaniemi *et al.* (2000), num estudo de neuroimagem, fornecem mais evidências sobre a localização das populações neuronais que geram a MMN em resposta a contrastes entre sons da fala. Por meio de tomografia por emissão de pósitrons (PET), foi observada uma atividade diferencial refletindo a MMN nos giros temporal superior e temporal medial do córtex auditivo esquerdo quando as vogais /e/ e /o/ eram apresentadas como estímulo-padrão e desviante, respectivamente. Quando os estímulos padrão e desviante eram os acordes A maior e A menor, a atividade relacionada à MMN foi observada no giro temporal superior do córtex auditivo direito.

No plano estabelecido por dois eixos representando os formantes F1 e F2, a vogal húngara /ɛ/ se sobrepõe às vogais finlandesas /e/ e /æ/ e a vogal finlandesa /e/ se sobrepõe às vogais húngaras /ɛ/ e /é/. Isto tornou possível para Winkler *et al.* (1999a) elaborarem um experimento em que foram usados um par de vogais que é inter-categoria em Húngaro e intra-categoria em Finlandês e um par que é inter-categoria em Finlandês e intra-categoria em Húngaro. Na “seqüência húngara” os estímulos desviante e padrão eram os membros do primeiro par e na “seqüência finlandesa” eram os membros do segundo par. Com fins de controle foi incluído um segundo estímulo desviante nas duas seqüências, a vogal /y/. O estímulo desviante principal ocorria com uma probabilidade de 0,15 em cada seqüência e a vogal /y/ com uma probabilidade de 0,025. As duas seqüências foram apresentadas a um grupo de falantes nativos do Finlandês e a um grupo de falantes nativos do Húngaro. Em ambos os grupos a vogal /y/ eliciou uma MMN significativa. Excluindo-se o efeito da vogal /y/, no grupo húngaro foram obtidas MMN’s significativas em ambas as seqüências, mas a

MMN observada na “seqüência húngara” (contraste inter-categoria) apresentou uma amplitude maior e uma latência menor. No grupo finlandês uma MMN significativa foi observada para a “seqüência húngara” (intra-categoria) e uma MMN maior apresentando dois picos sucessivos (o que pode indicar dois processos diferentes em operação) foi observada para a seqüência finlandesa (inter-categoria). Este trabalho forneceu evidências adicionais favoráveis às conclusões de Näätänen *et al.* (1997) em relação ao emprego tanto de representações acústicas quanto de representações categóricas na detecção pré-atencional de contrastes entre sons vocálicos. Em outro estudo, Winkler *et al.* (1999b) não observaram MMN em resposta à vogal desviante /ä/ que ocorria ocasionalmente em meio a uma seqüência de estímulos-padrão (/e/) em um grupo de sujeitos húngaros que não sabiam falar Finlandês. Já sujeitos húngaros que viveram anos na Finlândia e aprenderam a falar finlandês fluentemente apresentaram uma MMN semelhante à apresentada por sujeitos finlandeses em resposta ao mesmo contraste.

Dehaene-Lambertz, Pallir, Serniclaes, Sprenger-Charolles, Jobert e Dehaene (2004) realizaram um estudo de neuroimagem e MMN em que os sujeitos respondiam a provas compostas por quatro sons análogos a sílabas CV formados por ondas senoidais simples. De acordo com os autores, estes sons são ouvidos normalmente como zunidos eletrônicos pelos sujeitos, mas se for explicado que se tratam de imitações de sons da fala, os sujeitos os ouvem como sílabas CV. A tarefa era simplesmente responder em cada prova se há ou não uma diferença no último som enquanto registros eletroencefalográficos ou de ressonância magnética funcional eram efetuados. Os sujeitos foram testados antes e depois de serem informados a respeito do aspecto de “som da fala” dos estímulos. Assim, os autores tentaram fazer com que os sujeitos passassem de um “modo acústico” para um “modo de fala”. No “modo de fala”, em comparação com o “modo acústico” foram observados uma diferença

maior em relação à latência e à amplitude da MMN entre os contrastes intra- e entre-categorias (sendo a resposta correspondente ao último maior e mais rápida) e um aumento na atividade do sulco temporal superior esquerdo. Os resultados foram interpretados como evidências de que a percepção de fonemas é mediada por uma rede especializada no córtex cerebral esquerdo correspondente a um modo de processamento de informações relativas aos sons da fala.

Nos estudos de MMN, a forma mais comum de isolar o componente “não acústico” dos contrastes entre os sons da fala é comparar as respostas para condições em que os estímulos desviante e padrão pertencem à mesma categoria com as respostas para condições em que, embora a diferença seja fisicamente idêntica à da primeira condição, os estímulos desviante e padrão pertencem a categorias diferentes – as categorias são estabelecidas em uma tarefa de identificação de fonemas, em que os sujeitos classificam os sons de um *continuum* acústico entre pontos extremos das categorias de interesse. Se a MMN depende de uma representação acústica dos estímulos, espera-se que uma MMN seja encontrada nas duas condições. Se, por outro lado, no caso dos sons da fala, a MMN depende apenas de representações categóricas, uma MMN seria esperada apenas na segunda condição. O resultado mais comumente encontrado, entretanto, é uma MMN maior e/ou com menor latência na segunda condição, o que indica a presença tanto de representações acústicas quanto de representações categóricas (fonéticas e/ou fonológicas).

Phillips (2001), Phillips *et al.* (2000) e Phillips *et al.* (1995) argumentam que a partir de resultados obtidos por meio deste método, podemos inferir com segurança a existência de representações no nível fonético, mas não no nível fonológico, para o qual são necessárias evidências adicionais. Nesta linha argumentativa, representações fonológicas têm natureza

simbólica; são as unidades discretas que são combinadas nas formas lexicais. Enquanto no nível fonético as representações são gradientes e apresentam estrutura interna (ver, por exemplo, KUHL, 1991), sendo os exemplares agrupados dentro de cada categoria distintos entre si, no nível fonológico, dado o caráter discreto das representações, distinções entre membros de uma mesma categoria são irrelevantes.

Phillips *et al.* (1995), em um estudo de MEG (magnetoencefalografia), elaboraram um método que torna possível o acesso a representações fonológicas, isto é, unidades discretas de processamento que representam categorias fonológicas. Neste método, ao invés de um estímulo padrão e um estímulo desviante, há um conjunto de estímulos padrão e um conjunto de estímulos desviantes, dentro dos quais o parâmetro acústico de interesse varia em passos fixos, de forma que o que delimita os dois conjuntos não seja o parâmetro acústico em si, mas o limite entre as duas categorias da fala cujo contraste se pretende investigar. A relação “muitos-para-poucos” que define os conjuntos como “desviante” e “padrão” é estabelecida numa condição experimental em que a maior parte dos estímulos apresentados pertence a uma das categorias. Deste modo, uma resposta de MMN não poderia ser atribuída à variação acústica, pois esta ocorre entre os próprios estímulos-padrão. Phillips *et al.* (2000) observaram uma MMNm para o contraste entre consoantes oclusivas alveolares vozeadas e surdas.

As evidências aqui apresentadas de que a MMN em resposta a um contraste entre-categorias é em parte causada por um processo que envolve representações de categorias de sons da fala na memória esclarecem algumas questões relacionadas aos estudos comportamentais de percepção categórica. Em primeiro lugar, os resultados contrariam a tese de que o efeito de limite de categoria observado nos estudos em que são comparados os desempenhos dos sujeitos em tarefas de discriminação e de classificação tem origem não em processos

perceptivos, mas no nível da decisão, uma vez que a MMN é uma resposta relacionada a mecanismos pré-atencionais de discriminação. Em segundo lugar, verifica-se que na discriminação de sons da fala ocorrem comparações envolvendo traços na memória auditiva relacionadas às propriedades acústicas do sinal e comparações em que são empregadas representações de categorias dos sons da fala. Deste modo, é natural que nos estudos comportamentais de percepção categórica sejam observados desempenhos na discriminação superiores à previsão feita a partir dos resultados na tarefa de classificação. Outra observação interessante é a de que algumas categorias (provavelmente a maior parte) são estabelecidas durante a aprendizagem da língua (NÄÄTÄNEN *et al.*, 1997; WINKLER *et al.*, 1999a; WINKLER *et al.*, 1999b) e outras refletem descontinuidades inerentes ao sistema sensorial (SINEX; MCDONALD, 1989).

2.3 Esclarecimentos adicionais

Antes de seguir para a descrição metodológica do estudo é importante que alguns conceitos sejam esclarecidos.

Este estudo é motivado por uma observação referente ao conjunto de vogais do Português Brasileiro, qual seja, o fato de que ao mesmo tempo em que são atestados pares mínimos que sustentam a funcionalidade do contraste entre vogais médias abertas e vogais médias fechadas, verifica-se a ocorrência de variação entre as duas categorias, tanto no caso das vogais anteriores como no caso das vogais posteriores, no mesmo ambiente em que se observa o contraste em questão.

Da perspectiva de teorias de orientação formalista, em que regras ou restrições operam sobre representações simbólicas discretas (neste caso, obviamente, o conceito de “fonema” tem especial relevância), seria desejável que se busquem critérios que facilitem a escolha entre um sistema com sete e um sistema com cinco vogais de modo a minimizar os custos de, por um lado, conviver com a presença de formas como [estrə]/[ɛstrə] e [posə]/[pɔsə] ou, por outro lado, de retirar do contraste entre as médias abertas e fechadas o estatuto fonológico. Entretanto, evidências obtidas em estudos experimentais como o descrito neste texto dificilmente podem ser integradas a esse tipo de teoria. No caso das mesmas, o ideal seria alguma formulação como, por exemplo, de uma demonstração em que os segmentos, que aparentemente variam de maneira livre, ocorrem de modo previsível.

Teorias de orientação funcionalista são bem mais receptivas a resultados de pesquisas experimentais ou empíricas de um modo geral. Modelos e evidências sobre a percepção da fala desempenham um papel de extrema importância na formulação de teorias como a Fonologia de Uso (BYBEE, 2001) e a Fonologia Funcional de Boersma (1998). Vale notar que as observações em questão – isto é, aquelas relativas à concomitante ocorrência de pares mínimos e de variação entre vogais médias abertas e fechadas na sílaba tônica – não são incômodas para as teorias funcionalistas como para as teorias formalistas. A questão da variação se apresenta naquelas de forma natural, dada à incorporação de elementos probabilísticos e relacionados ao conceito de “gradiência” e a tese de que as unidades de armazenamento e processamento lingüístico são semelhantes às unidades verificadas nos outros domínios cognitivos.

Como foi salientado na introdução, embora este estudo seja motivado pelo fenômeno citado, não se pretende com o mesmo defender uma ou outra posição em relação à questão sobre a

adequação do sistema vocálico composto por sete ou por cinco elementos. O que se pretende investigar é simplesmente a possibilidade de que o contraste entre [o] e [ɔ] seja menos saliente que o contraste entre [o] e [u] no sistema perceptivo. Isto é, se o menor grau de consistência do primeiro contraste – em relação ao segundo contraste – verificado nas ocorrências lingüísticas se reflete na percepção (e no armazenamento). Muito dificilmente, as evidências obtidas nesse trabalho indicarão categoricamente em favor de um conjunto de cinco ou de sete vogais, pelo simples fato de que trabalhos empíricos dependem das condições de realização e devem ser interpretados à sua luz. Todavia, é interessante observar que se for o caso que, no caso de se verificar no experimento um contraste menos distinto entre as duas vogais médias em questão, pelo menos três hipóteses sobre a direção da causalidade podem ser consideradas. A permutação entre [o] e [ɔ] na sílaba tônica sem conseqüências semânticas poderia ser de alguma forma resultado de um limite relativamente pouco distinto entre as duas categorias armazenadas na memória e usadas na identificação do som. Alternativamente, um contraste menos pronunciado entre [o] e [ɔ] no sistema perceptivo poderia ter origem na experiência lingüística no que se refere aos padrões de ocorrência das vogais na língua (incluindo as ocorrências nas sílabas átonas). Se duas categorias de sons da fala numa língua ocorrem nos mesmos contextos, sem com isso estabelecerem diferenças de significado, é de se esperar que o falante desenvolva representações destas categorias menos distintas entre si. A terceira possibilidade é de que ambas sejam corretas. Poderia ser ainda que nenhuma delas fosse correta.

Inferências a respeito da organização das vogais na mente do falante serão realizadas a partir de medidas efetuadas sobre o desempenho dos sujeitos em certas tarefas de modo a estabelecer relações entre estímulos e respostas. Os objetos da investigação são categorias e limites de categorias dos sons da fala entendidos como entidades que se fazem presentes nos

processos perceptivos. Torna-se assim importante que sejam colocados em pauta alguns conceitos intimamente relacionados à idéia de que o desempenho dos sujeitos nas tarefas aqui empregadas nos informa sobre a organização do sistema perceptivo do falante do Português Brasileiro no que diz respeito aos processos de identificação, classificação e discriminação das vogais. Como o interesse se concentra sobre categorias armazenadas na memória representando sons (ou grupos de sons) da fala, que se manifestariam nos padrões observados nos dados do experimento, e os resultados serão interpretados em termos de representações na memória de longo prazo, um ponto a ser considerado é o conceito de representação.

A maior parte das teorias das ciências cognitivas admitem que nós seres humanos possuímos representações do mundo em que vivemos, sendo estas estados do sistema cognitivo que se referem a estados do ambiente e que são processados pelo mesmo sistema cognitivo. Por exemplo, quando uma pessoa percebe o som de uma campainha, considera-se que padrões de atividade neural desencadeados por este estímulo compõem uma representação do mesmo. Esta atividade promove modificações na estrutura das redes neurais de modo a impor restrições sobre fluxos de ativação provocados por estimulações posteriores. Em teorias simbólicas (como a gramática gerativa), representações e processos são abstraídos do substrato orgânico de tal forma que o funcionamento do sistema é descrito em termos de símbolos (unidades discretas e passivas de representação) e operações sobre estes símbolos.

Em relação à epistemologia, na versão mais simples as representações são como cópias das coisas que representam, de forma que a percepção proporcione um acesso direto ao ambiente. De acordo com uma versão um pouco mais sofisticada, a representação é uma função do que se apresenta no ambiente, isto é, a informação é de alguma forma transformada. A relação entre representação e coisa representada é também neste caso meramente referencial. Outras

concepções consideram que a representação é uma função do ambiente e do próprio sistema cognitivo. Este desempenha um papel ativo na formação das representações, formação esta que depende de disposições do organismo estabelecidas na interação do indivíduo (ou da espécie) com o ambiente.

