

FACULDADE DE PSICOLOGIA E DE CIÊNCIAS DA EDUCAÇÃO  
UNIVERSIDADE DO PORTO

## **ALFABETIZAÇÃO E PERCEÇÃO DA FALA**

CONTRIBUIÇÃO EXPERIMENTAL PARA O ESTUDO DOS EFEITOS  
DO CONHECIMENTO DA ESCRITA EM ASPECTOS DO PROCESSAMENTO DA  
LINGUAGEM FALADA

Dissertação de candidatura ao grau de  
Doutor elaborada sob a orientação do  
Professor José MORAIS da  
Universidade Livre de Bruxelas

MARIA DE SÃO LUÍS DE VASCONCELOS DA FONSECA E CASTRO SCHÖNER

1988

O trabalho apresentado nesta dissertação foi subsidiado por:

Fundação Calouste Gulbenkian;

Junta Nacional de Investigação Científica e Tecnológica;

"Laboratoire de Psychologie Expérimentale" (Universidade Livre de Bruxelas).

Esta dissertação é constituída por trabalho original que, salvo as excepções que a seguir se referem, não foram objecto de publicação prévia.

A Experiência II é apresentada em Castro, S. L. F. e Morais, J. (1987). Ear differences in illiterates. *Neuropsychologia* 25, 409-417. A Experiência III é apresentada em Morais, J., Castro, S. L. F., Scliar-Cabral, L., Kolinsky, R. e Content, A. (1987). The effects of literacy on the recognition of dichotic words. *The Quarterly Journal of Experimental Psychology* 39A, 451-465.

A parte que me coube na realização daqueles artigos consistiu na recolha e análise dos dados, preparação dos estímulos e redacção do artigo (Experiência II) e, em geral, discussão dos resultados e questões teóricas envolvidas, elaboração do produto final.

## AGRADECIMENTOS

O trabalho que se segue só foi possível mediante o apoio e colaboração de muitas pessoas e instituições; para todos os que, de uma maneira ou outra, contribuíram para o seu bom termo, vai a minha profunda gratidão.

Gostaria de expressar publicamente o meu agradecimento ao Professor José Morais, pela sua orientação atenta e perspicaz, e por me ter ajudado a dar os primeiros passos no meio científico internacional; ao Professor Paul Bertelson, pela hospitalidade com que sempre fui recebida no Laboratório de Psicologia Experimental da Universidade Livre de Bruxelas; ao Conselho Científico da Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto, que me libertou do serviço docente pelo período de 36 meses; ao Professor Bártolo Campos e ao Centro de Psicologia da Universidade do Porto, pelo encorajamento e apoio às iniciativas de investigação; aos Professores Alvin Liberman e Scott Kelso, por me terem recebido nos laboratórios que dirigem, e me terem dado acesso a todas as facilidades de equipamento aí existentes; à Dr<sup>a</sup> Teresa Barreiros Leal, pela ajuda inestimável nas várias fases desta investigação.

Agradeço muito especialmente o apoio recebido da Fundação Calouste Gulbenkian, desde o início desta investigação; agradeço também à Junta Nacional de Investigação Científica e Tecnológica o subsídio concedido.

A minha gratidão é particularmente viva para com as pessoas e instituições que me facultaram o acesso aos participantes nas experiências: na impossibilidade de as nomear exaustivamente, saliento apenas a colaboração da Direcção Geral de Educação de Adultos (Zona Norte), na pessoa da Dr<sup>a</sup> Maria Pinto, e os esforços incansáveis de Amadeu Almeida e Irene Oliveira.

A execução gráfica e preparação final desta dissertação ficam a dever-se ao trabalho empenhado e efectivo da Maria Amélia Santos e Rosa Maria Costa, cuja colaboração sempre pronta refiro com gratidão. Gostaria ainda de agradecer os préstimos do serviço de apoio à docência da Faculdade, particularmente os de Fernando F. Vieira.

Gostaria de salientar com apreço o apoio dos familiares e amigos que ao longo dos últimos anos criaram condições propícias ao desenvolvimento deste trabalho.

Finalmente, gostaria de dedicar este trabalho aos sujeitos participantes nas experiências realizadas, em especial àqueles que desinteressadamente abriram as portas de suas casas e me concederam o seu tempo e esforços.

Macieira, Verão de 1988

## APRESENTAÇÃO GERAL

Este trabalho procura elucidar se o conhecimento da escrita afecta alguns aspectos do processamento da fala. A relação entre fala e escrita do ponto de vista dos processos psicológicos envolvidos tem sido objecto de investigação experimental, nomeadamente no contexto da aprendizagem da leitura. Em geral, e de acordo com uma certa prioridade ontogenética, aborda-se a questão de como o tratamento da escrita envolve processos usados para tratar a fala. No presente trabalho toma-se a outra faceta da possível interacção entre fala e escrita: concentrar-nos-emos na percepção de fala e tentaremos desvendar aí a influência do conhecimento da escrita. A estratégia de investigação usada é simples: comparar o desempenho de adultos letrados e iletrados, e de crianças letradas e pré-letradas, em tarefas de percepção da fala. A investigação com sujeitos que não dispõem de uma representação gráfica das palavras, e para quem a representação consciente da fala como uma sequência de fonemas não é evidente, constitui um meio privilegiado para distinguir os processos de percepção da fala que são universais e em certa medida automáticos, dos processos dependentes da experiência e em larga medida opcionais.

Tanto quanto sabemos, não estão disponíveis na literatura muitos dados sobre o problema da relação entre fala e escrita como ela é aqui encarada. Em consequência, optámos por apresentar conceitos e questões no domínio da percepção da fala na primeira parte desta dissertação, e discutir aprofundadamente as observações relevantes para o nosso problema no contexto de uma segunda parte, devotada às experiências que realizámos.

Começaremos pois por tentar caracterizar o estímulo fala, recorrendo a noções de fonética articulatória e acústica psicologicamente relevantes (capítulo I), após o que debateremos questões gerais e específicas

de percepção da fala (capítulo II). Interessar-nos-á reflectir sobre a arquitectura do presumível sistema de percepção de fala no ser humano; não se trata pois de uma revisão exaustiva da literatura sobre percepção de fala, que pela sua assombrosa vastidão excederia o âmbito do presente trabalho. Ao longo do terceiro e último capítulo da parte introdutória, far-se-á a revisão dos trabalhos devotados a um fenómeno de percepção da fala, a fusão fonológica, que será objecto de uma das experiências por nós realizadas.

A nossa contribuição experimental para o estudo dos efeitos da alfabetização na percepção da fala será objecto da segunda parte deste trabalho. Começaremos por discutir a questão do suporte biológico da linguagem falada em relação com o domínio da escrita, tal como ela tem sido colocada na literatura; apresentaremos a este propósito duas experiências sobre efeitos de lateralidade auditiva em letrados e iletrados (capítulo IV). Trataremos a seguir o eventual papel da alfabetização na identificação de palavras faladas em condições difíceis (capítulo V); com vista ao seu esclarecimento, foram analisados os erros na identificação de palavras apresentadas dicoticamente a adultos letrados e iletrados (experiência III) e a crianças pré-letradas e crianças e jovens letrados (experiência IV), e ainda na identificação de palavras mascaradas por ruído (experiência V). No capítulo VI discutiremos o papel da experiência com a linguagem na identificação de categorias fonéticas e apresentaremos uma experiência sobre a nomeação de estímulos diferindo quanto ao vozeamento, com adultos letrados e iletrados (experiência VI). Finalmente, abordaremos questões mais específicas do conhecimento da escrita, como a representação ortográfica, e a sua influência na percepção de palavras faladas, através de uma experiência sobre a fusão fonológica em letrados e iletrados (capítulo e experiência VII).

A forma final de apresentação das experiências, que propomos ao leitor, é fruto de elaborações sucessivas sempre no sentido de uma maior concisão; procurou-se deliberadamente aligeirar o texto definitivo, libertando-o de detalhes menos importantes, que se remeteram, em parte, para os Anexos.

O conjunto das experiências realizadas permitiu distinguir aspectos da percepção de fala que sofrem a influência da alfabetização, de outros que lhe são imunes, como se refere na Conclusão.

Convém desde já esclarecer o que entendemos por "sujeitos iletrados". Trata-se de indivíduos cujo desconhecimento da escrita se deve a razões sócio-económicas, e não a dificuldades da aprendizagem. Foi tomado cuidado especial para não incluir sujeitos que não dominassem a escrita por problemas de ordem intelectual, ou por qualquer desordem física ou emocional. De facto, isto não foi tão difícil como poderá parecer; estes sujeitos habitavam na sua maioria em meio rural, onde eram membros activos e válidos; no processo de busca dos sujeitos, tirou-se partido do conhecimento "popular" de que a pessoa "x" ou "y" não tinha aprendido a ler porque era "fraca da cabeça". Essas pessoas obviamente não participaram nas experiências.

Trabalhámos apenas com sujeitos femininos, por ser difícil encontrar iletrados masculinos na faixa etária considerada (até um máximo de cerca de 65 anos, devido a problemas auditivos associados à idade). As mulheres participantes nas nossas experiências tinham em geral a mesma história: a escola longe, a família numerosa, os pais de opinião que "a escola é para homens" e precisando de ajuda nas lides domésticas e agrícolas; repetidamente ouvimos que "o meu pai achava que era perigoso aprender a ler e escrever, pois assim podia escrever bilhetinhos aos namorados".

Para isolar tanto quanto possível o efeito da alfabetização de outras experiências de vida, trabalhámos sempre com um grupo adicional de mulheres pertencentes a um meio sócio-económico semelhante ao das iletradas, mas que tinham conseguido frequentar a escola pelo período obrigatório de 4 anos. Designaremos estes sujeitos por semiletradas.