A noção de representação é hoje alvo de duras críticas baseadas na argumentação de que os seres vivos não mantêm interações instrutivas com o meio (por exemplo, MATURANA; VARELA, 2001). Isto é, um evento no ambiente nunca determina mudanças estruturais num organismo, podendo apenas estabelecer perturbações na superfície sensorial que desencadeiam alterações que são determinadas pela configuração da própria estrutura. Desta forma, não há possibilidade de um mapeamento entre estados internos e estados externos ao organismo, sendo, portanto, o conceito de representação baseado numa noção equivocada. Contudo, mesmo tendo como ponto de partida este fechamento operacional dos organismos, de acordo com Peschl e Riegler (1999), é possível justificar o emprego do conceito, desde que se renuncie ao aspecto de referência a um mundo externo, tratando-se apenas de padrões de atividade neural que se referem a padrões de atividade neural. Isto reduz a influência da estimulação provinda do ambiente à modulação de dinâmicas representacionais internas.

Winkler (1999) argumenta a respeito da questão estabelecida pelo aspecto de fechamento operacional dos organismos que embora estes sejam fechados no nível de descrição das suas operações, são abertos no nível de descrição das trocas de matéria e energia. Segundo Winkler (1999), o fenômeno representado é o padrão espaço-temporal da interação entre o sistema cognitivo e seu ambiente, e não o objeto em si, e o conceito de representação deve incluir uma relação de significado, o que implica em uma referência ao ambiente externo.

Neste trabalho, admite-se que objetos no exterior do organismo promovem perturbações nas superfícies sensoriais e que estas por sua vez causam padrões de atividade neural que modificam as estruturas em que ocorrem, e que regularidades estatísticas na atividade e na estrutura de redes neurais correspondem a regularidades estatísticas no ambiente em vários níveis de abstração. Portanto, representações são entendidas aqui como construtos mentais correspondentes a padrões no interior do organismo que mapeiam padrões ocorrentes no meio. Note-se que disto depende a justificativa do uso de medidas relacionando estímulo e resposta na investigação sobre o funcionamento e a organização de representações sonoras.

Tendo esclarecido o que é denotado aqui pelo termo “representação”, o próximo passo é um breve exame deste conceito aplicado aos objetos de interesse do estudo, quais sejam, as categorias que agrupam sons correspondentes às vogais médias posteriores e suas delimitações. Neste contexto, o “fonema” se apresenta como uma representação mental que se refere a uma classe de sons funcionalmente equivalentes. De acordo com a fonologia gerativa (CHOMSKY; HALLE, 1968), trata-se de um feixe de valores de traços distintivos que define a unidade sonora simbólica, em oposição aos sons como entidades concretas. Neste caso, o fonema, enquanto unidade simbólica, é destituído de conteúdo relativo a propriedades acústicas. No âmbito da competência lingüística, no que concerne aos segmentos, há apenas fonemas e não fones ou detalhamento fonético de qualquer espécie. Cadeias lineares de fonemas seriam então intenções do falante e impressões do ouvinte; as representações que estabelecem a interface entre a entrada acústica e o léxico. Desta perspectiva não só fazem sentido como se tornam necessárias considerações a respeito do estatuto fonológico do contraste entre as vogais médias fechadas e as vogais médias abertas, isto é, se /o/ e /ɔ/ (ou /e/ e /ɛ/) são realmente dois fonemas diferentes. Formulada a questão desta maneira, é discutível se evidências experimentais podem ou não proporcionar alguma contribuição.

Mesmo que existam efetivamente, não há como observar diretamente construtos mentais (como fonemas, imagens mentais, conceitos, etc.). Mesmo nas neurociências o que se pode observar são localizações e seqüências de atividade neural, e não construtos mentais. Em estudos como o apresentado aqui não há nada que justifique de forma inequívoca a afirmativa de que certo resultado que se venha a observar reflita a organização das categorias no nível de representação dos fonemas. Suponhamos um caso extremo em que nenhum limite entre as categorias /o/ e /ɔ/ se manifeste nos resultados da tarefa de classificação. Ainda assim seria possível argumentar, por exemplo, com base em noções gerativistas, que não há nenhuma garantia de que as respostas dos sujeitos reflitam a organização no nível fonológico (isto é, um nível de representações discretas e abstratas) e não apenas num nível fonético (superficial, contínuo e concreto).

Sendo a existência do fonema como entidade mental presumida nas teorias estruturalistas e gerativistas – definições puramente operacionais também foram propostas (ver, por exemplo, TWADELL, 1935) – é possível questionar o conceito. Port (no prelo) argumenta que a tese de que as palavras são formadas por segmentos estáticos, discretos e organizados em série, não se sustenta se forem considerados os dados de investigações sobre percepção e produção da fala. Fones e fonemas são estáticos, discretos e ordenados de forma serial e, portanto, não seriam adequados na descrição científica dos processos contínuos e sobrepostos característicos da fala, tanto no que se refere à articulação quanto ao sinal. No que concerne à percepção, as unidades de representação seriam quaisquer padrões auditivos, em várias escalas (de traços a sílabas), que tenham suficiente predominância estatística na fala. Port (no prelo) conclui que fones e fonemas são conceitos atrativos porque correspondem às intuições de pessoas com uma longa experiência de uso do alfabeto e não necessariamente a um fato psicológico.

Jaeger (1980 *apud* OHALA 1983) empregou uma técnica experimental da psicologia para o estudo da formação de conceitos para verificar se um grupo de falantes do inglês formava uma categoria que incluísse os alofones [k] e [k^h]. Numa primeira etapa era explicado ao sujeito que ele ouviria várias palavras e que algumas pertenceriam a certa categoria conforme o som das mesmas. A tarefa era indicar se cada som pertence ou não à categoria – sobre a qual não foram dadas indicações. Apenas um *feedback* era dado após cada resposta indicando se ela estava correta ou incorreta. O conjunto de estímulos correspondente à categoria correta era formado por palavras do inglês com [k] ou o [k^h] como fone inicial. Na tentativa de anular a influência da ortografia, palavras não iniciadas com [k] ou [k^h] e cuja forma escrita tem *k*, *c*, *ch* e *qu* como grafemas iniciais foram incluídas no grupo de palavras externas à categoria de interesse no experimento. O fone [k^h] não foi apresentado nesta etapa inicial. Na segunda etapa o *feedback* foi suspenso e foram incluídos exemplares do fone [k^h]. Os sujeitos consistentemente agruparam os fones [k] e [k^h] na mesma categoria. O resultado foi interpretado pela autora como evidência da realidade psicológica do fonema.

Outra questão que pode ser levantada é se o fonema é ou não a unidade perceptiva no processamento da fala. Convém salientar que esta é uma questão que tem mais importância em teorias que concebem as representações como símbolos. Em teorias que enfatizam processos subsimbólicos e não se comprometem com o pressuposto de que a mente processa unidades discretas, a questão se torna secundária ou carente de sentido. Baseados em evidências empíricas, alguns defendem a tese de que o fonema não é a unidade da percepção da fala (MASSARO, 1972; LOTTO; HOLT, 2000; PORT, no prelo), outros que defendem que sim (NEAREY, 1997; 2003). Fitzpatrick e Wheeldon (2000) apresenta um modelo de acesso lexical em que as unidades que estabelecem a mediação entre o sinal acústico e o

léxico são traços. Neste modelo, apenas após o acesso ao léxico são realizadas análises em sílabas e fonemas. Foss e Swinney (1973), em um experimento em que os sujeitos se engajavam em tarefas de monitoramento de fonemas e de sílabas cujos estímulos eram palavras dissílabas, observaram respostas mais rápidas na detecção de sílabas que na detecção de fonemas. Segundo McQueen e Cutler (1997), estes resultados indicam que a sílaba é a unidade básica da percepção da fala, mas os próprios autores (FOSS; SWINNEY,1973) discordam. No mesmo estudo foi verificado que a detecção de fonemas também é mais lenta que a detecção de palavras. De acordo com a interpretação dos autores, deve-se fazer uma distinção entre percepção e identificação, sendo a última relacionada à manifestação do estímulo na consciência. Talvez relutantes em considerar a possibilidade de que o reconhecimento de palavras ocorra em etapas anteriores do processamento em relação ao reconhecimento de fonemas, eles concluem que os resultados estariam refletindo a ordem da identificação e não da percepção.

No próximo capítulo serão descritos os experimentos por meio dos quais se busca, nesta investigação, elementos que proporcionem inferências a respeito da forma como o contraste entre as categorias [o] e [ɔ] – em comparação com o contraste entre [o] e [u] – se dá no nível das representações armazenadas na memória de longo prazo.

3 Metodologia

3.1 Sujeitos

Participaram deste estudo doze falantes nativos do Português Brasileiro (dialeto falado na região de Belo Horizonte), seis homens e seis mulheres, com idades entre 18 e 27 anos, nenhum dos quais relata problemas auditivos. Todas as tarefas foram realizadas por cada sujeito.

3.2 Estímulos

Para compor o *continuum* em F1 empregado neste estudo, 42 vogais foram sintetizadas por meio do programa Praat 4.3.12 (BOERSMA; WEENINK, 2005). A frequência do formante F1 variou em passos de 10 Hz entre as frequências 250 Hz e 660 Hz. As frequências dos formantes F2, F3, F4, ..., F10 foram fixadas, respectivamente, nos valores 900 Hz, 2500 Hz, 3500 Hz, 4500 Hz, 5500 Hz, 6500 Hz, 7500 Hz, 8500 Hz e 9500 Hz. As larguras de banda utilizadas para F1, F2, F3, ..., F10 são, respectivamente, 50 Hz, 100 Hz, 150 Hz, 200 Hz, 300 Hz, 400 Hz, 500 Hz, 600 Hz, 700 Hz e 800 Hz.

A fonte a partir da qual os sons do *continuum* foram criados é um sinal gerado de forma a simular um som produzido por uma série de pulsos glotais. A forma da curva que descreve o fluxo glotal em função do tempo (normalizado) foi estabelecida pelo ajuste *default* do programa Praat, a partir da equação

$$U(x) = x^3 - x^4$$

onde o eixo-x representa o tempo normalizado (variando entre 0 e 1) e $U(x)$ é o fluxo glotal normalizado em unidades arbitrárias. A duração do sinal da fonte é de 300 ms, durante os quais a frequência fundamental decai linearmente de 145 a 85 Hz, e a intensidade do sinal varia de 0 a 75dB (SPL) no intervalo entre 0 e 25 ms, se mantém até os 275 ms e decai então até 0 dB (SPL) nos 300 ms.

Para gerar cada som do *continuum*, essa fonte é então filtrada pela série de formantes descrita acima de forma a se obter sons semelhantes a vogais emitidas com uma voz masculina, com frequência fundamental descendente, a uma intensidade de cerca de 75 dB (SPL), por um intervalo de 300 ms.

3.3 Delineamento experimental

3.3.1 Tarefa de Classificação

Cada uma das 42 vogais do *continuum* em F1 foi apresentada vinte vezes em ordem pseudoaleatória (não foram permitidos estímulos idênticos sucessivos), o que resulta em 840 provas. Esta tarefa foi realizada em duas etapas de 420 provas com o mesmo número de ocorrências de cada som, isto é, dez. Em cada prova os sujeitos classificaram cada estímulo como “u”, “o” ou “ɔ”. Um intervalo de 1 s ocorre entre a resposta do sujeito e a apresentação do próximo som.

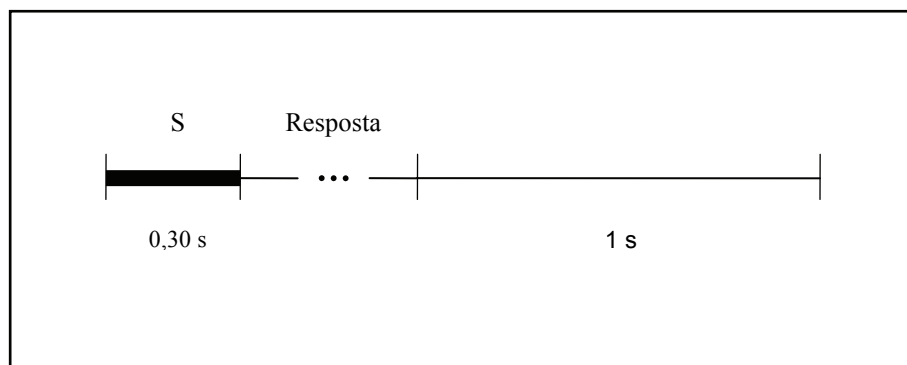


FIGURA 1 - Esquema de uma prova na tarefa de classificação, se iniciando com a apresentação do som S e terminando com um intervalo de 1s após a resposta.

3.3.2 Tarefa de discriminação 2AFC (escolha forçada com dois intervalos e duas alternativas)

Em cada prova é apresentado um estímulo composto por um par de sons. A diferença entre estes sons é fixa em 30 Hz. A cada ponto no *continuum* corresponde um par de sons. Cada par foi apresentado doze vezes. Na metade das vezes, o som com o F1 mais alto está no primeiro intervalo e, na outra metade, no segundo intervalo. Daqui em diante um par de sons em uma tarefa de discriminação será denotado pelo valor de F1 mais baixo do par. Assim, por exemplo, o par composto pelos sons com a frequência de F1 em 400 Hz e 430 Hz será denotado por “400”, independente da ordem de ocorrência dos sons.

Entre os dois sons em cada prova há um intervalo de 150 ms e entre a resposta e o início da prova seguinte, com a apresentação do próximo par de sons, ocorre um intervalo de 1s. A ordem dos estímulos é pseudoaleatória (não ocorrem estímulos idênticos em provas sucessivas). O sujeito deve escolher uma entre duas alternativas de resposta que indique a ordem de ocorrência dos estímulos em relação ao valor de F1. Não há qualquer tipo de

menção a este parâmetro ou à direção da diferença nas instruções ou nas alternativas de resposta. Estas se referem a ordens de ocorrência de categorias vocálicas.

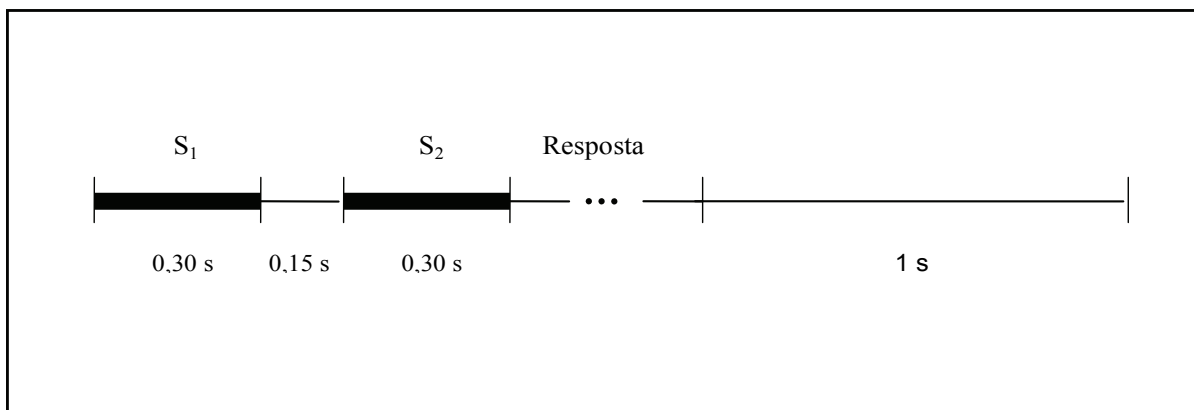


FIGURA 2 - Esquema de uma prova da tarefa de discriminação 2AFC.

Para esta tarefa, o continuum em F1 foi dividido em dois *continua* menores compostos por 21 sons, o primeiro entre 250 e 470 Hz e o segundo entre 440 e 660 Hz. Assim, para cada *continuum* temos vinte pares de sons. Na tarefa em que é empregado o primeiro *continuum*, o sujeito deve responder em cada prova se o par de sons apresentado está na ordem [u-o] ou [o-u]. Na tarefa em que é empregado o segundo *continuum*, o sujeito deve responder se os dois sons ocorreram nas ordens [o-ɔ] ou [ɔ-o]. Isto foi feito em razão da natureza das alternativas de resposta. O *continuum* de 250 a 660 Hz foi gerado para compreender três categorias de vogais, e as alternativas de resposta na tarefa 2AFC se referem necessariamente a ordens de ocorrência de duas categorias. Por exemplo, seria difícil para um sujeito entender que “u-ɔ” é a resposta correta para o par composto pelos sons com F1 de 430 Hz e 460 Hz, nesta ordem, pois estes sons estão localizados no *continuum* na região em que a resposta mais comum é “o”.

Esta parte do experimento é composta então por duas tarefas 2AFC, uma para cada *continuum*. Cada par de sons ocorre seis vezes na ordem ascendente (em F1) e seis vezes na

ordem descendente. Como cada tarefa compreende vinte pares de sons apresentados doze vezes, ao todo são 240 provas.

3.3.3 Tarefa de discriminação 4I2AFC (escolha forçada com quatro intervalos e duas alternativas)

Nesta tarefa o *continuum* de 250 a 660 Hz foi dividido em dois da mesma forma que na tarefa 2AFC descrita anteriormente. Foram também usados pares de sons com uma diferença fixa de 30 Hz.

Em cada estímulo, dois sons flanqueiam um par, isto é, um som ocorre anteriormente ao par e outro posteriormente. Estes sons são idênticos entre si e podem ser idênticos ao primeiro ou ao segundo membros do par. Com isto, tomado um par composto pelos sons A e B, temos os estímulos nos padrões AABA, ABAA, BABB ou BBAB. Cada par ocorre 24 vezes, seis vezes em cada um dos padrões. Desta forma, na metade das vezes o som com o F1 mais alto está no segundo intervalo e, na outra metade, no terceiro. Há um lapso de 150 ms entre o primeiro e o segundo e entre o terceiro e o quarto sons. Entre o segundo e o terceiro sons há um lapso maior, de 250 ms, de modo que os quatro sons sejam percebidos em pares. A ordem dos estímulos é pseudoaleatória (não ocorrem estímulos idênticos em provas sucessivas). O sujeito deve responder se o som diferente ocorreu no segundo ou no terceiro intervalo. Assim como nas demais tarefas, há um intervalo de 1s entre a resposta e a apresentação do primeiro som da prova seguinte.

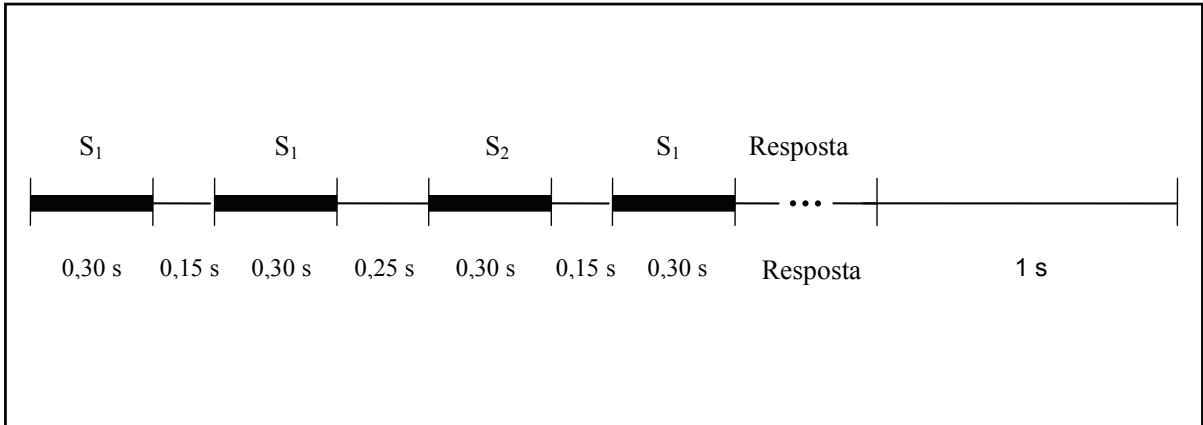


FIGURA 3 - Esquema de uma prova da tarefa de discriminação 4I2AFC em que o som diferente ocorre no terceiro intervalo.

Verificou-se em testes piloto que a tarefa se tornaria muito dispendiosa para os sujeitos se fossem empregados, tal como na tarefa 2AFC, todos os pares de sons possíveis no *continuum* (com a diferença fixa de 30 Hz). Por limitações de tempo não foi possível realizar esta tarefa em duas etapas. Decidiu-se então utilizar apenas a metade dos pares possíveis. Para isto, a distância entre dois pares foi aumentada de 10 Hz – na tarefa 2AFC – para 20 Hz na tarefa 4I2AFC. A diferença de 30 Hz entre os sons de cada par foi mantida. Os sons utilizados variam, no primeiro *continuum*, de 260 a 470 Hz e, no segundo *continuum*, de 440 a 650 Hz. Desta forma, cada tarefa 4I2AFC é composta por 240 provas.

3.4 Procedimentos

Todas as tarefas foram realizadas em uma câmara com isolamento acústico. Os estímulos foram apresentados por meio de um fone de ouvido conectado a um microcomputador. Os sujeitos responderam indicando com o *mouse* um dos campos de resposta presentes na tela do monitor.

Os sujeitos foram testados individualmente. Imediatamente antes de cada tarefa de discriminação foi dada uma versão reduzida da mesma, como treinamento. Em uma primeira seção, cada sujeito executou primeiramente as duas tarefas de discriminação 2AFC. A ordem das mesmas foi contrabalançada entre os sujeitos e deu-se um intervalo de vinte minutos entre as duas. Vinte minutos após o término da segunda tarefa 2AFC, foi dada a primeira etapa da tarefa de classificação e vinte e cinco minutos após esta foi dada a segunda etapa. Em uma segunda seção, duas semanas após a primeira, cada sujeito foi testado nas duas tarefas 4I2AFC. Como na primeira seção, a ordem das tarefas foi contrabalançada e foi dado um intervalo de vinte minutos entre as mesmas.

Nas tarefas 2AFC, os dois campos de resposta no centro da tela apresentam as opções “uo” e “ou” ou “oO” e “Oo”. Foi esclarecido aos sujeitos que “uo” e “ou” representam, respectivamente, os estímulos [u-o] ou [o-u] e “oO” e “Oo” representam os estímulos [o-ɔ] e [ɔ-o]. A cada oitenta provas o sujeito pode, se quiser, fazer uma breve pausa sem remover os fones ou se deslocar de sua posição em frente ao computador. Para a versão de treinamento desta tarefa foram tomados pares dentro dos quais e entre os quais a diferença é de 30 Hz, o que resulta em sete pares de sons. Cada um destes é repetido quatro vezes, duas em cada ordem. Logo antes do início do treinamento de cada uma das tarefas 2AFC foram dadas as seguintes instruções:

“Serão apresentados alguns pares de sons. Indique com o mouse se o que você ouviu se assemelha mais a [u o]’ ou [o u]’ (ou “[o ɔ]’ ou [ɔ o]’ ”). Ouça os dois sons antes de responder. O tempo da resposta não é importante.”

Na tarefa de classificação, nos três campos de resposta apresentados no centro da tela do monitor estão as letras “u”, “o” e “O”, nesta ordem. Foi esclarecido aos sujeitos que “o” representa a categoria [o] e “O” representa a categoria [ɔ]. Após cada 140 provas é permitida uma pequena pausa da mesma forma que na tarefa 2AFC. Logo antes do início da primeira etapa da tarefa de classificação foram dadas as seguintes instruções:

“Serão apresentados alguns sons que você deverá identificar como ‘[u]’, ‘[o]’ ou ‘[ɔ]’. Em cada tentativa aperte o botão da vogal que melhor corresponde ao som que você ouviu. O tempo da resposta não é importante.”

Os dois campos de resposta na tarefa 4I2AFC contêm os tipos “2^o” e “3^o”. A cada oitenta provas o sujeito pode fazer uma pequena pausa como nas demais tarefas. Para a versão de treinamento desta tarefa foram tomados pares dentro dos quais a diferença é de 30 Hz e entre os quais a diferença é de 60 Hz. Isto resulta em quatro pares de sons, cada um dos quais sendo repetido oito vezes, duas em cada padrão (AABA, ABAA, BABB e BBAB). Antes que o treinamento desta tarefa fosse iniciado foram dadas as seguintes instruções:

“Em cada prova será apresentado um grupo de quatro sons. Entre eles há três sons iguais e um som diferente. O som diferente pode ser o segundo ou o terceiro na série. Indique com o mouse se o som diferente é o terceiro ou o segundo som. Ouça os quatro sons antes de responder. O tempo da resposta não é importante.”

3.5 Plano de Análise

Na tarefa de classificação, a variável dependente a ser observada é a proporção da resposta para cada uma das três alternativas. Por exemplo, se no ponto “340” da variável independente o sujeito responde “u” dezesseis vezes, “o” quatro vezes e “c” nenhuma vez, a proporção da resposta é 0,80 para a alternativa “u”, 0,20 para a alternativa “o”, e 0,0 para a alternativa “c”. Como são interessantes tanto análises intra- como análises entre-sujeitos, estas proporções serão obtidas para cada sujeito separadamente e também para o grupo de sujeitos – neste caso é calculada a média dos sujeitos em cada ponto da variável independente.

Três transformações foram efetuadas sobre a proporção da resposta. Os valores de proporção de cada sujeito e também os valores observados nos resultados do grupo foram transformados em valores de probito de forma a linearizar a relação entre as variáveis dependente e independente. Sobre as curvas correspondentes às respostas “u” e “c” foi realizada uma análise de probito para cada sujeito e também uma análise entre-sujeitos. Por meio dessa análise foram obtidos coeficientes de inclinação e valores que representam pontos na variável independente (isto é, pontos no *continuum*) em que a probabilidade da resposta é de 50%. A curva da resposta “o” não foi incluída na análise por motivos óbvios: ela não é informativa, já que as proporções correspondentes às outras duas respostas esgotam toda a informação nos resultados.

Para verificar se a declinação de uma das curvas é mais abrupta que a outra, as médias dos coeficientes de inclinação das curvas das respostas “u” e “c” obtidas na análise intra-sujeito foram comparadas por meio de um teste t em amostras pareadas.

Os valores da variável “proporção da resposta” na análise entre-sujeitos foram transformados também em valores que expressam o grau de ambigüidade dos estímulos em relação às respostas “u” e “o”. Novamente as médias das duas curvas foram comparadas.

A última transformação foi feita de modo a converter os resultados da tarefa de classificação em uma previsão dos resultados nas tarefas de discriminação supondo que dois sons podem ser discriminados apenas se forem classificados em categorias diferentes. Maiores detalhes sobre os procedimentos de análise e as transformações mencionadas serão apresentados no capítulo 4 (Resultados), a seguir.

Para as tarefas de discriminação, necessitamos de uma medida da discriminabilidade entre os sons dos pares apresentados. A proporção de respostas corretas é uma medida bastante utilizada, mas é influenciada não só pela discriminabilidade entre os sinais, mas também pelos critérios de decisão que os sujeitos usam para responder. Uma medida pretensamente pura de discriminabilidade é o d' . Trata-se de uma medida paramétrica que implica em duas suposições. Em primeiro lugar, admite-se que um sinal (ou apenas o ruído) faz com que uma variável de decisão assumira um valor dentro de uma distribuição e que esta distribuição é normal. A segunda suposição é a de que as variâncias das distribuições correspondentes aos dois sinais a serem discriminados (ou ao sinal e ao ruído, nas tarefas de detecção) são equivalentes. Informações sobre a estatística d' em cada uma das tarefas serão dadas no capítulo 4.

O interesse nas tarefas de discriminação reside na relação das mesmas com a tarefa de classificação. Numa análise intra-sujeitos, a medida dessa relação aqui empregada foi o índice

de percepção categórica (VAN HESSEN; SCHOUTEN, 1999), que expressa a proximidade entre a curva da discriminação observada e a curva de previsão da discriminação a partir da classificação e a semelhança entre as suas respectivas formas. Na tentativa de observar alguma diferença no grau em que os *continua* [u-ɔ] e [o-ɔ] são percebidos categoricamente, as médias destes índices para os dois *continua* foram comparadas por meio de um teste t em amostras pareadas.

Numa análise entre-sujeitos, foram empregadas três medidas da relação entre discriminação e classificação, obtidas a partir das médias dos valores de d' dos sujeitos em cada ponto do *continuum*. São elas a correlação entre a discriminação observada e a previsão da discriminação, o desvio quadrático médio (DMQ's) entre a curva de discriminação observada e a curva da discriminação prevista. Foram obtidos os graus de significância correspondentes às diferenças entre os DMQ's e entre os coeficientes de correlação obtidos dois *continua* [u-ɔ] e [o-ɔ].

4 Resultados

A tarefa de classificação nos permite observar a maneira como os sujeitos dividem o *continuum* de vogais nas categorias correspondentes às alternativas de resposta na tarefa. Considera-se que os resultados contêm informações a respeito das representações das categorias em questão na memória e também do critério de decisão adotado para a resposta, embora não seja possível, apenas com base naqueles, estabelecer de maneira satisfatória uma distinção entre estas duas fontes de variabilidade.

Os aspectos a serem observados são os limites de categoria (as regiões que separam duas categorias vizinhas no *continuum*), a forma das curvas, o grau de ambigüidade dos estímulos, especialmente nas regiões de limite de categoria, e o grau em que os resultados da tarefa de classificação se relacionam com o desempenho dos sujeitos nas tarefas de discriminação. Comparações serão feitas entre os padrões observados nas regiões correspondentes aos dois contrastes – [u/o] e [o/o].

Para observar a relação entre os resultados da tarefa de classificação e das tarefas de discriminação, os dados da classificação foram transformados de forma a proporcionarem uma predição dos resultados da discriminação. Esta transformação supõe que dois sons podem ser discriminados se e somente se pertencerem a categorias diferentes. Como a medida de discriminabilidade empregada neste estudo é o d' , essa predição pode ser obtida calculando-se a diferença entre os escores z das probabilidades de uma mesma resposta na tarefa de classificação aos dois estímulos a serem discriminados (MACMILLAN *et al.*, 1977). Este procedimento converte os valores de probabilidade correspondentes a cada estímulo na tarefa de classificação em valores de d' .