## ÍNDICE

Agradecimentos	iii
Apresentação geral	iv
Índice	viii

## PARTE I - A PERCEPÇÃO DE FALA

CAPÍTULO I. CARACTERIZAÇÃO DO ESTÍMULO FALA: CONTRIBUTOS DA FONÉTICA ARTICULATÓRIA E ACÚSTICA	1
1. Termos básicos e convenções	2
2. Noções de fonética articulatória e produção da fala	6
2.1. Introdução	6
2.2. Mecanismos respiratórios	6
2.3. Mecanismos laríngeos	7
2.4. Mecanismos articulatórios	10
2.5. Coarticulação e adaptação	18
3. Noções de fonética acústica	21
3.1. Introdução	21
3.2. Estrutura acústica da fala	22
3.3. Teoria fonte-filtro da produção de fala	24
3.4. Índices acústicos de categorias perceptivas	28
4. Variação acústica dependente do contexto	33
4.1. O problema da segmentação	33
4.2. O problema da invariância	34
CAPÍTULO II. O SISTEMA DE PERCEPÇÃO DE FALA	41
1. Extensão e limites do conceito e alguns problemas sobre o processo	41
2. Percepção categórica da fala	47
2.1. Generalidades	47
2.2. A descoberta da percepção categórica	48
2.3. Extensões da descoberta inicial	53
2.4. Factores explicativos: modo fala, aprendizagem e sensibilidade auditiva	58
3. O modo fonético	66
3.1. Os efeitos de contexto	66

3.2. Integração perceptiva de índices acústicos	67
3.3. Dissociação de perceptos fala e não-fala para o mesmo acontecimento acústico	69
3.4. Integração audiovisual	70
4. Teoria motora e modularidade	71
4.1. Asserções fundamentais (formulação inicial e teoria revista)	72
4.2. Modularidade e percepção da fala	75
 CAPÍTULO III. FUSÃO FONOLÓGICA: EXPLORAÇÕES EXPERIMENTAIS E TEÓRICAS	 80
1. Características gerais da fusão fonológica	81
2. Explorações experimentais	82
2.1. Introdução; fala natural e fala sintética	82
2.2. Variáveis de tipo acústico-auditivo	84
2.3. Variáveis de tipo linguístico	87
3. Explorações teóricas	96
3.1. A fusão fonológica no quadro de outros fenómenos de fusão de sons	96
3.2. A questão das diferenças individuais	104
4. Conclusão	121
 PARTE II - CONTRIBUIÇÃO EXPERIMENTAL PARA O ESTUDO DOS EFEITOS DA ALFABETIZAÇÃO NA PERCEPÇÃO DA FALA	
 CAPÍTULO IV. LATERALIZAÇÃO DAS FUNÇÕES LINGUÍSTICAS	 126
1. Introdução geral	124
1.1. Variáveis experienciais e lateralização	125
1.2. O papel da aprendizagem da leitura e da escrita	130
2. Audição dicótica de palavras dissimilares e similares por adultos letrados, semiletrados e iletrados	142
2.1. Introdução	142
2.2. Experiência I	146
2.3. Experiência II	157
3. Conclusão geral	165

CAPÍTULO V. RECONHECIMENTO DE PALAVRAS FALADAS EM CONDIÇÕES DIFÍCEIS	167
1. Introdução geral	167
1.1. Influência do conhecimento da escrita em certas tarefas com linguagem falada (sujeitos letrados)	168
1.2. Análise da fala em sujeitos iletrados	170
1.3. Consciência da estrutura fonética da fala e processos perceptivos	175
2. Identificação de palavras dicóticas - a: Análise dos erros de adultos letrados, semiletrados e iletrados (Experiência III)	177
2.1. A técnica de audição dicótica como meio de abordar os processos perceptivos	177
2.2. Fusão de traços fonéticos	179
2.3. Predições quanto ao tipo de erros nos vários grupos	181
2.4. Método	182
2.5. Resultados	182
2.6. Discussão	188
3. Identificação de palavras dicóticas - b: Análise dos erros de crianças pré-letradas e crianças e jovens letrados (Experiência IV)	193
3.1. Objectivos	193
3.2. Método	194
3.3. Resultados e discussão	195
3.4. Conclusão	207
4. Inteligibilidade de palavras em ruído: comparação de letradas, semiletradas e iletradas (experiência V)	211
4.1. Introdução	211
4.2. Método	214
4.3. Resultados e discussão	218
4.4. Resumo e conclusão	237
5. Discussão geral e conclusão	240
CAPÍTULO VI. IDENTIFICAÇÃO DE ESTÍMULOS DIFERINDO QUANTO AO VOZEAMENTO	247
1. Introdução	248

1.1. Ontogénese e percepção de sons da fala	248
1.2. Identificação de categorias fonémicas quanto ao vozeamento	250
2. Nomeação de <i>continua</i> de vozeamento velar e labial por letradas, semiletradas e iletradas (Experiência VI)	255
2.1. Introdução	255
2.2. Método	258
2.3. Resultados	263
2.4. Discussão e conclusão	269
 CAPÍTULO VII. INTEGRAÇÃO DE INFORMAÇÃO FONOLÓGICA E ASPECTOS ESPECÍFICOS DO CONHECIMENTO DA ESCRITA	 278
Fusão fonológica e ortografia (experiência VII)	278
1. Introdução	278
2. Método	284
2.1. Sujeitos	285
2.2. Estímulos	288
2.3. Procedimento	289
3. Resultados	289
3.1. Identificação em monaural	289
3.2. Apresentação dicótica de pares fundíveis	289
4. Discussão e conclusão	292
 CONCLUSÃO	 307
REFERÊNCIAS	314
ANEXOS	338

## PARTE I - A PERCEPÇÃO DA FALA

### CAPÍTULO I

#### CARACTERIZAÇÃO DO ESTÍMULO FALA: CONTRIBUTOS DA FONÉTICA ARTICULATÓRIA E ACÚSTICA

O objectivo deste primeiro capítulo é esclarecer em que consiste o estímulo fala, tanto quanto possível de um ponto de vista funcional. Encaramos a fala como uma categoria de acontecimentos físicos com um determinado estatuto perceptivo: o que distingue a fala de outras estimulações a que o ser humano responde não são só propriedades da estimulação independentes da resposta, é também o facto do percepto ser de um tipo específico, que poderíamos subsumir no termo genérico fala. A ideia de que para compreender o comportamento face a uma determinada estimulação é importante caracterizá-la do ponto de vista do que é funcionalmente relevante, e não só do que é conhecido e saliente para o observador, tem vindo a ser explicitada em vários contextos, psicológicos e outros (*e.g.*, Gibson, 1979; Schöner e Kelso, 1988). Ela pode parecer um tanto circular se se entende que o importante é explicar a resposta (pois para a explicar apelaríamos ao estímulo definido ele próprio em termos de resposta); todavia se se pretende compreender os processos psicológicos envolvidos na mediação entre estimulação e resposta, é inevitável e crucial que se atenda explicitamente ao que vai ser relevante nessa mediação.

Um dos aspectos mais *sui generis* da fala é que ela constitui uma classe de acontecimentos que até há bem pouco tempo só existiam se fossem produzidos pelo ser humano (exceptuando o caso curioso de algumas aves como o papagaio). Este facto pode sem dúvida por si só suscitar especulações filosóficas sobre a relação íntima entre produzir e perceber\* fala (quem percebe é quem produz, e quem produz é quem percebe); em termos de caracterização da fala enquanto estímulo, ele sugere uma via intuitivamente atraente, a de se averiguar como ela é produzida.

Começaremos pois por abordar noções básicas de fonética articulatória e produção da fala, passando em seguida para uma caracterização da fala em termos acústicos. Não se trata de uma revisão do estado actual da pesquisa no domínio da produção de fala, nem de um exame detalhado da fonética acústica, que ultrapassariam largamente o âmbito deste trabalho. Procuraremos apenas esclarecer noções-chave daqueles domínios, indispensáveis à compreensão do que é a percepção da fala em geral, e de certos aspectos das nossas experiências em particular.

Nem sempre nos referiremos apenas a características que são funcionalmente relevantes em termos perceptivos. Esse é justamente um problema que a investigação em ciência da fala tem procurado resolver, e que está ainda em aberto. Discutiremos neste contexto o problema dos índices acústicos na percepção da fala, e a questão da invariância.

## 1. Termos básicos e convenções

É conveniente explicitar alguns termos básicos e convenções que serão usados ao longo deste trabalho. Começemos pela conhecida trilogia, fonema, alofone e fone.

\* Usaremos perceber no sentido estrito de perceber, e não de compreender.

Um fonema é um conjunto de sons da fala que numa determinada língua serve para indicar diferenças de sentido. A título exemplificativo, tomemos as consoantes portuguesas "p" e "b". Trata-se de dois fonemas, pois "pata" e "bata" têm significados distintos. Se um inglês tenta dizer aquelas palavras em português, ele provavelmente pronunciará o "p" com aspiração e dirá [p<sup>h</sup>atα]. Um falante português estranhará a pronúncia, mas entenderá o mesmo significado: o "p" aspirado não é um fonema da língua portuguesa. É-o noutras línguas, como o hindi e o thai; nesta última existem as seguintes palavras: [bàa], ombro, [pàa], floresta e [p<sup>h</sup>àa], dividir (Borden e Harris, 1982). Um fonema é pois uma entidade abstracta, só passível de concretização num contexto linguístico específico. A descrição do sistema e padrões de sons de uma determinada língua constitui a Fonologia.

Segundo o contexto, um mesmo fonema pode ser realizado através de sons diferentes. Compare-se o som do "l" em "lar" e em "mal": o primeiro é como que mais claro (é realizado com a ponta da língua tocando a zona alveolar), e o segundo tem uma tonalidade mais profunda (o pós-dorso da língua eleva-se em direcção ao véu palatino). Todavia, a não ser por esforço de análise, esta distinção passaria despercebida, e é irrelevante de um ponto de vista pragmático. Chamam-se alofones às variantes de um mesmo fonema, como o "l" alveolar [l] e o "l" velarizado [ɫ]. Os alofones de um fonema constituem uma família de sons muito parecidos, cuja variação é dependente do contexto e irrelevante semanticamente.

Alguns investigadores da ciência da fala usam um termo adicional, o fone, para se referirem aos sons propriamente ditos dos fonemas (e.g., Borden e Harris, 1982, p. 9). Neste sentido, alofone e fone são conceitos muito próximos, mas não equivalentes. Um alofone define-se em relação a um fonema, enquanto um fone descreve um som da fala concreto; o primeiro

remete a uma entidade linguística (a categoria da qual é membro), enquanto o segundo reenvia antes para noções acustico-perceptivas.

O fone é aliás definido por alguns linguistas (*e.g.*, Chomsky e Halle, 1968) e psicolinguistas da tradição generativa (Gleitman e Rozin, 1977; Liberman e Studdert-Kennedy, 1978) em termos essencialmente perceptivos: são os segmentos da mensagem acústica que identificamos como pertencentes à categoria abstracta de "sons da fala", esses segmentos sem significado que ouvimos como vogais e consoantes. Cada fone ou segmento fonético pode ser descrito por uma lista de especificações que definem as suas propriedades articulatórias e acústicas, chamadas traços fonéticos\*.

A noção de traço fonético está aparentada com o conceito linguístico de "traço distintivo", desenvolvido por N. S. Trubetzkoy em 1958 e R. Jakobson em 1962, que constitui a base teórica da fonologia generativa. O traço distintivo baseia-se na ideia de que o fonema não é um todo indivisível, pode antes ser decomposto em características componentes que servem para distinguir de outros fonemas. Estas características ou traços definem as relações estruturais do sistema fonológico (para estudos da fonologia do português nesta perspectiva, cf. Mateus, 1975; Pardal, 1977).

No contexto dos trabalhos sobre percepção e produção da fala, o traço fonético é uma noção importante quer para definir o estímulo fala, quer para tentar entender os processos de percepção e produção; não reenvia necessariamente para a matriz fonológica abstracta da língua.

Apesar da distinção entre Fonologia (o estudo do sistema de sons da língua) e Fonética (o estudo de sons da fala) ser nítida em termos logico-formais, em termos psicológicos nem sempre é claro se um determinado processo deveria ser apelidado de fonético ou fonológico, ou se a entidade

\* "Traço" no sentido de conjunto de características ou "feature".

relevante para a sua caracterização é o fone ou o fonema. De facto, na literatura psicológica recorre-se por vezes quase indistintamente àqueles termos. Ao longo deste trabalho, utilizaremos o termo genérico segmento, quando não for possível ou relevante esclarecer se se trata de segmento fonético ou fonémico. Noutras situações preferiremos usar o classificativo fonético sempre que subsista dúvida sobre o estatuto fonológico do fenómeno em questão: de facto, o fonológico é necessariamente fonético, mas o fonético não é necessariamente fonológico.

Sempre que se queira registar um fonema ou transcrever conjuntos de fonemas pertencentes a palavras ou frases, é de convenção fazê-lo entre barras oblíquas: *e.g.*, /b/, /lar/, /mal/. Este tipo de transcrição é denominado transcrição fonémica; ela é usada quando só interessa representar as variações dos sons que têm implicações semânticas. Quando, pelo contrário, é importante registar mais fielmente as variações do som recorre-se à transcrição fonética. Esta é feita sempre no interior de parêntesis rectos, *e.g.* [laɾ], [mal]. As transcrições fonéticas variam quanto ao grau de precisão com que definem os sons, sendo as menos detalhadas vulgarmente designadas por transcrições largas ou latas, e as mais exactas por restritas. A escolha de um ou outro tipo depende dos objectivos. As transcrições fonéticas aqui apresentadas usam o alfabeto proposto pela "International Phonetic Association", IPA, de que se apresentam em Anexo os símbolos relevantes para o português (cf. Anexo 1/IPA).

## 2. Noções de fonética articulatória e produção da fala

### 2.1. Introdução

O modo fácil, directo e sem esforço como pronunciamos uma palavra ou frase contrasta fortemente com a complexidade das operações requeridas para transformar a corrente de ar expelido dos pulmões nos padrões acústicos da fala. Com efeito, a produção dos mais simples vocábulos exige a acção conjugada de mecanismos de tipos diferentes: respiratórios, laríngeos ou de fonação, articulatórios e de ressonância. Passaremos a seguir uma breve revista a estes mecanismos, concentrando-nos naqueles que são utilizados pela família de línguas indo-europeias, especificamente o português. Assim, entre outros, não discutiremos a produção de consoantes por fluxo aéreo glótico (em vez do pulmonar), nem a aspiração (característica das línguas germânicas).

### 2.2. Mecanismos respiratórios

Quer a inspiração, quer a expiração são modificadas em situação de fala relativamente à situação de repouso. Em vez do carácter automático do ciclo respiratório, quando falamos exercemos um certo grau de controlo voluntário: para emitir sem interrupção uma frase mais longa, aumentamos o volume inspiratório (através da acção extra dos músculos que promovem a elevação do esterno e das costelas), e prolongamos a duração da expiração (através de contracções musculares activas, tanto inspiratórias como expiratórias). Enquanto em repouso cada ciclo respiratório consiste em aproximadamente 40% de inspiração e 60% de expiração, e o volume de ar expelido é de 10% da capacidade vital, durante a fala e/ou canto aquela relação transforma-se em 10% (inspiração) e 90%

(expiração), e o volume de ar sobe para 25% da capacidade vital, ou mesmo 40% quando se fala muito fortemente.

Da pressão do ar saído dos pulmões depende em grande parte a intensidade, e por vezes a frequência, da emissão vocal. Este ar em movimento passa pelos brônquios e traqueia e chega à laringe.

### 2.3. Mecanismos laríngeos

#### 2.3.1. Vozeamento

Na laringe situa-se um dos mecanismos fundamentais para a produção da fala, as cordas vocais. Estas encontram-se por detrás da cartilagem tiroideia na direcção anterior-posterior, e são constituídas por várias estruturas musculares (processos musculares, processos vocais, ligamento vocal) e a membrana mucosa que os envolve. O espaço horizontal entre as cordas vocais é denominado glote, sendo a pressão que aí é exercida pelo ar vindo dos pulmões chamada pressão subglótica ( $P_s$ ). É desta  $P_s$  que depende a intensidade de vibração das cordas vocais, vibração essa sucintamente designada por vozeamento.

É sabido já desde os finais do século passado que as "cordas" vocais não se comportam como cordas de um instrumento musical, como se fossem postas em vibração por um impulso nervoso de efeito análogo às dedilhações de um guitarrista. Conforme proposto por Helmholtz e Müller na teoria aerodinâmica mioelástica da fonação, é hoje aceite que as cordas vocais são activadas pela corrente de ar expelido pelos pulmões (daí o "aerodinâmica"), vibrando a determinadas frequências em virtude da sua elasticidade ("mioelástica").

Quando o ar expulso dos pulmões encontra as cordas vocais fechadas ou quase (pela acção dos músculos aritnóides), e a pressão subglótica é maior que a supraglótica, a força exercida nas cordas vocais desencadeia a sua abertura ou abdução. A expulsão do ar através da constrição glótica suscita por sua vez uma diminuição da pressão nas paredes interiores da glote. Trata-se do efeito Bernouilli: o aumento de velocidade de um gás ou líquido resulta na diminuição da pressão exercida por esse gás. Devido a esta diminuição de pressão entre as cordas vocais, estas fecham-se novamente (adducção), obstruindo a passagem do ar e recriando um aumento de pressão subglótica. Por sua vez, esta promove a abertura das cordas vocais, que de novo resulta no seu fecho e assim sucessivamente. Assim, em virtude destas duas forças aerodinâmicas, a pressão subglótica e o efeito Bernouilli, é instaurado um ciclo respectivamente de abertura ou abdução, e fecho ou adducção das cordas vocais, que periodicamente se repete. Esta vibração das cordas vocais cria ciclos de compressão e rarefacção nas moléculas do ar em movimento, mais exactamente uma onda acústica quase-periódica e complexa, constituída por uma frequência fundamental ( $f_0$ ) e as suas várias harmónicas (frequências múltiplas da  $f_0$ ).

O número de aberturas ou de oclusões da glote por segundo constitui a  $f_0$ . Esta frequência depende da massa, elasticidade e tensão das cordas vocais. No sexo masculino, estas variam entre 17 a 24 milímetros (mm) e vibram a frequências entre 80 e 200 Herz (Hz). No sexo feminino, as cordas vocais são mais curtas (13-17 mm) e a frequência é mais elevada ( $f_0 > 200$  Hz), o mesmo acontecendo nas vozes infantis ( $f_0 > 300$  Hz). O carácter grave ou agudo duma determinada voz depende pois da frequência de vibração das cordas vocais. Evidentemente, cada falante pode variar dentro de certos limites a altura tonal da sua voz, tornando-a mais "grossa" ou mais "fina": por exemplo, ao entoar uma pergunta a altura tonal sobe na

parte final, e numa frase declarativa ela desce. Na maioria dos casos, isto consegue-se através da variação de tensão das cordas vocais executada por acções musculares várias, nomeadamente do músculo crico-tiroideu. Porém um aumento de pressão subglótica, que normalmente resulta num aumento de intensidade, pode também provocar um aumento de altura tonal se não forem feitos ajustamentos musculares compensatórios.