Como o *continuum* [u -o] empregado na tarefa de classificação foi dividido em dois para as tarefas de discriminação, gerando os *continua* [u -o] e [o -o], a previsão da discriminação foi calculada a partir dos resultados da classificação na parte do *continuum* correspondente a cada tarefa de discriminação. Foram descartadas eventuais provas em que respostas “u” foram dadas a estímulos envolvidos na previsão da discriminação no *continuum* [o-o], ou respostas “o” a estímulos envolvidos na previsão da discriminação no *continuum* [u-u], admitindo que os raros casos em questão se devam a erros de resposta. Houve muito pouca sobreposição entre as curvas correspondentes às probabilidades das respostas “u” e “o” no *continuum* da tarefa de classificação, sendo assim o efeito desse procedimento insignificante.

Para os cálculos de d' , escores z iguais a $-\infty$ e $+\infty$ foram evitados substituindo as proporções 1 e 0 por 0,99 e 0,01 (cf. SCHOUTEN, 1999; SCHOUTEN; VAN HESSEN, 1992; VAN HESSEN; SCHOUTEN, 1992; GERRITS; SCHOUTEN, 2004; SCHOUTEN; GERRITS; VAN HESSEN, 2003; GERRITS, 2001). Segundo Brown e White (2005), o melhor procedimento para eliminar as proporções 0 e 1 é adicionar uma constante entre 0,25 e 0,5 aos valores de frequência absoluta correspondentes a cada resposta possível a cada estímulo. Entretanto, como a quantidade de repetições de cada estímulo não é a mesma nas tarefas de classificação, 2AFC e 4I2AFC, a transformação teria efeitos diferentes sobre os resultados de cada uma. A alternativa aqui escolhida proporciona uma aproximação suficiente aos propósitos deste trabalho, já que o interesse se concentra não sobre os valores de d' em si, mas sobre a relação entre os resultados nas tarefas de classificação e discriminação, especialmente nas regiões próximas aos limites de categoria.

No GRÁF.1 são apresentadas as curvas que representam as proporções de cada resposta possível para cada estímulo da tarefa de classificação, cada ponto indicando a média sobre os

sujeitos. Pode-se observar declinações abruptas nas curvas que indicam as proporções de cada uma das três alternativas de resposta ao longo do *continuum*. As duas regiões em que essas ocorrem correspondem aos limites entre as três categorias. As curvas correspondentes às categorias “u” e “o” se cruzam entre os estímulos “360” e “370” e as curvas correspondentes às categorias “o” e “ɔ” se cruzam entre os estímulos “510” e “520”. As proporções próximas de 1,0 nas demais regiões mostram que os sons dentro das categorias foram classificados com muita facilidade pelos sujeitos. Se estes sons são percebidos como mais similares ou menos discrimináveis entre si, isto é algo que não pode ser decidido apenas com base nos resultados de uma tarefa de classificação.

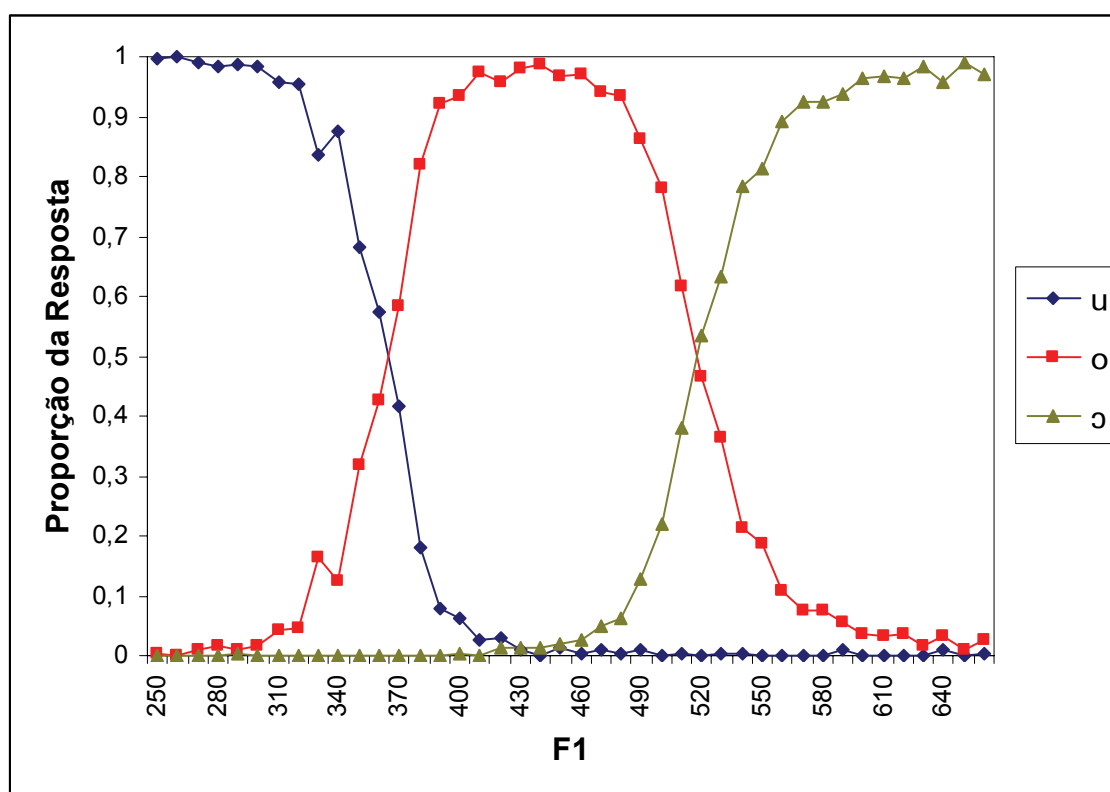


GRÁFICO 1 - Proporções das respostas “u”, “o” e “ɔ” ao longo do *continuum* em F1.

Com o objetivo de comparar as inclinações das curvas correspondentes às categorias, os valores dos pontos das curvas das respostas “u” e “o” na tarefa de classificação foram, para cada sujeito, transformados em valores de probito¹. Por meio de uma análise de probito, foram obtidos valores de inclinação para as curvas. Estes valores são coeficientes que indicam a inclinação de uma reta ajustada aos dados relacionando o probito (proporção da resposta) da variável dependente à variável independente (valor de F1). Desta forma, eles podem ser interpretados como uma medida do impacto da variável independente sobre a variável dependente – no caso, em quantas unidades o probito da proporção da resposta se altera quando o valor de F1 é alterado em uma unidade².

As médias dos coeficientes de inclinação para “u” e “o” são, respectivamente, -0,0371 e 0,0293. Foi observado por meio de um teste t para amostras pareadas que a média do valor absoluto dos coeficientes de inclinação é significativamente maior para a curva correspondente à resposta “u” ($p = 0,033$).

Foram também transformados em valores de probito os valores que compõem o GRÁF.1, isto é, as médias das proporções de resposta dos sujeitos em cada ponto no *continuum* (análise *across-subject*). O GRÁF.2 expressa a relação entre estes valores de probito e o valor de F1.

¹ Na definição aqui empregada, o probito de p é igual a $\Phi^{-1}(p)$. A função Φ é a função de distribuição cumulativa.

² Note-se que há aqui uma abstração, já que não se considera o fato de que a existência de limiares perceptivos e “diferenças apenas perceptíveis” implica numa relação não-continua entre o objeto físico e o objeto perceptivo.

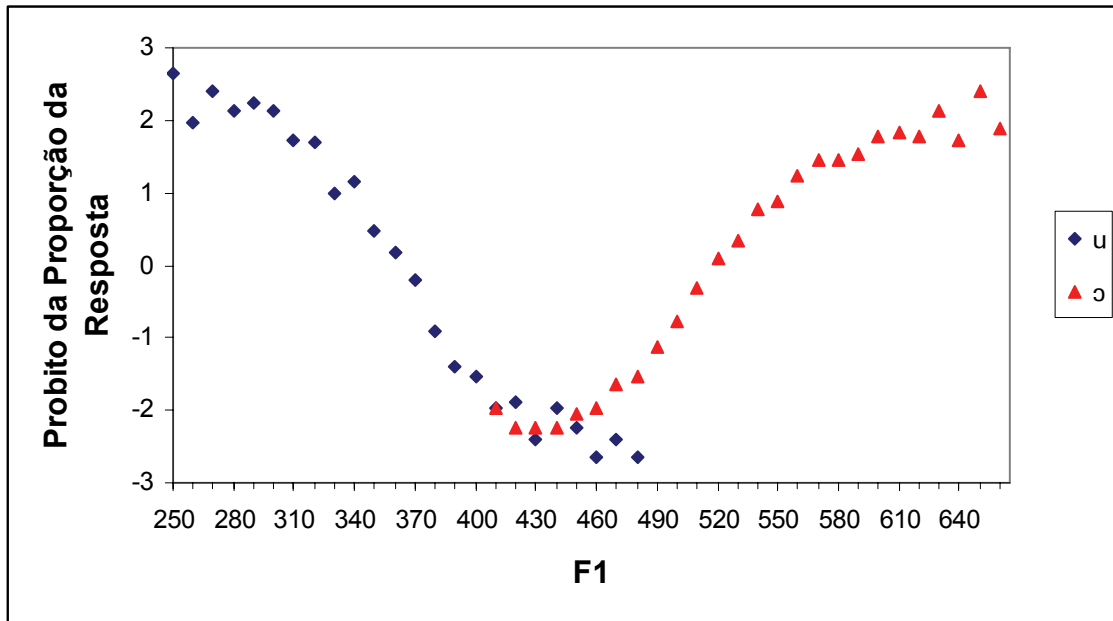


GRÁFICO 2 - Valores de probito das proporções das respostas ao longo do *continuum* em F1. Para compor este gráfico, as caudas inferiores das curvas foram cortadas de forma a terem o mesmo tamanho das caudas superiores. Para isto, em cada curva o número de pontos com valores abaixo de -1,6449 (valor correspondente à proporção 0,05) foi limitado pelo número de pontos com valores acima de 1,6449 (valor correspondente à proporção 0,95). Além disso, para evitar as proporções 0 e 1, onde não ocorreu nenhuma resposta foi registrado 0,5 e onde ocorreram 20 respostas foi registrado 19,5.

De acordo com a análise de probito aplicada a estas médias sobre todos os sujeitos, 362,84 é o valor de F1 correspondente à probabilidade 0,50 de ocorrência da resposta “u”, e 524,52 é o valor correspondente à probabilidade 0,50 da resposta “c”. A maneira como se estabeleceram os pontos correspondentes à probabilidade 0,50 pode ser considerada uma estimativa razoável dos limites de categoria observados (Nittrouer e Miller, 1996).

O grau de ambigüidade de um estímulo foi definido, de modo a variar entre 0 e 1, como o quanto a proporção de uma resposta se desvia de 0,5, de acordo com a equação

$$A_i = 1 - \frac{|p(i | R) - 0,5|}{0,5} \quad (1)$$

onde $p(i|R)$ é a probabilidade de que a resposta ao estímulo i seja R . O valor 0,5 foi escolhido admitindo-se que um estímulo ambíguo gera indecisão entre duas alternativas de resposta (ou entre uma alternativa e a sua negação – como, por exemplo, “u” e “não-u”). Trata-se de uma versão da equação apresentada por Massaro (1987) modificada de modo que os valores possíveis do grau de ambigüidade variem entre zero e um.

Foram obtidos os graus de ambigüidade dos sons no *continuum* em relação às respostas “u” e “ɔ” a partir das médias dos resultados do grupo de sujeitos em cada ponto. Foi realizado um teste t em amostras independentes para comparar as médias dos dois grupos de valores – quais sejam, 0,0974 para “u” e 0,1283 para “ɔ”. No GRÁF.3 estão representados os graus de ambigüidade dos estímulos ao longo do *continuum* em F1. As duas médias não diferem significativamente ($p = 0,51$).

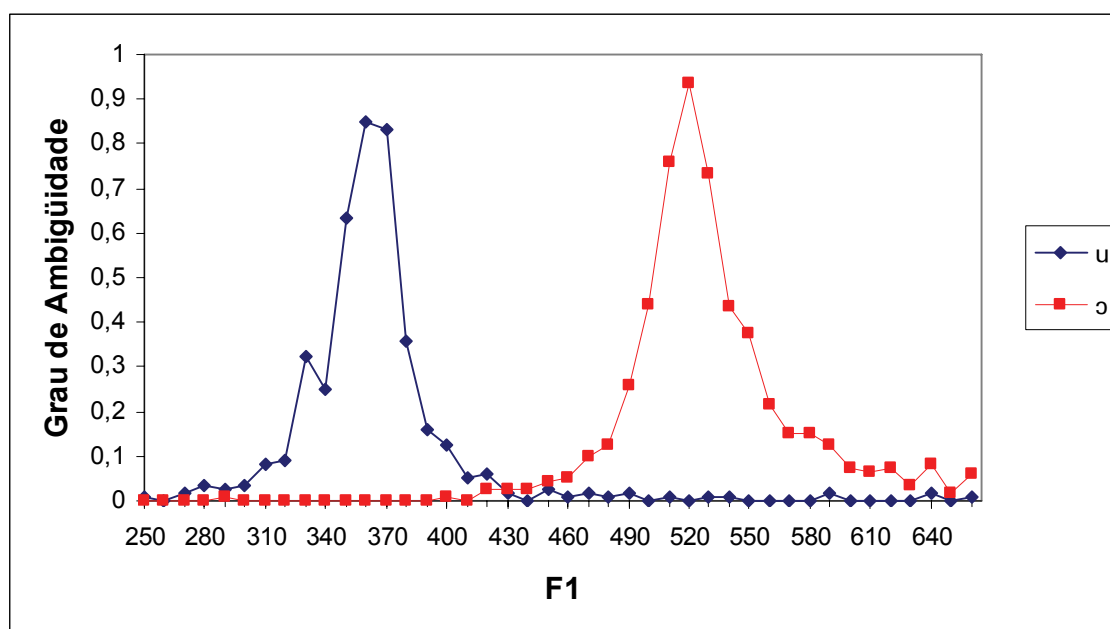


GRÁFICO 3 - Graus de ambigüidade dos estímulos do *continuum* em F1.

O mesmo teste foi realizado limitando-se o número de pontos com valores de proporção abaixo de 0,05 ao número de pontos com valores de proporção acima de 0,95, de modo a tornar as curvas menos assimétricas em relação ao tamanho das caudas. Novamente não há entre as médias para “u” e “s” – 0,1670 e 0,2067, respectivamente – uma diferença significativa ($p = 0,59$).