### 2.3.2. Sons sonoros e surdos

Até aqui temos tratado de sons da fala em que há vibrações das cordas vocais. Estes sons chamam-se vozeados ou sonoros, e são constituídos por todas as vogais e algumas consoantes (e.g., /b, d, g, z, v/). No entanto, há sons da fala cuja fonte não é vibração das cordas vocais; são todos consonânticos e designam-se por não vozeados ou surdos (e.g., /p, t, k, s, f/). Nestes sons, as cordas vocais estão bem abertas, impedindo a pressão subglótica de exceder a supraglótica, e o ar vindo dos pulmões vai criar ruídos na cavidade oral. Deparamo-nos pois com uma distinção básica dos sons da fala, consoante a fonte da onda acústica: já a nível laríngeo pela vibração das cordas vocais, ou a nível do tracto oral por mecanismos articulatorios que a seguir veremos. Os primeiros são sonoros, os segundos surdos.

Na fala corrente, consoantes surdas seguem-se a sons sonoros, e vice-versa, a ritmos extraordinariamente velozes. Isto exige aberturas glóticas rápidas para interromper o vozeamento, que são levadas a cabo pela acção dos músculos posteriores crico-aritnóides. A alternância surdo/sonoro é de difícil execução, e os falantes têm tendência em manter o vozeamento sempre que tal seja aceitável linguisticamente. Por exemplo, a desinência de plural "s" é pronunciada diferentemente em "os pares e os bares": no

primeiro caso, onde a consoante seguinte é surda, dizemos [ ʃ ], também surdo; no segundo, em que a consoante seguinte é sonora, dizemos [ ʒ ], também sonoro.

## 2.4. Mecanismos articulatorios

### 2.4.1. Tracto vocal (oral e nasal). Articuladores

Acima da laringe e cordas vocais, o ar passa pelo tracto vocal, *i.e.*, as cavidades constituídas pela faringe e boca (tracto oral) e nariz (tracto nasal). O tracto vocal desempenha um papel fundamental na produção da fala, pois não só permite a passagem do ar e/ou onda sonora, como determina as suas próprias características. *Grosso modo*, ele constitui uma cavidade de ressonância que modifica o espectro da onda acústica. Estas modificações dependem da forma do tracto vocal, ou seja, ocorrem através da formação de cavidades de ressonância particulares que amplificam certas frequências e atenuam outras. Por isso, é usual conceber o tracto vocal como um tubo acústico, constituído por uma secção oral com aproximadamente 17 centímetros (cm) de comprimento e uma área transversal variando de zero (se houver oclusão completa) a 20 cm<sup>2</sup>, e uma secção nasal com aproximadamente 12 cm de comprimento e um volume de 60 cm<sup>3</sup>.

A forma do tracto vocal é alterada por acção dos articuladores. Na cavidade oral, distinguem-se os articuladores inferiores e os superiores. Os primeiros são constituídos pelo lábio inferior e língua, dividida em ápice ou ponta, dorso (anterior, médio e posterior) e raiz; os segundos são constituídos pelo lábio superior, dentes (nomeadamente os incisivos), palato rígido, subdividido em alvéolos, palato anterior, médio e posterior, e véu palatino ou palato mole, subdividido em anterior, posterior e úvula. O véu palatino é

**QUADRO 1:** A relação entre fonte, ressoador, tipo de som e maneira de articulação.

FONTE	RESSOADOR	SOM	MANEIRA	EXEMPLOS
Cordas Vocais	Tracto Vocal	Quase- -periódico (P)	Vogais Ditongos Aproximantes Nasais	[a] [i] [ui] [ou] [j] [w] [m] [n]
Tracto Vocal	Tracto Vocal	Aperiódico (Ap)	Oclusivas Surdas Fricativas Surdas	[p] [k] [f] [s]
Cordas Vocais e Tracto Vocal	Tracto Vocal	Misto (P+Ap)	Oclusivas Sonoras Fricativas Sonoras	[b] [g] [v] [z]

uma formação muscular que pode ser elevada e pressionada contra a parede posterior da faringe, impedindo a saída de ar pelo nariz. Esta acção, a oclusão nasal ou naso-faríngea, ocorre em todos os sons orais; o tracto nasal só está aberto durante a emissão de sons nasais.

#### 2.4.2. Sons consonânticos. Maneira e lugar de articulação

Para compreender a acção dos articuladores é preciso considerar pelo menos dois critérios: a maneira e o lugar de articulação.

As maneiras de articulação são definidas através do grau de obstrução oferecido à passagem do ar, uma obstrução completa produzindo as consoantes oclusivas e uma obstrução mínima produzindo as vogais. Como no caso das vogais os articuladores nunca se aproximam muito e a passagem do ar é relativamente desobstruída, distinguem-se basicamente as articulações consonânticas das articulações vocálicas. Nestas, a única fonte sonora são as vibrações quase periódicas das cordas vocais, que vão sofrer modificações devidas apenas às ressonâncias do tracto vocal. Nas consoantes entra em jogo outra fonte sonora, de natureza aperiódica, no local do tracto vocal onde é feita a obstrução (cf. Quadro 1).

Nas consoantes oclusivas há um fecho completo dos articuladores, que interrompe a passagem do fluxo aéreo por um tempo relativamente curto. Este fecho pode ocorrer em vários locais, e ser ou não acompanhado por oclusão naso-faríngea, dando assim origem a diferentes consoantes. Se os dois lábios se juntarem formando uma oclusão bilabial, a reabertura (ou libertação) desta oclusão pode produzir três sons diferentes, /p, b, m/. No último som, o tracto nasal está em comunicação com o oral, enquanto nos dois primeiros o som só passa através do tracto oral. A distinção entre /p/ e /b/ é apenas de vozeamento: em /b/ ocorrem as vibrações periódicas

originadas nas cordas vocais juntamente com as não periódicas originadas na oclusão, enquanto em /p/ a fonte do som reside exclusivamente no tracto oral. Sons como estes, com o mesmo lugar de articulação, são chamados homorgânicos.

Se a oclusão se der mais atrás, com a ponta da língua encostada aos incisivos ou ao palato alveolar, os sons resultantes são /t, d, n/; os dois primeiros orais, um surdo, outro sonoro, e o último nasal (e sonoro, tal como /m/). São as oclusivas e nasais dental-alveolares.

Finalmente, se a parte posterior da língua for encostada ao véu palatino, as consoantes resultantes serão /k, g, ŋ/, oclusivas velares, análogas às anteriores. Em português há ainda a oclusiva nasal palatal [ɲ], em, por exemplo, "sonho" ['soɲu].

As consoantes fricativas constituem um outro grupo de sons caracterizado por uma maneira de articulação específica: dois articuladores são aproximados por forma a obstruir parcialmente o fluxo aéreo. Esta obstrução parcial resulta na produção de turbulência, *i.e.*, vibrações aperiódicas das moléculas aéreas. Consoante o local em que se efectua a obstrução parcial, são produzidas diferentes fricativas. Se o lábio inferior se aproximar dos incisivos superiores, resultam os sons /f, v/, conforme a vibração das cordas vocais esteja, respectivamente, ausente ou presente. No Quadro 2 pode observar-se o efeito do lugar de articulação na produção de outras fricativas.

Uma terceira maneira de articulação consiste em aproximar os articuladores de modo a que a constrição seja relativamente aberta e não haja produção de turbulência. Aos sons assim produzidos reserva-se o termo

aproximantes\* (Ladefoged, 1982). Eles ocorrem em palavras como "quando" ['kwɔ̃du], "muito" ['muʃtu], "lado" ['ladu], "julho" ['ʒuɫu]. Os últimos dois fonemas, /l/ e /ʎ/ são vulgarmente designados por laterais, porque a obstrução é feita centralmente e o ar passa pelos lados: o ápice da língua aproxima-se da zona dental-alveolar em /l/, e da zona palatal em /ʎ/, impedindo o ar de sair pelo centro do tracto oral e permitindo apenas passagens laterais. Os dois primeiros fones, /j, w/, são centrais, com articulação respectivamente palatal e labio-velar. A sua semelhança com os sons vocálicos /i, u/ é evidente, e traduz-se por espectrogramas muito parecidos com o dos ditongos. Em ambos, "aproximantes" centrais e ditongos, se observam transições de formante (cf. 3.2) muito graduais; devido a esta característica os aproximantes centrais são chamados semivogais.

Resta-nos considerar dois tipos particulares de articulações consonânticas, os trilos ou vibrantes (em inglês, "trills") e os batimentos ("flaps"). Exemplos de trilos ocorrem em "ramo" ['ɾɔmu], "carro" ['kaɾu]. O fonema /ɾ/ tanto pode ser realizado pondo o ápice da língua contra a zona alveolar ou contra os incisivos superiores (dental-alveolar), como pondo o pós-dorso da língua contra a úvula (como uvular). Trata-se em qualquer dos casos de colocar um articulador próximo de um outro, de modo a que a corrente de ar o(s) ponha em vibração. Batimentos ou "flaps" podem ser observados em "caro" ['kaɾu], "pera" ['peɾɔ]. São realizados por um rápido golpe de língua contra a região alveolar. Genericamente, um batimento é pois caracterizado pelo desferimento de um golpe rápido de um articulador contra um outro\*\*.

\* Alguns autores denominam esta categoria "semivogais" e/ou "glides". Outros reservam estes termos para subcategorias dentro da classe geral dos aproximantes. Esta última alternativa é adoptada por Ladefoged (e.g., 1982, p.10), cuja classificação seguiremos.

\*\* Seguimos mais uma vez o esquema classificatório de Ladefoged (1982). Outros autores referem-se a estas duas consonantes como vibrantes (e.g. Cintra, 1987), o [ɾ] como vibrante de batimento único e o [R] como vibrante de batimentos múltiplos (cf. Barbosa, 1983; p. 170-171).

Uma visão de conjunto dos vários tipos de sons consonânticos é dada na carta das consoantes, onde por convenção se dispõe verticalmente a maneira de articulação e horizontalmente o lugar (cf. Quadro 2, onde estão presentes as categorias relevantes para línguas indo-europeias).

### 2.4.3. Sons vocálicos

Enquanto os sons consonânticos se caracterizam pela existência de aperiodicidade, as vogais consistem numa onda acústica quase periódica originada pela vibração das cordas vocais. Determinante do timbre de cada vogal oral ou nasal são as posições da língua e dos lábios, que modificam a forma do tracto oral sem o obstruir nem criar pontos de grande constrição.

As vogais podem ser classificadas *grosso modo* consoante a posição do local mais elevado da língua: anteriores ou palatais /i, e, ê/, centrais /α, a/ e posteriores ou velares /u, o, ɔ/. Considerando agora cada uma destas regiões, verifica-se que o grau de abertura varia conforme a altura da língua: em /i, u/ a abertura é mínima, em /ɛ, ɔ/ máxima, em /e, o, α/ média\*. Estes dois critérios combinados, como que uma dimensão longitudinal quanto ao local de elevação acrescida por uma dimensão em altura, serviram para definir o clássico quadrilátero das vogais. Ocasionalmente, esta representação geométrica é referida como o triângulo das vogais, considerando o /a/, vogal central de abertura máxima, como vértice.

Finalmente, um terceiro gesto articulatório relevante para a produção das vogais é o arredondamento dos lábios. Este varia entre mínimo em /i/ e máximo em /u/, sendo usado em conjunção com os dois critérios

\* Em terminologia anglo-saxónica, este critério de abertura mínima vs máxima é expresso pela distinção "high" vs "low".

anteriores para caracterizar as várias vogais. *E.g.*, o /i/ é definido como vogal anterior, de abertura mínima e não redonda, o /u/ como posterior, abertura mínima e redonda.

O grau de arredondamento dos lábios é sem dúvida o gesto articulatorio mais evidente e acessível à análise fenomenológica da produção das vogais. As restantes mudanças do tracto oral são mais subtis, o que é evidenciado, por exemplo, na dificuldade que os surdos têm em pronunciar vogais contrastantes. Apesar disto, o procedimento típico dos fonéticos para descrever as vogais de diferentes línguas ou dialectos consistia em usar aqueles critérios articulatorios e inscrever as vogais observadas no conhecido quadrilátero.

Para facilitar esta tarefa, D. Jones propôs em 1957\* as chamadas vogais cardinais, pontos de referência fixos e em relação aos quais se poderiam localizar as vogais reais. As vogais cardinais (1) e (5) são definidas em termos estritamente articulatorios. A primeira é produzida com os lábios estendidos e a língua tão à frente e tão elevada quanto possível sem criar fricção audível. Isto resulta numa vogal muito parecida com o [i], mas mais extrema; é representada também por [i]. A vogal cardinal (5) é articulada com os lábios em posição neutra, e a língua tão atrás e tão baixa quanto possível. Este som é semelhante à vogal [ɑ] no inglês "father, hot", distinguindo-se do [a] português por a posição da língua ser mais posterior. Todas as outras vogais cardinais são definidas em termos relativos, como pontos perceptivamente equidistantes.

Não é este o lugar para entrarmos numa discussão pormenorizada sobre o sistema das vogais cardinais. Interessa salientar que

---

\* A referência clássica, que não consultámos directamente, é a obra de Daniel Jones intitulada "An Outline of English Phonetics", publicada em 1957 pela editora Heffer, Cambridge, Mass.

ele permite apreender uma propriedade fundamental do conjunto de sons vocálicos: é que as variações que permitem distinguir uma vogal de outra são mais de tipo quantitativo e contínuo do que categórico e discreto. É possível começar a dizer /i/ e mudá-lo gradualmente até /u/, mas não se pode variar continuamente de /b/ até /z/. Enquanto para representar as consoantes se utiliza uma matriz dividida em categorias discretas, para as vogais recorre-se a uma área definida por dimensões contínuas. Diferentes línguas "escolheram" locais específicos daquela área como categorias vocálicas, de acordo com o que se poderia chamar o "princípio de separação perceptiva suficiente" (Ladefoged, 1982, p. 205). Não são conhecidas línguas onde todas as vogais se aglomerem numa parte restrita do quadrilátero, por exemplo, sendo todas anteriores; em geral, cada língua tem aproximadamente o mesmo número de vogais anteriores e posteriores, eventualmente com mais uma ou duas centrais. Um caso exemplificativo simples é o do espanhol, que contrasta apenas cinco vogais como que maximizando o grau de separação mútua.

Outro aspecto que se evidencia progressivamente ao longo da apresentação do mapa das vogais é que para o definir é preciso recorrer a critérios auditivos para além dos articulatórios. De facto, estudos radiográficos da posição da língua ao produzir séries de vogais mostraram que ela não é fielmente representada no quadrilátero, nomeadamente quanto à altura (cf. Borden e Harris, 1982). Hoje é sabido que o mapa das vogais dá apenas indicações grosseiras quanto aos gestos articulatórios subjacentes. Por isso, é preferível falar no grau de abertura das vogais do que na maior ou menor altura da língua\*.

\* A ambiguidade é mais aparente na literatura anglo-saxónica, onde se usam indistintamente os termos "height" e "vowel height". Esta altura das vogais não deverá ser interpretada como traduzindo directamente a altura da língua.

Mesmo se os critérios do mapa vocálico não reproduzem fielmente as dimensões articulatórias, eles gozam de indiscutível realidade perceptiva. Sem dúvida que os fonéticos clássicos, treinados em fazer discriminações auditivas finas, não representaram as vogais num sistema de coordenadas arbitrárias. Estudos mais recentes de fonética acústica (e.g., Peterson e Shoup, 1966) revelaram uma clara correspondência entre as distinções tradicionais e as características acústicas das vogais. Concretamente, o critério de abertura máxima a mínima tem como correlato acústico a frequência do primeiro formante, que é tanto mais elevada quanto maior for a abertura. A distinção anterior *vs* posterior corresponde estreitamente a uma variável acústica relacional, *i.e.*, à diferença de frequência entre o segundo e o primeiro formantes. Mais à frente trataremos com mais pormenor alguns aspectos de fonética acústica. Por ora, acentue-se que o sistema de relações entre vogais, que os fonéticos clássicos julgavam ser de natureza fisiológica, é também de natureza acústica e permanece um poderoso instrumento para a compreensão dos sons vocálicos (Ladefoged, 1982).

#### 2.4.4. Ditongos

O carácter potencialmente contínuo da produção vocálica é evidenciado pela existência dos ditongos. Um ditongo é uma "vogal" cuja qualidade varia dentro de uma única sílaba; ou, se se preferir uma definição articulatória, trata-se de uma articulação de tonalidade vocálica que se realiza alterando a ressonância do tracto vocal no interior da mesma sílaba.

Para clarificar a noção de ditongo é conveniente introduzir a distinção entre vogais brandas e tensas. As primeiras são mais curtas e com ajustes articulatórios menos pronunciados, enquanto às segundas são mais

longas e exigem movimentos de língua mais extremos, estando muito próximas das semivogais. Na palavra "cerveja", por exemplo, o primeiro "e", [ɐ], é brando e o segundo é tenso. Nas vogais tensas, a forma do tracto vocal pode manter-se idêntica ao longo da emissão, ou variar: se se puxar a língua para cima e à frente, o som resultante é [ei], se para cima e trás, [eu]. Os ditongos constituem assim o subgrupo das vogais tensas em que há uma mudança do tracto vocal. Em geral, a mudança é em direcção à vogal de abertura mínima, anterior ou posterior. Em português, os ditongos partem das vogais tónicas em direcção a /i/ ou /u/, acompanhados ou não pela abertura do tracto nasal. Os ditongos em fala corrente são realizados como vogal mais semivogal, ou vice-versa, (e.g., "pai", [paj]), o que ilustra bem a proximidade entre as chamadas vogais tensas e as semivogais.

## 2.5. Coarticulação e adaptação

Temos até aqui considerado apenas segmentos fonémicos e fonéticos em isolação. Porém, na fala corrente em geral, e mesmo na emissão de segmentos como as consoantes oclusivas, os vários sons sucedem-se rapidamente ou são pronunciados como totalidades. No caso óbvio das oclusivas, elas têm sempre de ser articuladas em conjunto com uma vogal; e mesmo os outros sons da fala fluente não são produzidos em sequência estrita, como se os fonemas fossem contas de um rosário. Pelo contrário, articulações adjacentes sobrepõem-se no tempo, *i.e.*, são coarticuladas.