As tarefas de discriminação, especialmente quando investigadas em relação à tarefa de classificação, proporcionam informações importantes a respeito da forma com que os sujeitos percebem os sons do *continuum* e como as categorias são representadas na memória do falante. O principal aspecto a ser observado é a relação entre a discriminação e a classificação nas duas regiões do *continuum* em que as categorias contrastam. Como a tarefa 2AFC força uma estratégia classificatória, é esperada uma relação mais estreita com os resultados da tarefa de classificação.

De acordo com a Teoria da Detecção do Sinal (SDT – *Signal Detection Theory*; GREEN; SWEETS, 1969) uma medida de sensibilidade livre de tendências (ou propensões) que o sujeito possa apresentar em relação a uma ou outra resposta, é o d' , que mede a distância, em unidades de desvio padrão, entre as médias das duas distribuições numa variável de decisão correspondentes aos dois sinais a serem discriminados. Entretanto, para que o d' não seja afetado pelas tendências de resposta, as distribuições devem ser idealmente ambas normais e apresentar a mesma variância. Desta forma, ao utilizarmos esta medida, admitimos que tais condições são satisfeitas.

Para a tarefa 2AFC, o d' foi calculado pela seguinte fórmula (GREEN; SWEETS, 1969; MACMILLAN; CREELMAN, 2004):

$$d' = \frac{z(H) - z(F)}{\sqrt{2}} \quad (2)$$

onde “H” é a proporção de respostas “u-o” (ou “o-o”) dado um estímulo com os dois sons apresentados em ordem crescente (em relação ao valor de F1), e “F” é a proporção das mesmas respostas dado um estímulo em que os dois sons são apresentados em ordem decrescente. Este procedimento supõe que a estratégia empregada pelo sujeito envolve uma operação equivalente à subtração $X_1 - X_2$, em que X_1 e X_2 são os intervalos de cada prova da tarefa. A variável de decisão é então o resultado desta subtração, cujos termos correspondem cada um a uma distribuição. Isto resulta numa terceira distribuição com uma média igual à diferença entre as médias das duas distribuições envolvidas na subtração e um desvio padrão $\sqrt{2}$ vezes maior que o desvio padrão das mesmas. Daí a divisão por $\sqrt{2}$ na equação (2).

Duas análises foram realizadas para a tarefa 4I2AFC. De acordo com Macmillan e Creelman (2004) a tarefa 4I2AFC deve ser analisada da mesma forma que a tarefa 2AFC. Como o primeiro e o quarto sons não correspondem a nenhuma alternativa de resposta, um sujeito ideal agiria como se estivesse fazendo uma tarefa 2AFC. Assim, o d' foi calculado pela fórmula citada acima para a análise da tarefa 2AFC.

Como será visto a seguir, os resultados foram muito diferentes dos obtidos na tarefa 2AFC, tanto no escore quanto no grau de percepção categórica. Com isso, uma análise adicional foi realizada a partir da hipótese de que os sujeitos se comportaram na tarefa 4I2AFC da mesma forma que o fariam numa tarefa 4IAX – uma tarefa em que 4 sons são apresentados, um dos quais sendo diferente, e o sujeito deve indicar se este se deu no primeiro ou no segundo par. Vale lembrar que na tarefa 4I2AFC realizada neste estudo, o intervalo de tempo entre o

segundo e o terceiro som é um pouco mais longo que os demais, de modo que os sujeitos ouçam claramente dois pares de sons. O d' , nesta análise alternativa, foi calculado então a partir da fórmula proposta por Macmillan *et al.* (1977) para calcular o valor de d' na tarefa 4ifx:

$$H = \Phi\left(\frac{d'}{2}\right) + \Phi\left(-\frac{d'}{2}\right) = 1 - F \quad (3)$$

onde Φ é a função de distribuição cumulativa. Esta equação supõe que não há tendências de resposta (H igual a F). Assim, para se obter um valor de d' livre dessas tendências, foi usada uma tabela (semelhante à apresentada por Kaplan, MacMillan e Creelman, 1978) que relaciona valores da subtração $z(H) - z(F)$ a valores obtidos por meio da fórmula (3). Os resultados relativos à análise alternativa da tarefa 4I2AFC serão indicados pelo termo “4I2AFC”.

Pode-se observar nos GRAF. 4, 5, 6 e 7 as curvas previstas (classificação) e as curvas obtidas nas duas tarefas de discriminação.

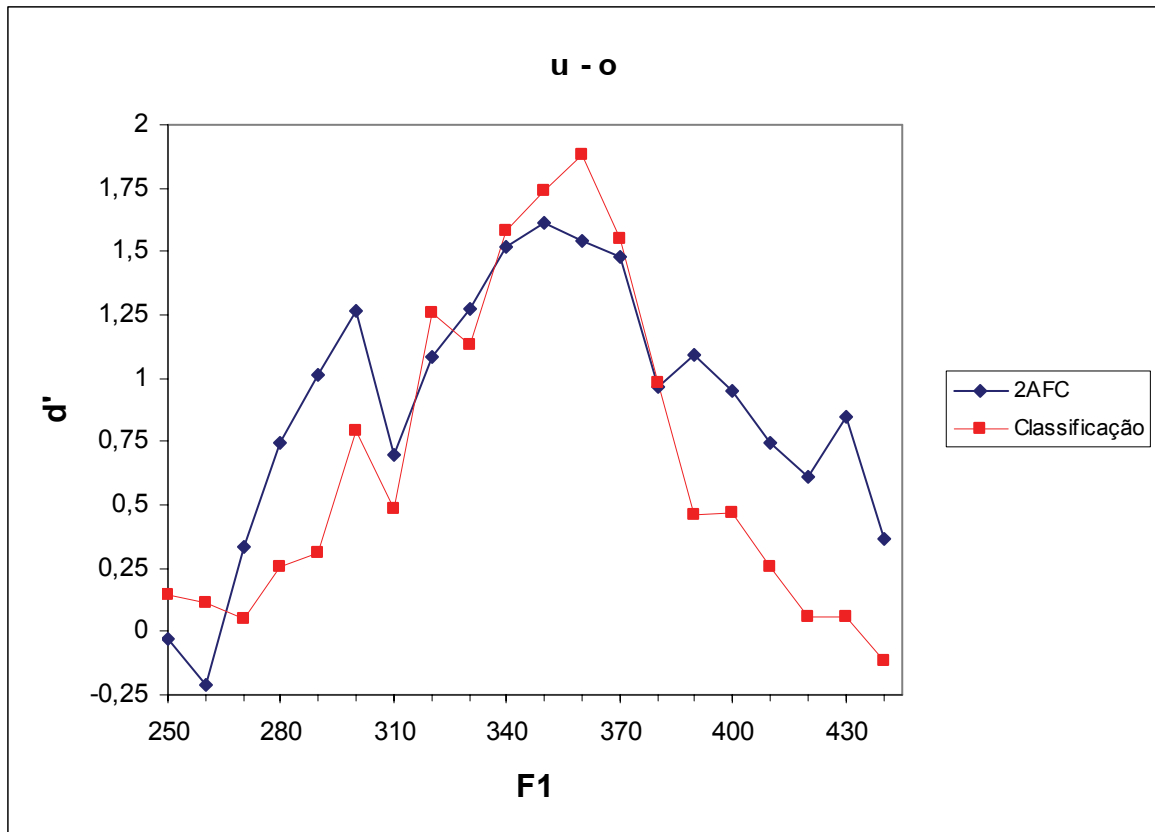


GRÁFICO 4 - Discriminação obtida e discriminação prevista (classificação) no *continuum* [u-o]; tarefa 2AFC.

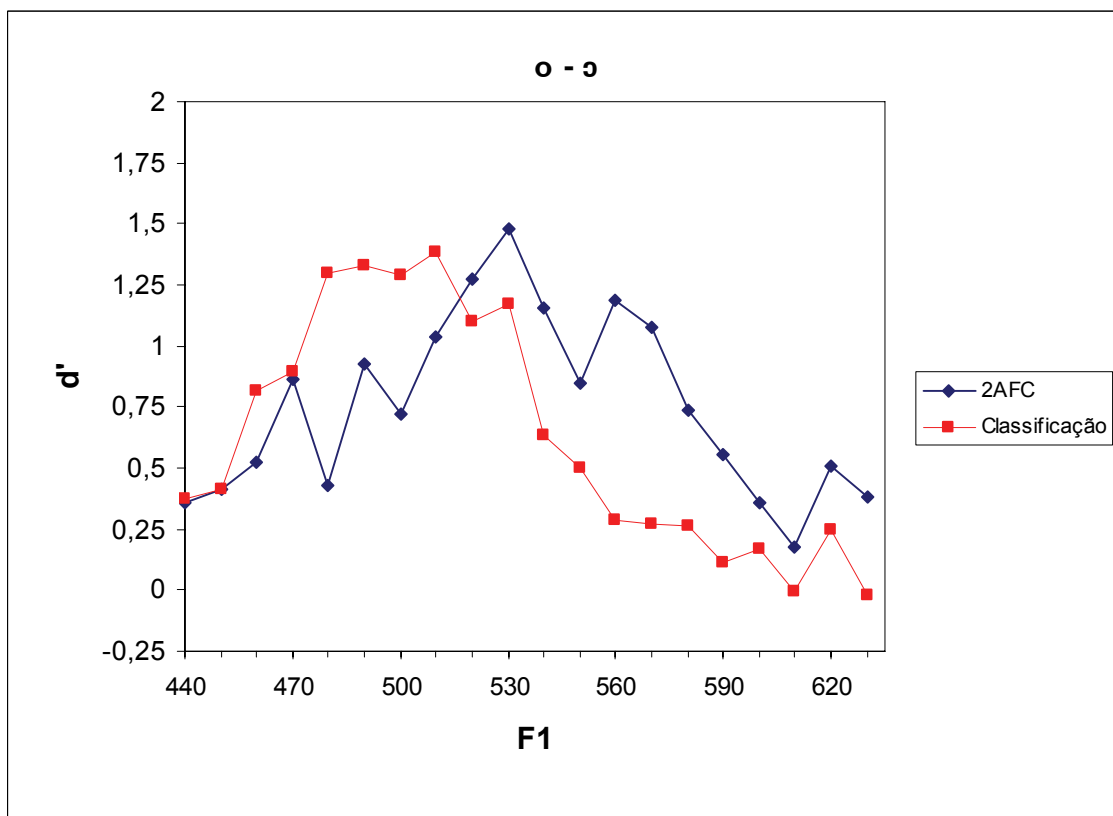


GRÁFICO 5 - Discriminação obtida e discriminação prevista (classificação) no *continuum* [o-ɔ]; tarefa 2AFC.

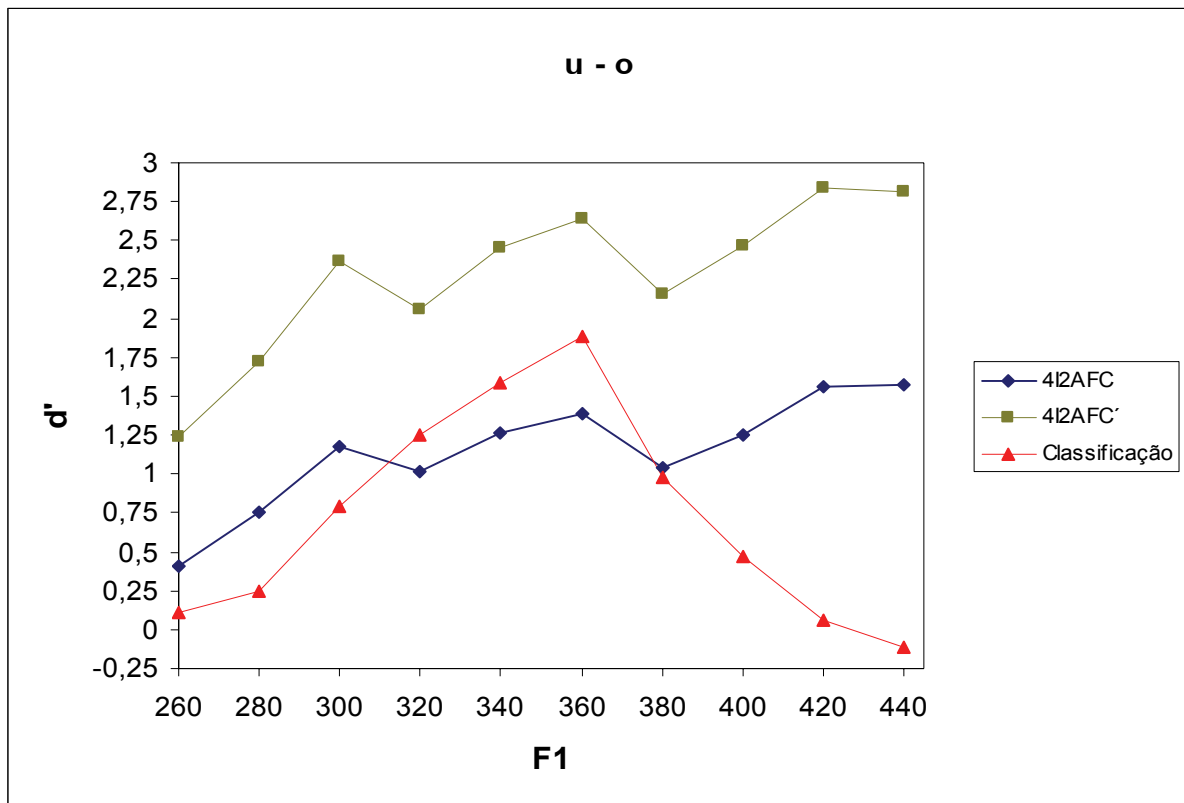


GRÁFICO 6 - Discriminação obtida e discriminação prevista (classificação) no *continuum* [u-o]; tarefa 4I2AFC.