Um dos exemplos mais conhecidos é o da sílaba /stru/: ao pronunciá-la, os lábios estão arredondados desde o início, antecipando a vogal /u/ três fonemas mais cedo (comparem-se os movimentos articulatorios usados para produzir /stru/ e /stra/). Este fenómeno, em que

um dos articuladores não envolvido na produção de um determinado som se movimenta para a posição correspondente ao som seguinte, é chamado coarticulação antecipatória. Durante uma oclusão, por exemplo, os restantes articuladores estão numa posição que antecipa a vogal: ao dizer /títu/, as duas oclusões são afectadas pelo carácter anterior *vs* posterior das vogais seguintes. No primeiro /t/, tudo o que é necessário, é efectuar uma oclusão dental-alveolar com a ponta da língua; o restante corpo daquela está livre para especificar o /i/ seguinte, portanto puxado acima e à frente; no segundo /t/, o restante corpo da língua está pelo contrário puxado atrás em antecipação do /u/. Possivelmente o primeiro /t/ será realizado como dental e o segundo como alveolar.

Estes exemplos ilustram como a fala não é produzida serialmente através da concatenação de posições estáticas e rigidamente determinadas. Cada fonema é globalmente especificado pelo vozeamento, maneira e lugar de articulação que lhe correspondem, estando os restantes mecanismos articulatorios livres para preparar fonemas adjacentes. Por isso, é conveniente caracterizar um fonema em termos de configurações articulatorias-alvo ou "posições-alvo", que podem ser realizadas de modos ligeiramente diferentes consoante as configurações articulatorias antecedentes e subsequentes (Stevens, 1978).

A coarticulação pode ser encarada como um caso especial do fenómeno mais geral que, com Borden e Harris (1982, p. 126), designaremos por adaptação. Esta é entendida como a variação dos movimentos articulatorios segundo o contexto, não apenas num determinado momento (simultaneidade de gestos para diferentes segmentos, coarticulação), como ao longo de um intervalo temporal relativamente longo (influência dos gestos articulatorios precedentes nos subsequentes): ao produzir a configuração B

parte-se da configuração A, e não de um qualquer ponto de partida neutro e invariante.

O caso mais flagrante é o da articulação das vogais em fala normal rápida: ela é tão pouco diferenciada que as posições-alvo podem nem mesmo ser atingidas antes de a língua se mover para o fonema seguinte. Este fenómeno é tão característico da emissão da fala rápida (e, em geral, do aumento de velocidade da fala) que ele é designado na literatura com um termo próprio, o "undershoot" ou redução articulatória, a que corresponde um "undershoot" ou redução acústica\*. Outro exemplo clássico de adaptação foi sugerido por P. Macneilage: falar com o cachimbo na boca. Todos os gestos articulatórios necessários à emissão de fala têm de se adaptar à nova posição relativa dos dentes e maxilares.

Um caso extremo de adaptação ocorre na assimilação, quando o efeito dos segmentos vizinhos é tão poderoso que chega a mudar as características das articulações alvo dos fonemas adjacentes. É o caso já referido da desinência de plural /s/, pronunciada como surda ou sonora consoante o fonema seguinte fôr surdo ou sonoro. Trata-se de uma assimilação de vozeamento. Um exemplo de assimilação ao lugar de articulação pode ser observado em "clamor", em cuja realização o /l/ é velarizado por influência do /k/. Compare-se com o /l/ de "plano".

Estes fenómenos (coarticulação, adaptação e assimilação) têm consequências acústicas e perceptivas que serão discutidas adiante. Os exemplos relatados ilustram como a fala resulta de uma coordenação espaço-temporal complexa, determinada pelas interacções entre as várias configurações-alvo, e não pela especificação rígida de um movimento

\* "Undershoot" significa literalmente atirar, visando não o centro do alvo, mas mais abaixo ou ao lado. Neste contexto, é um modo metafórico de exprimir que o "alvo" não é realmente atingido. Um termo análogo em português poderia ser "esbatimento".

invariante e específico correspondente a cada fonema. A actualização dos vários segmentos fonéticos depende do contexto. Assim, não há correspondência termo-a-termo nem relação simples entre a descrição de uma língua em termos de fonemas e a descrição da fala em termos de posições articulatórias. As variações alofónicas de um fonema podem resultar de um esbatimento de produção "visando" o mesmo alvo (como no exemplo do /t/ alveolar e do /t/ dental), mas também da existência de diferentes alvos consoante, por exemplo, a posição na palavra (como o /l/ alveolar e o /l/ velarizado). Entre fonemas e configurações articulatórias há sem dúvida uma correspondência; só que não é termo-a-termo. Fonemas são unidades abstractas e invariantes que servem para descrever as línguas; porém aquela invariância não é mantida ao nível da concretização da língua na fala corrente, onde variações de produção dependentes do contexto são regra. Grande parte da investigação sobre produção da fala tem procurado resolver o problema de como a variabilidade de produção pode transportar invariantes perceptivos e linguísticos (cf. e.g., Perkell e Klatt, 1986).

### 3. Noções de fonética acústica

#### 3.1. Introdução

Uma das maneiras de tentar compreender o fenómeno fala, talvez a mais intuitiva, consiste em esclarecer como são produzidos os vários sons nela usados. Outra, complementar da primeira, é a de inquirir sobre as características dos sons propriamente ditos: estes dependem evidentemente de especificidades de produção, mas são certamente passíveis de um outro tipo de caracterização, de natureza acústica. Trata-se de, em detrimento da perspectiva "egocêntrica" de produtor/falante (Fonética Articulatória), adoptar a perspectiva "neutra" dos aparelhos de medida do produto/fala

(Fonética Acústica). A vantagem mais evidente desta segunda alternativa é a aplicação dos conhecimentos obtidos na produção artificial ou síntese da fala, e ainda em máquinas de leitura para invisuais (para uma história detalhada, cf. Cooper, Gaitenby e Nye, 1984). Outras adicionais são as de possibilitar a caracterização de sons difíceis de distinguir em termos de movimentos articulatórios, ou explicar por que certos sons são sistematicamente confundidos.

### 3.2. Estrutura acústica da fala

Pioneiro na pesquisa sobre acústica da fala foi H. von Helmholtz (1821-1894). A ele se deve a formulação de alguns princípios fundamentais, hoje conhecimento "standard" do principiante em ciência da fala: que os "sopros" de ar escapando através das cordas vocais são a fonte sonora da voz, que as componentes harmônicas desta fonte são sujeitas às ressonâncias das cavidades faríngeas, orais e nasais, e que as vogais se caracterizam por padrões de ressonância específicos. Ele próprio desenvolveu uma técnica para analisar as frequências componentes de um tom complexo, utilizando globos e tubos de vidro de diferentes ressonâncias (*i.e.*, que ampliam frequências próximas da sua frequência natural de ressonância e atenuam as restantes). Estes instrumentos ficaram conhecidos por ressoadores de Helmholtz. Ajustando firmemente um ressoador a um ouvido, e tapando o outro com cera de modo a impedir a passagem do som, em vez do som complexo original ouve-se um som cujas frequências são filtradas pelo ressoador; através da manipulação dos ressoadores, Helmholtz conseguiu identificar a frequência fundamental e as harmônicas da voz, bem como as principais ressonâncias do tracto vocal.