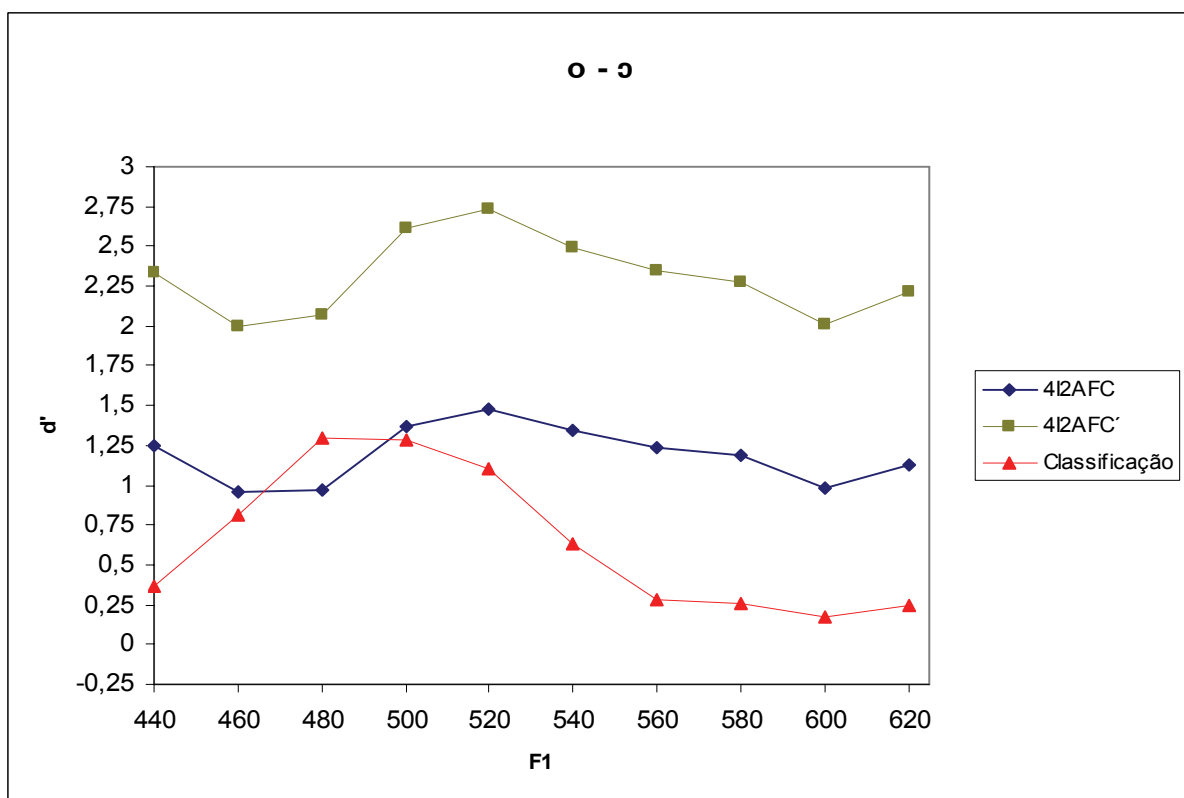


GRÁFICO 7 - Discriminação obtida e discriminação prevista (classificação) no *continuum* [o-ɔ] tarefa 4I2AFC.

Percebe-se facilmente no GRÁF. 1 uma relação entre os valores de d' na tarefa de classificação e na tarefa 2AFC para o *continuum* [u-o]. Um pico pronunciado se apresenta em ambas as curvas na região do limite entre as categorias [u] e [o] (em torno dos estímulos “350” e “360”), e o desempenho decai abruptamente na medida em que se afasta dessa região, para níveis abaixo do previsto pelos resultados da classificação. No GRÁF. 2, a relação entre discriminação e classificação no *continuum* [o-ɔ] se torna menos clara. Os picos nas duas curvas não coincidem e parecem menos distintos. No que diz respeito aos resultados das tarefas 4I2AFC, não se verifica nos GRÁF. 6 e 7 relações evidentes entre o desempenho observado nas tarefas de discriminação e o que é previsto pelos resultados da classificação.

Para verificar o grau de dependência da discriminação em relação à classificação foi calculado o índice de percepção categórica, proposto por van Hessa e Schouten (1999), que leva em conta a “semelhança” e a “proximidade” entre a curva de discriminação obtida e a curva de discriminação prevista pelos resultados da tarefa de classificação:

$$PC = 100 \left(\frac{r}{1 + 0,2|d'(\text{class}) - d'(\text{disc})|} \right) \quad (4)$$

O numerador é o coeficiente de correlação entre as duas curvas e o denominador contém um termo que é a média das diferenças absolutas entre o d' previsto e o d' obtido em cada ponto do *continuum* multiplicado por uma constante. Desta forma, o grau de percepção categórica varia entre 0 e 100 (ou -100, no caso de correlação negativa), aumentando se a correlação entre as curvas aumenta e se a distância diminui.

Índices de percepção categórica foram calculados para cada sujeito em cada uma das tarefas de discriminação. A TAB. 1 mostra as médias dos índices de percepção categórica obtidos em cada tarefa e em cada *continuum*. Para as tarefas 4I2AFC foram calculados dois valores, cada um admitindo uma hipótese diferente sobre as operações do sujeito na realização da tarefa. Por meio de um teste t em amostras pareadas foi comparada para cada tarefa, a média dos índices de percepção categórica obtida no *continuum* [u-o] com a média obtida no *continuum* [o-o]. Não foram observadas diferenças significativas ($p > 0,10$ em todas as tarefas).

TABELA 1

Médias dos índices de percepção categórica (PC) nas tarefas 2AFC e 4I2AFC para os *continua* [u-o] e [o-o] e significância (p bicaudal) da diferença entre os valores correspondentes aos dois *continua* (análise intra-sujeitos).

Tarefa	Continuum		p
	u-o	o-o	
2AFC	32,0241	16,579	0,139273
4I2AFC	11,3742	5,7965	0,475431
4I2AFC'	10,6636	7,5745	0,626308

Na TAB. 2 estão os índices de percepção categórica e a correlação entre discriminação obtida e prevista para as quatro tarefas de discriminação, calculado a partir das médias sobre todos os sujeitos dos valores de d' para cada ponto nos *continua*. Adicionalmente, são apresentados os coeficientes de correlação entre a discriminação prevista e a discriminação obtida e os desvios quadráticos médios da discriminação obtida em relação à discriminação prevista. Para transformas estes desvios quadráticos médios em valores em unidades de d' basta tomar a raiz quadrada dos mesmos, obtendo-se nos *continua* [u-o] e [o-o], respectivamente, 0,4158 e 0,4550 na tarefa 2AFC , 0,8128 e 0,6895 na tarefa 4I2AFC e 1,7127 e 1,6828 para a análise alternativa 4I2AFC'.

TABELA 2

Índice de percepção categórica (PC), coeficiente de correlação entre discriminação obtida e prevista (r) e desvio quadrático médio (DQM) da discriminação obtida em relação à discriminação prevista, calculados a partir das médias de d' do grupo de sujeitos (análise entre-sujeitos) nas tarefas 2AFC e 4I2AFC para os *continua* [u-o] e [o-o].

Tarefa	Continuum					
	u-o			o-o		
	PC	R	DQM	PC	r	DQM
2AFC	77,3663	0,828 (p < 0,0001)	0,1729	44,3904	0,4773 (p = 0,0333)	0,2071
4I2AFC	10,3382	0,1163 (p = 0,6253)	0,6607	18,7728	0,2116 (p = 0,3706)	0,4754
4I2AFC'	11,2575	0,3473 (p = 0,5350)	2,9335	26,0607	0,1475 (p = 0,1335)	2,8318

O índice de percepção categórica é consideravelmente mais elevado na tarefa 2AFC. O índice é maior para o *continuum* [u-o] do que para o *continuum* [o-o] na tarefa 2AFC, mas a relação se inverte na tarefa 4I2AFC.

As únicas correlações significativas ($p < 0,05$) foram observadas na tarefa 2AFC. Os demais coeficientes apresentam valores p acima de 0,10. A transformação de Fisher de r para z foi usada na estimação de um valor p que indique a significância da diferença entre os dois coeficientes de correlação observados na tarefa 2AFC. A diferença entre os coeficientes correspondentes aos *continua* [u-o] e [o-o] é marginalmente significativa ($p = 0,054$).

Em resumo, as curvas nos GRAF. 1 e 3 demonstram que as respostas dos sujeitos se deram de acordo com três categorias bem delimitadas dentro das quais os estímulos variam em relação

ao nível de ambigüidade, que se aproxima de zero nas regiões do *continuum* correspondentes aos centros das categorias. Não foi observada diferença significativa entre os níveis médios de ambigüidade correspondentes às curvas das respostas “u” e “o”. Numa análise em que, para cada sujeito, foram obtidos um coeficiente de inclinação para a curva “u” e um para a curva “o”, verificou-se uma diferença significativa ($p > 0,05$) entre as médias destes coeficientes.

Se a cada sujeito for atribuído um índice de percepção categórica em cada uma das tarefas de discriminação e forem comparadas as médias no *continuum* [u-o] com as médias no *continuum* [o-o] não se observa diferença significativa. Se forem tomadas as médias dos valores de d' do grupo de sujeitos em cada ponto dos *continua* são verificados coeficientes de correlação entre discriminação e classificação significativos ($p > 0,05$) apenas para a tarefa 2AFC. Nesta tarefa, a diferença entre os coeficientes de correlação observados nos dois *continua* não atinge o nível de significância 0,05, embora se aproxime bastante, sendo o valor observado no *continuum* [u-o] maior. Os maiores índices de percepção categórica foram observados na tarefa 2AFC, entre os quais o maior foi observado no *continuum* [u-o]. Os desvios médios quadráticos (DQM's) das diferenças entre os valores previstos pela classificação e os valores observados nas tarefas de discriminação são uma medida do ajustamento da previsão às observações. Não foram observadas diferenças significativas entre os DQM's obtidos nos *continua* [u-o] e [o-o].

5 Discussão

Dadas as inconsistências observadas no comportamento das vogais médias do Português Brasileiro em relação à funcionalidade dos contrastes entre vogais médias abertas e vogais médias fechadas, o objetivo deste estudo foi investigar possíveis diferenças nos modos como estes contrastes são representados no sistema perceptivo do falante de nossa língua. Para isto, foi observado um *continuum* de sons vocálicos variando em um único parâmetro (F1) ao longo do qual se estabelecem dois contrastes, um entre as vogais médias posteriores aberta e fechada (contraste [o/ɔ]) e outro – sobre o qual não há dúvidas no que se refere à sua funcionalidade na língua – entre a vogal média posterior fechada e a vogal alta posterior (contraste [u/o]). A comparação entre os resultados em tarefas de classificação e discriminação referentes aos contrastes [u/o] e [o/ɔ] foram o foco desta investigação.

Um limite de categoria menos distinto e abrupto entre as categorias [o] e [ɔ] do que o observado entre as categorias [o] e [u] e relações menos evidentes entre discriminação e classificação no *continuum* [o-ɔ] do que no *continuum* [u-ɔ], são resultados que indicariam que a distinção entre [o] e [ɔ] não é representada na memória do falante da mesma forma que ocorre com a distinção entre [o] e [u], sendo esta última melhor estabelecida e mais saliente. Isto poderia ser considerado uma evidência favorável à hipótese de que as categorias [o] e [ɔ] não correspondem a dois fonemas distintos como as categorias [o] e [u].

Padrões semelhantes nas observações referentes aos dois contrastes contrariam a hipótese de que [o] e [ɔ] não possuem representações fonológicas distintas na memória como ocorre com [o] e [u]. Entretanto, é importante salientar que, como os testes estatísticos foram empregados para verificar a presença de diferenças significativas em certos níveis de significância, a

ausência das mesmas não permite concluir que não existem diferenças reais entre os modos como são representadas as duas distinções no sistema perceptivo do falante.

Os resultados da tarefa de classificação são ambivalentes. Por um lado, a média dos valores absolutos dos coeficientes de inclinação dos sujeitos na resposta “u” é significativamente maior do que a média observada para a resposta “o”, sugerindo que há alguma diferença na forma como os sujeitos operam em relação aos contrastes [u/o] e [o/o]. Por outro lado, as médias dos graus de ambigüidade dos estímulos do *continuum* em relação às respostas “u” e “o” são semelhantes. Isto quer dizer que é possível que haja uma diferença real entre estas médias, mas pequena demais para atingir o nível de significância, assim como é possível que a diferença observada entre os coeficientes de inclinação seja devida a variações aleatórias ou algum fator não considerado no estudo. Torna-se deste modo importante considerar estes resultados à luz do que foi observado nas tarefas de discriminação.

Como dissemos, a tarefa de discriminação 2AFC encoraja uma estratégia classificatória, de modo que é esperado um grau mais alto de percepção categórica. De acordo com Massaro (1987), não é de grande interesse uma relação estreita entre a tarefa de classificação e aquelas tarefas de discriminação que encorajam estratégias classificatórias, já que o sujeito estaria executando as mesmas operações ou operações muito semelhantes. Com a ausência (ou baixos níveis) de percepção categórica para tarefas em que os sujeitos podem atentar para os traços acústicos, como a tarefa 4I2AFC, observada neste e em outros estudos recentes (Gerrits e Schouten, 2004; Schouten, Gerrits e van Hessen, 2003; Gerrits, 2001), o conceito de percepção categórica se enfraquece ainda mais. Entretanto, este ganha novo alento com as pesquisas psicofisiológicas com o potencial evocado *Mismatch Negativity*, que vêm demonstrando efeitos de limite de categoria quando se observa respostas em níveis

relativamente básicos do processamento auditivo no córtex cerebral, que não dependem de mecanismos de decisão e atenção, conforme vimos no capítulo de revisão bibliográfica. A este respeito vale ainda mencionar o estudo realizado por Dehaene-Lambetz *et al.* (2005), também empregando a *Mismatch Negativity*, que forneceu evidências de dois processos distintos no cérebro que correspondem a um modo fonêmico, ou classificatório, e ao modo acústico na discriminação de sons da fala, em consonância com certos modelos psicofísicos de duplo-processo (FUJISAKI; KAWASHIMA, 1971; SCHOUTEN; VAN HESSEN, 1992; VAN HESSEN; SCHOUTEN, 1992). Torna-se então razoável entender a tarefa 2AFC não como uma tarefa de classificação com roupagem de discriminação, mas como uma tarefa de discriminação em que um modo classificatório de percepção se faz necessário. De qualquer forma, o interesse deste estudo, no que diz respeito à tarefa de discriminação, está nas comparações que podem ser feitas entre os valores que expressam a relação entre discriminação e classificação obtidos nos *continua* [u-o] e [o-o].

Analisando as tarefas de discriminação e a relação destas com a tarefa de classificação, as evidências que não corroboram a hipótese que investigamos de que há diferenças entre os contrastes [u/o] e [o/o] seriam (a) a ausência de diferenças significativas entre as médias dos índices de percepção categórica dos sujeitos nos *continua* [u-o] e [o-o]; e (b) a ausência de diferenças significativas entre os desvios quadráticos médios (DQM's) da discriminação observada e os DQM's da discriminação prevista, obtidos na análise das médias dos valores de d' do grupo de sujeitos em cada ponto dos *continua*.

Ainda em relação à análise entre-sujeitos – isto é, aquela em que são examinados os valores médios de d' no grupo de sujeitos – os índices de percepção categórica foram, como se esperava, maiores na tarefa 2AFC, sendo o valor mais alto observado no *continuum* [u-o]. Na

tarefa 2AFC, a diferença entre os coeficientes de correlação com a tarefa de classificação obtidos nos *continua* [u-o] e [o-ɔ], cujo nível de significância é superior, mas bem próximo, a 0,05, indica que o *continuum* [o-ɔ] pode ser percebido de forma menos categórica.