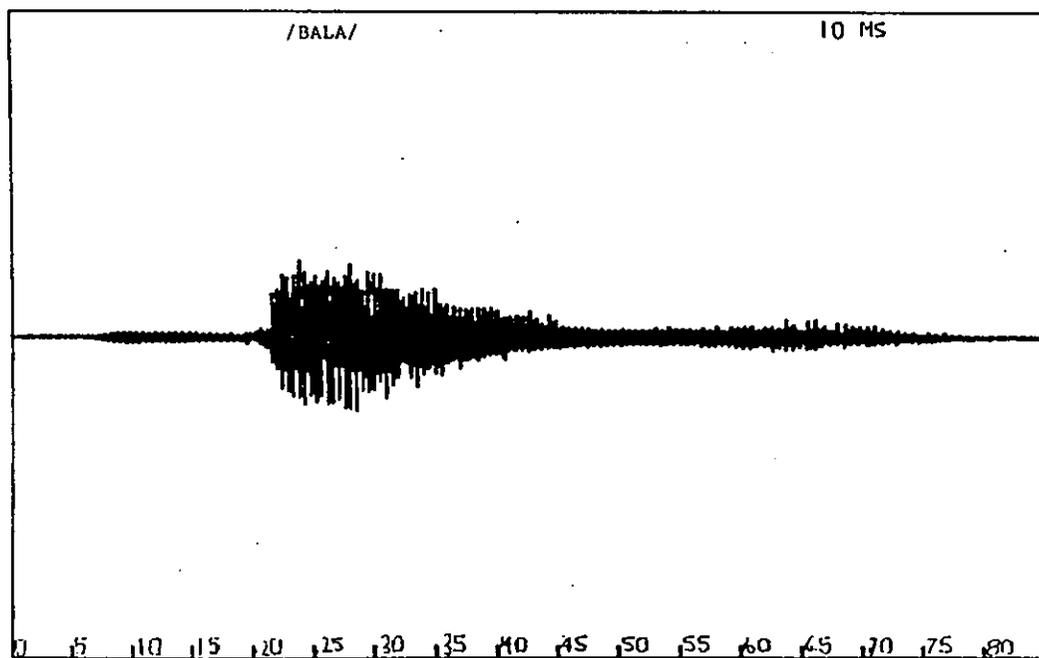
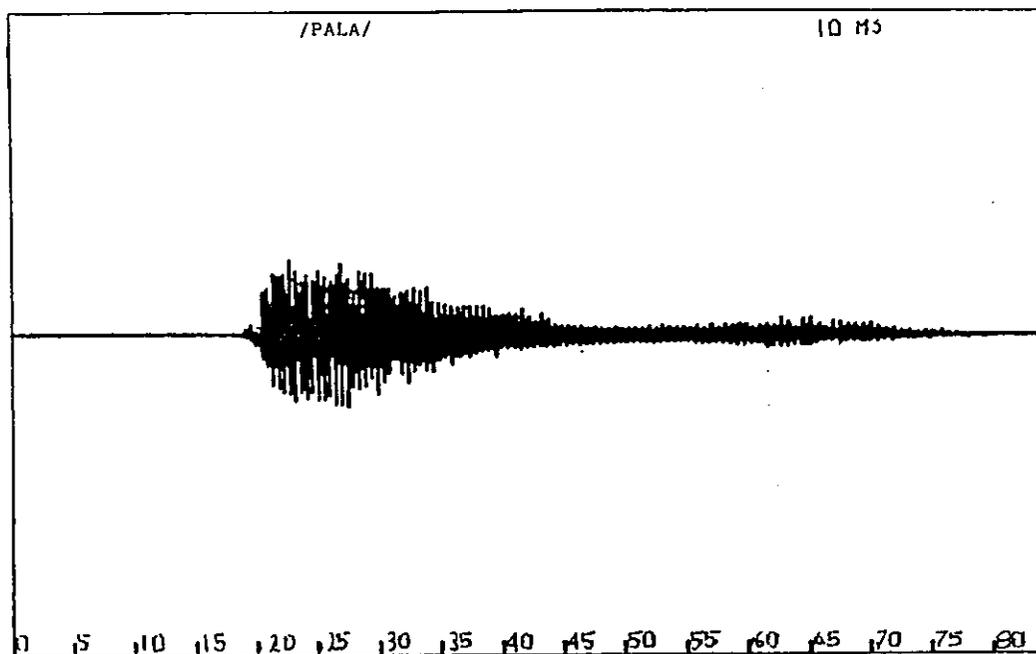
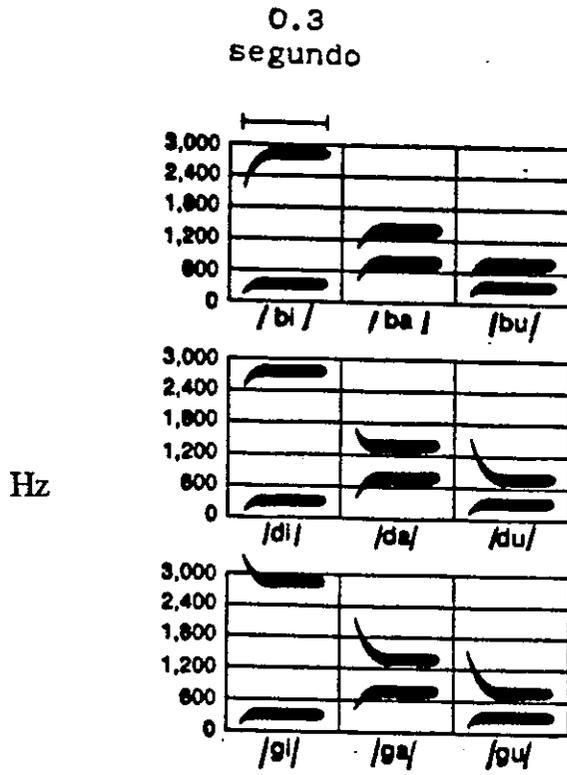


Figura 1: Os sonogramas de /pala/ e /bala/. Fala portuguesa digitalizada e analisada pelo sistema PCM dos Laboratórios Haskins.

Hoje, graças ao desenvolvimento tecnológico, é possível analisar automaticamente a estrutura acústica do som, e obter facilmente representações gráficas da fala segundo certas características acústico-temporais. Uma destas representações é o sonograma. Nele são registadas as variações de pressão das moléculas atmosféricas (em ordenada) num determinado período de tempo (em abcissa). O modo mais comum de o obter é através de um osciloscópio e/ou oscilógrafo. Na Figura 1 mostram-se os sonogramas das palavras "bala" e "pala". Eles permitem apreciar a forma da onda acústica; repare-se na existência de vibrações anteriores à explosão no caso de /bala/ e na sua inexistência em /pala/. Trata-se do correlato acústico da distinção fonémica entre /p/ e /b/ em português, língua em que as consoantes sonoras são "pré-vozeadas".

Outra maneira útil de descrever a estrutura acústica da fala é através do espectrograma: em abcissa está o tempo, em ordenada a frequência, sendo a intensidade de cada frequência indicada pelo grau de sombreamento. Um dos primeiros espectrogramas publicados foi o da frase "Speech we can see". A escolha desta frase traduz uma preocupação genuína por parte dos investigadores da época: de facto, por altura das primeiras tentativas em traduzir visualmente o estímulo acústico, na década de 40, esperava-se conseguir que pessoas surdas pudessem vir a usar o telefone. Infelizmente, veio a verificar-se que "ler" espectrogramas é bem mais difícil do que se julgava, por não haver correspondência invariante e termo-a-termo entre porções do espectrograma e fonemas (voltaremos a este ponto mais adiante).

A banda de frequências que mais contribui para a inteligibilidade da fala distribui-se quase simetricamente em torno dos 1900 Hz, com quatro a cinco oitavas entre 200 e 4000 Hz. Esta é justamente a região de maior sensibilidade na curva dos limiares absolutos do sistema auditivo humano.



**Figura 2:** Espectrogramas de sílabas sintéticas consoante-vogal, de dois formantes, percebidas como se indica. (Adaptado de Miller, 1981, p. 67)

Uma das características mais salientes dos espectrogramas são as barras mais escuras, correspondentes às zonas de maior intensidade de certas bandas de frequências. Estas zonas são chamadas formantes; convencionou-se referir os formantes (F) de determinado som numerando-os de baixo para cima. Na Figura 2, estão representadas sílabas sintéticas de dois formantes. É possível distinguir em cada formante uma zona em que ele se mantém mais ou menos constante, e outra em que é ascendente ou descendente. No primeiro caso, estamos em presença de um formante estável ("steady-state"), no segundo de uma transição de formante ("formant transition"). As frequências em que se situam os vários formantes resultam das características de ressonância do tracto vocal. A este propósito é conveniente fazer um breve exame dos mecanismos acústicos envolvidos na emissão da fala.

### 3.3. Teoria fonte-filtro da produção de fala

Segundo a teoria fonte-filtro da produção da fala ("source-filter theory"), proposta já em 1848 pelo fisiólogo J. Müller e posteriormente desenvolvida e elaborada por engenheiros de sistemas, uma emissão vocal engloba os dois processos seguintes: (1) acção de uma fonte sonora, as cordas vocais, que produzem uma onda acústica; esta onda constitui a função-fonte, com um determinado espectro, o espectro de origem; (2) transferência e filtragem daquela onda acústica pelas cavidades ressonadores do tracto vocal; nisto consiste a função-filtro, com picos de ressonância específicos. Estes picos de ressonância vão sobrepor-se ao espectro de origem, imprimindo-lhe a estrutura de formantes.

Vejamos um exemplo, extraído de Miller (1981, pp. 40-43). Suponhamos um falante masculino cuja voz tem uma  $f_0$  de 100 Hz, a

pronunciar a vogal neutra [ə]\*. A fonte sonora é uma vibração com  $f_0$  de 100 Hz e harmónicas de 200, 300, 400 Hz, etc; a energia acústica é maior na  $f_0$ , e decresce linearmente com o aumento de frequência. Este é o espectro acústico da fonte. Para pronunciar o som [ə], o tracto vocal tem uma forma que se aproxima de um tubo uniforme de 17 cm de comprimento e aberto numa extremidade. Esta característica aparece em radiografias feitas ao tracto oral de locutores a produzir este som. Um tubo acústico com aquela forma e comprimento tem frequências de ressonância a 500, 1500 e 2500 Hz\*\*. Quando o som da fonte sonora é filtrado por aquelas ressonâncias, as suas harmónicas próximas daquelas frequências vão ser amplificadas e as restantes atenuadas. O espectro resultante, o produto da função-fonte com a função-filtro, é percebido como [ə], quer seja originado por um falante humano, quer por meios artificiais.

Uma questão imediatamente se coloca: a de saber como esta mesma vogal será emitida por um falante com um comprimento do tracto vocal diferente (logo, com outras ressonâncias), por exemplo uma criança. Com a mesma forma do tracto vocal, e por analogia com o caso anterior, é fácil prever que os picos de ressonância e os formantes resultantes poderiam estar desviados à direita ( $f$  mais elevadas), mas manteriam o mesmo padrão: três picos de formantes, espaçados de 1:3. Inversamente, um mesmo falante pode cantar uma vogal idêntica ao longo da escala musical. Neste caso, a  $f_0$  da sua voz varia, mas não a forma do tracto vocal e

\* Esta vogal é designada por "schwa". Ela aparece em "about" [əbaut], "commit" [kəmit]; assemelha-se ao [ə] português diferindo deste apenas por ser mais central. É um dos fonemas mais estudados das línguas inglesa e francesa.

\*\* As frequências de ressonância de um tubo fechado numa extremidade são múltiplos ímpares da sua frequência de ressonância mais baixa. Esta, por sua vez, varia conforme o comprimento do tubo: a frequência de ressonância mais baixa tem um comprimento de onda 4 vezes o comprimento do tubo. O comprimento de onda (distância ocupada por um ciclo) depende inversamente da frequência da onda e directamente da velocidade de propagação, *i.e.*,  $\lambda=c/f$ . Assim, no exemplo referido, a frequência de ressonância mais baixa é de  $33500 \text{ cm}/(17/4)\text{cm}=500 \text{ Hz}$ .

a estrutura dos formantes; por isso, percebemos a mesma qualidade auditiva, e.g. [ə], apesar da variação de altura tonal. Estes exemplos ilustram a complexa relação entre estimulação sonora e percepto fala, e sugerem que para explicar a correspondência entre estes dois níveis é necessário invocar variáveis relacionais, como o padrão acústico, e não uma ou outra característica do som.