Quanto à relação entre a tarefa de classificação e a tarefa 4I2AFC, os baixos valores do índice de percepção categórica, a ausência de picos coincidentes com os picos nas curvas de discriminação prevista a partir da classificação e a ausência de correlações significativas entre discriminação e classificação, todos constituem evidências de que os sujeitos empregaram predominantemente estratégias baseadas nos traços acústicos dos estímulos. Isto corrobora a literatura citada anteriormente de que esta tarefa exige uma estratégia acústica.

Embora os resultados não permitam conclusões seguras em relação à questão colocada neste trabalho, dada a falta de evidências suficientemente confiáveis do ponto de vista estatístico, é importante notar que as diferenças observadas, estatisticamente significantes ou não, se deram geralmente na mesma direção. A inclinação da curva da resposta “ɔ” na tarefa de classificação é menor do que a inclinação da curva da resposta “u”; o nível médio de ambigüidade correspondente a esta mesma curva é também um pouco menor; na tarefa 2AFC, as medidas que relacionam classificação e discriminação, apresentam no *continuum* [o-ɔ] valores que, comparados com seus correspondentes no *continuum* [u-o], sugerem uma associação mais fraca entre a discriminação prevista e a discriminação observada (menor índice de percepção categórica, menor correlação entre classificação e discriminação e maior DQM). A única exceção é a tarefa 4I2AFC na análise entre-sujeitos. Todavia, isto não é inconsistente com as demais observações, já que os resultados desta tarefa indicam que os sujeitos não empregaram uma estratégia classificatória – isto é, não recorreram a representações de categorias de vogais na memória de longo prazo.

Isto nos leva a entender que, ainda que os resultados não sejam conclusivos a respeito das diferenças na forma como o sistema perceptivo do falante do Português Brasileiro opera em relação ao contraste [o-ɔ], eles chamam a atenção para a importância de estudos detalhados a respeito dos contrastes entre as vogais médias abertas e fechadas. É necessária uma investigação aprofundada a respeito da possibilidade de que o falante do Português Brasileiro, ao operar no modo fonêmico, discriminando sons a partir de categorias armazenadas na memória de longo prazo, não estabeleça o contraste entre as vogais médias abertas e fechadas de forma tão consistente quanto o faz em relação a outros contrastes entre categorias vocálicas.

Como dois contrastes que se estabelecem num *continuum* em F1 foram comparados em relação a descontinuidades nos resultados observadas em regiões de limite de categoria, deve ser considerada a possibilidade de que tais descontinuidades ocorram por razões não relacionadas às categorias vocálicas e suas delimitações. Descontinuidades inerentes ao sistema perceptivo que causem variações na discriminabilidade entre os estímulos ao longo do *continuum* poderiam determinar resultados como a diferença verificada na tarefa 2AFC entre os coeficientes de correlação entre classificação e discriminação nos *continua* [u-o] e [o-ɔ] ou a diferença entre as médias dos coeficientes de inclinação das curvas das respostas “u” e “ɔ”. Se, dada uma diferença fixa em termos de F1, os estímulos nas regiões em torno do limite entre as categorias [o] e [u] forem menos discrimináveis entre si do que os estímulos nas regiões em torno do limite entre [o] e [ɔ] em decorrência de características do sistema perceptivo, poderia ocorrer que a probabilidade de que os sujeitos procedam no modo fonêmico aumente quando é mais difícil efetuar discriminações em termos de traços acústicos, o que naturalmente resultaria em uma maior associação entre classificação e discriminação no

continuum [u-o]. Se ocorrer o inverso, isto é, se a discriminabilidade é menor em torno do limite entre [o] e [ɔ], o que implica em maiores diferenças minimamente perceptíveis nesta região, isto poderia contribuir para uma menor inclinação na curva da resposta “ɔ” na tarefa de classificação.

A este respeito seria útil examinar os resultados da tarefa 4I2AFC nos dois *continua*, já que nesta os sujeitos parecem ter empregado estratégias baseadas em traços acústicos. Nos GRÁF. 6 e 7 não há indícios de que uma dessas duas possibilidades tenha ocorrido. Entretanto, sabe-se que os limiares para a discriminação de formantes aumentam com a frequência dos formantes (FLANAGAN, 1955; KEWLEY-PORT; WATSON, 1996). Adicionalmente, Kewley-Port e Watson (1996) observaram saltos nos limiares de diferença quando um harmônico se alinhava exatamente à frequência central de um formante. Kewley-Port e Zheng (1998) avaliam modelos auditivos para sons vogais que podem fornecer métricas que reduzam a variabilidade observada nos limiares de diferença em tarefas de discriminação de formantes ao longo dos valores de frequência do formante de teste e de F0. Os estímulos empregados no presente estudo foram sintetizados de modo a variar na frequência fundamental de 145 a 85 Hz para evitar este tipo de interação (e também para obter sons mais naturais). Note-se que se descontinuidades sensoriais fossem responsáveis por descontinuidades observadas nos resultados das tarefas de classificação e 2AFC, estas deveriam se refletir também na tarefa 4I2AFC, já que, ao que tudo indica, esta é a tarefa em que as respostas são baseadas apenas na informação acústica. Não foi o caso.

A maior contribuição deste trabalho é a indicação de que devem ser realizados estudos detalhados a respeito do contraste entre as vogais médias abertas e fechadas. Os resultados destas pesquisas seriam de grande interesse para a fonética, fonologia e psicolinguística do

Português, pois estabeleceriam novas perspectivas e linhas de evidência que poderiam orientar formulações teóricas a respeito do sistema vocálico desta língua.

Seria interessante que estudos semelhantes ao presente fossem realizados com um número maior de sujeitos e/ou de provas para cada estímulo, de forma que a evitar que a variabilidade nos resultados impeça que diferenças efetivas atinjam os níveis de significância. O contraste entre as vogais médias anteriores deve ser também investigado. Além disso, é importante que sejam realizados experimentos em que, de alguma forma, as diferenças entre os sons sejam estabelecidas em unidades de diferenças minimamente perceptíveis, de forma que o *continuum* reflita uma dimensão perceptiva e não física (presumindo que a diferença minimamente perceptível corresponda a uma alteração constante na sensação).

Outro tipo de estudo que pode ser empreendido é a observação de respostas psicofisiológicas aos contrastes entre as vogais médias posteriores e entre as vogais médias anteriores. O potencial evocado *Mismatch Negativity* já se mostrou muito útil na pesquisa a respeito de representações de categorias dos sons da fala no córtex cerebral humano. É possível comparar a magnitude, a latência e a lateralização desta resposta aos contrastes [u/o] e [o/ɔ] (ou talvez [i/e] e [e/ɛ]). Isto forneceria evidências mais diretas e menos sujeitas a influências indesejadas como os efeitos dos mecanismos de atenção e de decisão nos resultados de experimentos envolvendo medidas comportamentais.

6 Conclusão

Com base no exame dos resultados deste estudo à luz de resultados anteriores no âmbito da pesquisa sobre a percepção da fala, conclui-se que há indícios de que a distinção entre [o] e [ɔ] não é tão bem definida quanto a distinção entre [o] e [u] no nível das representações das categorias vocálicas armazenadas na memória e empregadas nos processos perceptivos do falante do Português Brasileiro. Estudos adicionais a respeito desta possibilidade são necessários.

O esclarecimento desta questão é de grande interesse, sobretudo para a psicolinguística, a fonética e as teorias fonológicas de orientação funcionalista.

Referências Bibliográficas

- AALTONEN, O.; NIEMI, P.; NYRKE, T.; TUHKANEN, M. Event-related brain potentials and the perception of a phonetic continuum. *Biological Psychology*, v. 24, n. 3, p. 197-207, Jun.1987.
- ADES, A. A. Vowels, consonants, speech, and non-speech. *Psychological Review*, v. 84, n. 6, p. 524-530, Nov. 1977.
- ALHO, K. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear & Hearing*, v. 16, n. 1, p. 38-51, Fev. 1995.
- ALVES, M. M. *As vogais médias em posição tônica nos nomes do português brasileiro*. 1999. 136 f. Dissertação (Mestrado em Estudos Lingüísticos). Universidade Federal de Minas Gerais, Belo Horizonte, 1999.
- AULANKO R.; HARI, R.; LOUNASMAA O.V.; NÄÄTÄNEN R.; SAMS M. Phonetic invariance in the human auditory cortex. *Neuroreport*, v. 4, n. 12, p. 1356-1358, Set. 1993.
- BECKMAN, J. *Positional Faithfulness*. 1998. 270 f. Tese (Doutorado) - University of Massachusetts Amherst, Amherst, 1998.
- BARCLAY, J. R. Non-categorical perception of a voiced stop: A replication. *Perception & Psychophysics*, v.11, p.269-273, 1972.
- BISOL, L. Neutralização das átonas. *D.E.L.T.A.* v.19, n.2, p.267-276, 2003.
- BOERSMA, P. *Functional Phonology: formalizing the interactions between articulatory and perceptual drives*. 1998. 493 f. Tese (Doutorado) – Faculteit der Geesteswetenschappen, Universiteit van Amsterdam, Amsterdam, 1998.
- BOERSMA, P., WEENINK, D. *Praat: doing phonetics by computer*. Versão 4.3.12, 2005. Disponível em: <www.praat.org>
- BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001. 238 p.
- BRAIDA, L.; DURLACH, N. Intensity Perception II: Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, v. 51, n. 2, p. 583-502, Fev. 1972.
- BRAIDA, L.; LIM, J.; BERLINER, J.; DURLACH, N.; RABINOWITZ, W.; PURKS, S. Intensity Perception XIII: Perceptual-anchor model of context-coding. *Journal of the Acoustical Society of America*, v.76, n. 3, p. 722-731, Set. 1984.
- BROWN, G.; WHITE, G. The optimal correction for estimating extreme discriminability. *Behavior Research Methods*, v. 37, n. 3, p. 436-449, Ago. 2005.
- CAGLIARI, L. C. *Fonologia do Português: análise pela geometria de traços*. 2.ed. Campinas: Edição do Autor, 1997. v. 1. 150 p. (Série Lingüística).

CHOMSKY, N.; HALLE, M. *The Sound Pattern of English*. New York: Harper & Row, 1968. 470 p.

CUNHA, V. *Um traço do vocalismo português: a metafonía*. 1991. 154 f. Dissertação (Mestrado em Estudos Lingüísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 1991.

CUTTING, J.; ROSNER, B. Categories and boundaries in speech and music. *Perception & Psychophysics*, v. 16, p.564-570, 1974.

DEHAENE-LAMBERTZ, G. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, v. 8, n. 4, p. 919-924, Mar. 1997.

DEHAENE-LAMBERTZ, G.; PALLIER, C.; SERNICLAES, W.; SPRENGER-CHAROLLES, L.; JOBERT, A.; DEHAENE, S. Neural correlates of switching from auditory to speech perception. *Neuroimage*, v. 24, n. 1, Jan. 2005.

DURLACH, N.; BRAIDA, L. Intensity perception. I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, v. 46, n. 2, 372-283, Ago.1969.

EIMAS, P.; MILLER, J.; JUSCZYK, P. On infant speech perception and the acquisition of language. In: S. HARNAD (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press, 1987. p. 161-195.

FRY, D.; ABRAMSON, A.; EIMAS, P., LIBERMAN, A. The identification and discrimination of synthetic vowels. *Language and Speech*, v. 5, p. 171-189, 1962.

FOSS, D. SWINNEY, D. On the psychological reality of the phoneme: Perception, identification and consciousness. *Journal of Verbal Behavior*, v. 12, n. 3, p. 246-257, Jun. 1973.

GERRITS, E. *The categorization of speech sounds by adults and children*. 2001, 131 f. Tese (Doutorado) – Universiteit Utrecht, Utrecht, 2001.

GERRITS, E.; SCHOUTEN, M. E. H. Categorical perception depends on the discrimination task. *Perception & Psychophysics*, v. 66, n. 3, p. 363-376, Abr. 2004.

GREEN, D.; SWEETS, J. *Signal Detection Theory and Psychophysics*. New York: John Wiley and Sons, 1966. 455 p.

HARNAD; S. Introduction: Psychophysical and cognitive aspects of categorical perception: A critical review. In: S. HARNAD (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press, 1987. p. 1-25.

HARY, J; MASSARO, D. Categorical results do not imply categorical perception. *Perception & Psychophysics*, v. 32, n. 5, p. 409-418, Nov. 1982.

HEALY, A.; REPP, B. Context independence and phonetic mediation in categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, v.8, n.1, p.68-80, Fev. 1982.

IVERSON, P.; KUHL, P. K. Perceptual magnet and phoneme boundary effects in speech perception: Do they rise from a common mechanism? *Perception & Psychophysics*, v.62, n.4, p.874-886, Maio. 2000.

JACOBSEN, T.; SHRÖGER, E.; ALTER, K. Pre-attentive perception of vowel phonemes from variable speech stimuli. *Psychophysiology*, v. 41, n. 4, p. 654-659, Jul. 2004.

KEWLEY-PORT, D; WATSON, C. S. Formant-frequency discrimination for isolated English vowels. *Journal of the Acoustical Society of America*, v. 95, n. 1, p 485-496, Jan. 1994.

KAPLAN, H., MACMILLAN, N., CREELMAN, D. Tables of d' for variable-standard discrimination paradigms. *Behavior Research Methods & Instrumentation*, v. 10. n. 6, 796-813, 1978.

KUHL, P. The special-mechanisms debate in speech research: Categorization tests on animals and infants, In: S. HARNAD (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press, 1987. p. 355-386.

KUHL, P. Human adults and human infants show a “perceptual magnet effect” for the prototype of speech categories, monkeys do not. *Perception & Psychophysics*, v. 50, n. 2, p. 93-107, Ago. 1991.

LEE, S. H. Mid Vowel Alternations in Verbal Stems in Brazilian Portuguese. *Journal of Portuguese Linguistics*, v.2, n.2, 2003.

LIBERMAN, A.; HARRIS, K.; HOFFMAN, H.; GRIFFITH, B. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, v.58, p. 358-368, 1957.

LIBERMAN, A.; MATTINGLY, I. The Motor Theory of Speech Perception Revised. *Cognition*, v. 21, n. 1, p. 1-36, Out. 1985.

LOTTO, A.; KLUENDER, K.; HOLT, L. Depolarizing the perceptual magnet effect. *Journal of the Acoustic Society of America*, v. 103, n.6, p.3648-3654, Jun. 1998.

MACMILLAN, N. Beyond the categorical/continuous distinction: A psychophysical approach to processing modes. In: S. HARNAD (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press, 1987. p. 53-85.

MACMILLAN, N.; CREELMAN, C. D. *Detection Theory: A user's guide*. 2. ed. Mahwah: Lawrence Erlbaum Associates, 2005. 492 p.

MACMILLAN, N.; GOLDBERG, R.; BRAIDA, L. Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua. *Journal of the Acoustical Society of America*, v. 84, n. 4, 1262-1280, Out. 1988.

MAGALHÃES, J. O. *Une étude de certains processus de la phonologie portugaise dans le cadre de la Théorie du Charme et du Gouvernement*. 1990. 322 f. Tese (Doutorado) - Université de Montreal, Montreal, 1990.

MASSARO, D. *Speech Perception by Ear and Eye: A paradigm for psychological inquiry*. Hillsdale: Lawrence Erlbaum Associates, 1987a. 336 p.

MASSARO, D. Categorical partition: A fuzzy logical model of categorization behavior. In: S. HARNAD (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press, 1987b. p. 254-283.

MASSARO, D. Categorical perception: Important phenomenon or lasting myth? In: 5th INTERNATIONAL CONGRESS OF SPOKEN LANGUAGE PROCESSING, 1998, Sydney: *Proceedings...* Sydney [s. n.], 1998. p.2275-2278,

MASSARO, D.; COHEN, M. M. Categorical or continuous speech perception: a new test. *Speech Communication*, v.2, n. 1, p.15-35, Maio. 1983.

MATTOSO CÂMARA JR., J. *Para o estudo da fonêmica portuguesa*. 2.ed. Rio de Janeiro: Padrão, 1977.144 p.

MATURANA, H.; VARELA, F. *A árvore do conhecimento: As bases biológicas para a compreensão humana*. São Paulo: Palas Athena, 283 p.

McQUEEN; J. M.; Cutler, A. Cognitive process in speech perception. In: W. J. Hardcastle; J. Laver (eds.) *The handbook of phonetic sciences*. 1997, p. 566-585.

MILLER, G. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, v. 63, n. 2. p. 81-97, Mar. 1956.

NÄÄTÄNEN, R.; GAILLARD, A. W. K., MÄNTYSALO, S. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, v. 42, n. 4, p. 313-329, Jul.1978.

NÄÄTÄNEN, R.; JACOBSEN, T.; WINKLER, I. Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, v. 42, n. 1, p. 25-32, Jan. 2005.

NÄÄTÄNEN, R.; LEHTOKOSKI, A.; LENNES, M.; CHEOUR, M.; HUOTILAINEN, M.; LIVONEN, A.; VAINIO, M.; ALKU, P.; LIMONIEMI, R. J.; LUUK, A.; ALLIK, J.; SINKKONEN, J.; ALHO, K. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, v. 385, n. 6615, p. 432-434, Jan. 1997.

NÄÄTÄNEN, R.; WINKLER, I. The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*. v.125, n. 6, p. 826-859, Nov.1999.

NEAREY, T. Speech perception as pattern recognition. *Jouranal of the Acoustical Society of America*, v. 101, n. 6, p. 3241-3254, Jun. 1997.

NEAREY, T. On the factorability of phonological unites in speech perception. In: J. Local, R. Ogden, R. Temple (eds.) *Papers in laboratory phonology IV: Phonetic interpretation*. Cambridge: Cambridge University Press, 2003, p. 197-221.

PISONI, D. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, v. 13, p. 253-260, 1973.

PISONI, D. Auditory short-term memory and vowel perception. *Memory and Cognition*, v. 3, p. 7-18, 1975.

PASTORE, R.; AHROON, W.; BAFFUTO, J.; FRIEDMAN, C.; PULEO, J.; FINK, E. Common factor model of categorical perception. *Journal of Experimental Psychology: Human Perception and Performance*, v. 3, n. 4, p. 686-896, Nov. 1977.

PASTORE, R. Categorical Perception: Some psychophysical models. In: S. HARNAD (ed.) *Categorical perception: The groundwork of cognition*. Cambridge: Cambridge University Press, 1987. p. 29-52.

PAAVILAINEN, P.; JARAMILLO, M.; NÄÄTÄNEN, R.; WINKLER, I. Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, v.265, n. 3, 179-182, Abr. 1999.

PAAVILAINEN, P.; SIMOLA, J.; JARAMILLO, M.; NÄÄTÄNEN, R.; WINKLER, I. Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology*, v. 38, n. 2, p.359-365, Mar. 2001.

PESCHL, M.; RIEGLER, A. Does representation needs reality? Rethinking epistemological issues in the light of recent developments and concepts in cognitive sciences. In: A. RIEGLER; M. PESCHL; A. VON STEIN (eds.) *Understanding representation in the cognitive sciences*. New York: Kluwer Academic/Plenum, 1999, p. 9-17.

PHILLIPS, C., Levels of representation in the electrophysiology of speech perception. *Cognitive Science*, v. 25, n. 5, p.711-731, Set.-Out. 2001.

PHILLIPS, C.; MARANTZ, A.; MCGINNIS, M.; PESETSKY, D.; WEXLER, K. YELLIN, E.; PÖPPEL, D.; ROBERTS, T.; ROWLEY, H. Brain Mechanisms of Speech Perception: A Preliminary Report. *MIT Working Papers in Linguistics*, v. 26, 125-163, 1995.

PHILLIPS, C.; PELLATHY, T.; MARANTZ, A.; YELLIN, E.; WEXLER, K.; PÖPPEL, D.; MCGINNIS, M.; ROBERTS; T. Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study. *Journal of Cognitive Neuroscience*, v. 12, n. 6, p. 38-55, Nov. 2000.

REPP, B. Categorical Perception: Issues, methods, findings. In: N. J. LASS (ed.) *Speech and Language: Advances in Basic Research and Practice*. Orlando: Academic Press, 1984. v.10, p. 243-335.

ROTHER-NEVES, R. Notas sobre o método na Neurolinguística. In: E. A. M. Mendes; P. M. Oliveira; V. Benn-Ibler. (org.). *O novo milênio: interfaces lingüísticas e literárias*. Belo Horizonte: Faculdade de Letras da UFMG, 2001, p. 189-195.

SCHOUTEN, M. E. H.; GERRITS, E.; VAN HESSEN, A. The end of categorical perception as we know it. *Speech Communication*, v. 41, n. 1, p. 71-80, Ago. 2003.

SCHOUTEN, M. E. H.; VAN HESSEN, A. J., Modeling phoneme perception I: Categorical perception. *Journal of the Acoustical Society of America*, v.92, n.4, p. 1841-1855, Out.1992.

SHARMA A.; KRAUS N.; MCGEE T.; CARRELL T.; NICOL T. Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential. *Electroencephalography and Clinical Neurophysiology*, v. 88, n. 1, p. 64-71, Jan-Fev.1993.

SHARMA, A.; DORMAN; M. Exploration of the perceptual magnet effect using the mismatch negativity auditory evoked potential. *The Journal of the Acoustical Society of America*, v.104, n.1, p.511-517, Jul. 1998

SHARMA, A.; DORMAN, M. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *Journal of the Acoustical Society of America*, v. 106, n. 2, p. 1078-1083, Ago. 1999.

SHARMA, A.; DORMAN, M. Neurophysiologic correlates of cross-language phonetic perception *The Journal of the Acoustical Society of America*, v. 107, n. 5, p. 2697-2703, Maio.2000.

SINEX, D.G.; McDONNARD, L. P. Average discharge rate representation of voice onset time in the chinchilla auditory nerve. *The Journal of the Acoustical Society of America*, v. 83, n. 5, p.1817-1827, Maio. 1988.

STUDDERT-KENNEDY, M.; LIBERMAN, A.; HARRIS, K; COOPER, F. Motor theory of speech perception: a reply to Lane's critical review. *Psychological Review*, v. 77, n. 3, p. 234-249, Mai. 1970.

SUSSMAN, E.; RITTER, W; VAUGHAN, H. Predictability of stimulus deviance and mismatch negativity. *Neuroreport*, v. 9, n. 18, p. 4167-4170, Dez. 1998.

TERVANIEMI, M.; MEDVEDEV, S. V.; ALHO, K.; PAKHOMOV, S. V.; ROUDAS, M. S.; VAN ZUIJEN, T. L.; NÄÄTÄNEN, R. Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping*, v.10.n. 2, p.74-79, Jun. 2000.

VAN HESSEN, A. J.; SCHOUTEN, M. E. H. Categorical perception as a function of stimulus quality. *Phonetica*. v. 56, n. 1-2, p. 56-72, Jan.-Jun.1999.

VAN HESSEN, A. J; SCHOUTEN, M. E. H.; Modeling phoneme perception II: A model of stop consonant discrimination. *Journal of the Acoustic Society of America*, v. 92, n. 4, p. 1856-1868, Out.1992.

WETZELS, L. Mid vowel neutralization in Brazilian Portuguese. *Cadernos de Estudos Lingüísticos*, Campinas, v. 23, p. 19-55, Jul.-Dez. 1992.

WINKLER, F. The space-time unity and the representation problem. In: COMPUTING ANTICIPATORY SYSTEMS, 1999, Woodburg. Conference Proceedings. Woodburg: American Institute of Physics, p. 131-141.

WINKLER, I.; LEHTOKOSKI, A.; ALKU, P.; VAINIO, M.; CZIGLER, I.; CSEPE, V.; AALTONEN, O.; RAIMO, I.; ALHO, K.; LANG, H.; IIVONEN, A.; NÄÄTÄNEN, R. Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, v. 7, n. 3, p. 357-369, Jan. 1999a.

WINKLER, I.; KUJALA, T.; TIITINEN, H.; SIVONEN, P.; ALKU, P.; LEHTOKOSKI, A.; CZIGLER, I.; CSEPE, V.; ILMONIEMI, R.; NÄÄTÄNEN, R. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, v. 36, n. 5, p. 638-642, Set. 1999b.

ANEXO A - Proporções das respostas “u”, “o” e “ɔ” na tarefa de classificação

TABELA A 1

F1	Resposta		
	u	o	ɔ
250	0,9958	0,0042	0,0000
260	1,0000	0,0000	0,0000
270	0,9917	0,0083	0,0000
280	0,9833	0,0167	0,0000
290	0,9875	0,0083	0,0042
300	0,9833	0,0167	0,0000
310	0,9583	0,0417	0,0000
320	0,9542	0,0458	0,0000
330	0,8375	0,1625	0,0000
340	0,8750	0,1250	0,0000
350	0,6833	0,3167	0,0000
360	0,5750	0,4250	0,0000
370	0,4167	0,5833	0,0000
380	0,1792	0,8208	0,0000
390	0,0792	0,9208	0,0000
400	0,0625	0,9333	0,0042
410	0,0250	0,9750	0,0000
420	0,0292	0,9583	0,0125
430	0,0083	0,9792	0,0125
440	0,0000	0,9875	0,0125
450	0,0125	0,9667	0,0208
460	0,0042	0,9708	0,0250
470	0,0083	0,9417	0,0500
480	0,0042	0,9333	0,0625
490	0,0083	0,8625	0,1292
500	0,0000	0,7792	0,2208
510	0,0042	0,6167	0,3792
520	0,0000	0,4667	0,5333
530	0,0042	0,3625	0,6333
540	0,0042	0,2125	0,7833
550	0,0000	0,1875	0,8125
560	0,0000	0,1083	0,8917
570	0,0000	0,0750	0,9250
580	0,0000	0,0750	0,9250
590	0,0083	0,0542	0,9375
600	0,0000	0,0375	0,9625
610	0,0000	0,0333	0,9667
620	0,0000	0,0375	0,9625
630	0,0000	0,0167	0,9833
640	0,0083	0,0333	0,9583
650	0,0000	0,0083	0,9917
660	0,0042	0,0250	0,9708

ANEXO B - Tabelas de d'

TABELA B 1

Valores de d' no continuum [u-o]

F1	d'			
	Classificação	2AFC	4I2AFC	4I2AFC'
250	0,1439	-0,0316
260	0,1157	-0,2132	0,4064	1,2406
270	0,0507	0,3312
280	0,2514	0,7447	0,7504	1,7194
290	0,3097	1,0166
300	0,7908	1,2654	1,1747	2,3625
310	0,4862	0,6963
320	1,2559	1,0821	1,0106	2,0583
330	1,1284	1,2700
340	1,5838	1,5171	1,2609	2,4587
350	1,7369	1,6164
360	1,8801	1,5402	1,3890	2,6420
370	1,5508	1,4809
380	0,9832	0,9688	1,0382	2,1580
390	0,4579	1,0927
400	0,4662	0,9494	1,2461	2,4656
410	0,2514	0,7447
420	0,0588	0,6076	1,5618	2,8413
430	0,0568	0,8479
440	-0,1179	0,3628	1,5704	2,8118
Média	0,6720	0,8945	1,1409	2,2758

TABELA B 2Valores de d' no continuum [0-3]

F1	d'			
	Classificação	2AFC	4I2AFC	4I2AFC'
440	0,3711	0,3589	1,2435	2,3349
450	0,4156	0,4136
460	0,8167	0,5229	0,9552	1,9963
470	0,8983	0,8604
480	1,2987	0,4304	0,9712	2,0737
490	1,3313	0,9279
500	1,2872	0,7240	1,3682	2,6161
510	1,3807	1,0341
520	1,1038	1,2700	1,4798	2,7313
530	1,1734	1,4770
540	0,6350	1,1583	1,3455	2,4903
550	0,4986	0,8486
560	0,2840	1,1837	1,2361	2,3483
570	0,2679	1,0736
580	0,2595	0,7339	1,1863	2,2773
590	0,1095	0,5585
600	0,1680	0,3542	0,9880	2,0127
610	-0,0028	0,1773
620	0,2453	0,5058	1,1262	2,2186
630	-0,0242	0,3796
Média	0,6259	0,7496	1,1900	2,3099

ANEXO C – Scripts criados para a síntese das vogais

Gerar fonte

```
Create PitchTier... 145_85 0.00 0.30
Add point... 0.0 145
Add point... 0.3 85
To PointProcess
To Sound (phonation)... 22100 1.0 0.05 0.7 0.03 3.0 4.0
Create IntensityTier... risefall25 0.0 0.30
Add point... 0.000 0
Add point... 0.025 75
Add point... 0.275 75
Add point... 0.300 0
plus Sound 145_85
Multiply... yes
```

Gerar *Continuum* de vogais variando F1.

form continuum

```
positive f_1_minimo_(Hz) 250
positive f_1_maximo_(Hz) 700
positive tamanho_do_passo_(Hz) 10
positive f_2_(Hz) 900
sentence nome_da_fonte 145_85_int
```

endform

for i from 1 to ((f_1_maximo' - f_1_minimo')/tamanho_do_passo') + 1

```
f1_locus = 'f_1_minimo' + 'tamanho_do_passo' * (i - 1)
```

```
Create FormantTier... 'f1_locus' 0.00 0.30
```

```
Add point... 0.3 'f1_locus' 50 'f_2' 100 2500 150 3500 200 4500 300 5500 400 6500
500 7500 600 8500 700 9500 800
```

```
plus Sound 'nome_da_fonte$'
```

```
Filter
```

```
Rename... 'f1_locus'
```

```
select FormantTier 'f1_locus'
```

```
Remove
```

endfor

for x from 1 to ((f_1_maximo' - f_1_minimo')/tamanho_do_passo') + 1

```
n = 'f_1_minimo' + 'tamanho_do_passo' * (x - 1)
```

endfor