Escolhemos a vogal central [ə] por ela exemplificar a situação mais simples, a de um tubo acústico uniforme. Para outras vogais, anteriores ou posteriores, é preciso considerar as ressonâncias de tubos não uniformes. Por exemplo, em [i], como a língua é puxada à frente e acima, a cavidade oral é mais pequena e estreita do que a faríngea. Tal resulta num padrão de ressonâncias em que a distância entre  $F_1$  e  $F_2$  é grande, o primeiro de aproximadamente 300 Hz e o segundo de 2300 Hz. Em [u], devido ao arredondamento dos lábios, que prolonga ligeiramente a cavidade oral e reduz a sua abertura, as ressonâncias são mais baixas;  $F_1$  é de aproximadamente 300 Hz como em [i] (o dorso da língua está elevado nos dois casos), mas  $F_2$  é consideravelmente mais baixo,  $\approx$  900 Hz.

As relações entre características de ressonância do tracto vocal e produção de vogais foi extensamente estudada. Veja-se a clássica monografia de Chiba e Kajiyama publicada em 1941\*, que aplicaram leis acústicas de ressoadores à derivação das vogais, bem como a obra de Fant "Acoustical Theory of Speech Production", publicada pela primeira vez em 1960\*\*. Trata-se de um estudo pormenorizado sobre a acústica das vogais com base em medições de radiografias ao tracto vocal durante a sua emissão. A aplicação da teoria fonte-filtro à produção de consoantes não é tão evidente, pois é necessário considerar formas mais complexas do tracto

\* Não consultámos directamente esta obra, cuja referência é Chiba, T. e Kajiyama, M. (1941). *The Vowel, its Nature and Structure*. Tokyo: Kaiseikan Publ. Co.

\*\* Não consultámos directamente esta obra, publicada pela editora Mouton, de Haia.

vocal, bem como fontes sonoras múltiplas e aperiódicas. Discuti-la ultrapassaria o objectivo deste resumo simplificado de acústica articulatória.

Concluamos pois com duas notas. Primeiro, o espírito e as ideias fundamentais desta teoria desempenham um papel marcante na conceptualização e planeamento da investigação sobre produção e acústica da fala; ela capta indubitavelmente modos de funcionamento do sistema fonador, pelo menos dentro de certas condições (nomeadamente produção de vogais), e foi usada com sucesso em síntese de fala.

Segundo, uma nova perspectiva parece estar a despontar como alternativa à teoria anterior. De facto, a teoria fonte-filtro não explica certas não-linearidades do fluxo aéreo dentro e fora da cavidade oral: por exemplo, quando se observam dentro da cavidade oral certas frequências que estão ausentes à saída da voz, ou vice-versa, transformações típicas de um sistema multiplicativo e não aditivo. Trata-se de considerar o tracto vocal essencialmente como um instrumento de sopro, em que são criados sons devido à interacção da(s) corrente(s) de ar com as paredes envolventes. Em linguagem técnica, as paredes do tracto vocal funcionam como condições-fronteira, ou "boundary conditions", na determinação do comportamento do fluxo aéreo. O estatuto de "fonte sonora" deixa de estar limitado à "clássica" vibração das cordas vocais ou a uma constrição específica. Esta proposição centra-se na aplicação das ideias de dinâmica de fluidos aos comportamentos fonadores. Argui-se que é preciso considerar, além das fontes clássicas, os vórtices e jactos independentes da vibração glótica, gerados pela passagem turbulenta da corrente de ar pelas cavidades vocais (Kaiser, 1986).

### 3.4. Índices acústicos de categorias perceptivas

Apresentam-se a seguir os correlatos acústicos da hierarquia de ações fonadoras e articulatórias discutidas na secção anterior, cuja relevância perceptiva foi estabelecida experimentalmente. Estes correlatos acústicos devem ser encarados como indicações gerais e algo grosseiras, pois a forma concreta das variações acústicas depende do contexto, como se evidenciará ao longo da apresentação.

#### 3.4.1. Sonoro *vs* surdo: o "Voice Onset Time"

No espectrograma, um som sonoro distingue-se do surdo pela existência de estrias verticais correspondentes às vibrações glóticas. As oclusivas sonoras distinguem-se das surdas principalmente através do "Voice Onset Time", VOT, o tempo que medeia o início da vibração das cordas vocais e o relaxamento da oclusão (Lisker e Abramson, 1964). Nas sonoras, o segmento do espectrograma correspondente à oclusão é atravessado por um pequeno número de harmónicas de baixa frequência que traduzem o vozeamento, enquanto nas surdas aquele segmento tende a ser vazio (excepto se houver aspiração).

A distinção em termos de VOT é válida para várias línguas: diferentes oclusivas ocupam regiões distintas num *continuum* de VOT (Lisker e Abramson, 1967). O caso típico do inglês, idioma em que foi realizada a maioria dos estudos, é o seguinte: para /b, d, g/ em posição inicial, as cordas vocais começam a vibrar ligeiramente antes do relaxamento da oclusão, enquanto para as surdas aspiradas [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>] a vibração começa entre 20 a 100 milisegundos, ms, após aquele relaxamento. Em sílabas sintéticas constituídas por uma consoante e uma vogal, ou

sílabas CV, diferindo apenas no VOT, percebe-se /b, d, g/ se o VOT for inferior a 30 ms, e /p, t, k/ se for superior (Abramson e Lisker, 1967). Em português e espanhol a situação é mais extrema, na medida em que o vozeamento das sonoras se inicia cerca de 40 a 130 ms antes do relaxamento da oclusão. Como o VOT é negativo, alguns autores chamam-lhes pré-sonoras, ou pré-vozeadas.

Diferenças de VOT são concomitantes com outras variações acústicas. Uma oclusiva surda tem geralmente uma explosão mais forte, e, no inglês, um período de aspiração mais longo e um  $F_1$  menos intenso\*. Estas diferenças podem ser atribuídas ao facto de as cordas vocais estarem separadas por mais tempo, o que permite gerar uma maior pressão na cavidade oral (logo, explosão mais forte).

Este caso ilustra a asserção de que não existe uma correspondência termo-a-termo entre característica articulatória e propriedade acústica. Parece mais correcto falar numa dimensão articulatória e configuração acústica multidimensional resultante (e.g., Lisker, 1975; 1978-a e -b). Lisker e Abramson (1967) mostraram que estas propriedades acústicas são de facto usados na percepção do vozeamento (por exemplo, a diminuição da intensidade de  $F_1$  induz a percepção de surdo).

### 3.4.2. Vogais

As vogais constituem os sons da fala para os quais é mais fácil encontrar correlatos acústicos rapidamente identificáveis e relativamente invariantes: são os formantes estáveis. Cada vogal se distingue das outras

---

\* De facto, os dados a que nos referiremos provêm essencialmente de estudos na língua inglesa; na falta de prova em contrário, supomos serem extensíveis ao português, salvo nos casos evidentes como o da aspiração.

por um padrão de ressonâncias, logo de formantes, característico (para análise acústica das orais tónicas portuguesas, cf. Martins, 1971).

O primeiro formante está relacionado com a abertura da vogal: é mais baixo quando a abertura é mínima ([i, u]), mais elevado em [a, ɛ, ɔ]. Mais genericamente, a frequência de  $F_1$  aumenta quando a boca se abre e a constrição é posterior, e diminui quando a abertura labial e o alargamento da faringe são acompanhados por elevação da língua. A frequência de  $F_2$  é elevada quando a constrição é na cavidade oral, e baixa quando esta está mais aberta ou alongada. A projecção dos lábios, alongando o tracto vocal, contribui para diminuir as ressonâncias. Assim, quanto mais posterior é a vogal, menor é a diferença entre  $F_1$  e  $F_2$ . É devido a estas relações articulatório-acústicas que o quadrilátero vocálico pode ser reconvertido em parâmetros acústicos, com  $F_1$  como ordenada e a diferença entre  $F_1$  e  $F_2$  como abcissa.

Assim, a percepção de vogais depende essencialmente do padrão de frequências dos vários formantes; as frequências concretas dependem das propriedades do tracto vocal emissor. Em fala corrente a situação complica-se, pois a velocidade de articulação impede frequentemente que as posições-alvo sejam atingidas, o que se traduz num "esbatimento" acústico (e.g., Darwin, 1976).

### 3.4.3. Maneiras de articulação

Os ditongos e aproximantes têm uma estrutura de formantes semelhante à das vogais, com uma transição gradual para a vogal seguinte. É a velocidade lenta da transição dos  $F_2$  e  $F_3$  que indica a presença destes sons: /ba/ pode ser mudado para [ua] passando por [wa] apenas diminuindo

a velocidade de transição. As laterais têm formantes próximos dos 250, 1200 e 2400 Hz, sendo os mais elevados menos intensos.

As fricativas são facilmente identificáveis pela inexistência de estrutura de formantes; em vez desta, o padrão é de ruído aleatório, especialmente nas altas frequências. Dentro das fricativas, as altas frequências são mais intensas em [s, z, ʃ, ʒ] do que em [f, v, θ, ð]; por isso, os sons constantes do primeiro conjunto são chamados sibilantes. O mais sibilante é o [s], em que há relativamente pouca energia abaixo dos 4 KHz. Em [ʃ], o ruído desce até  $\approx$  2KHz, e nas não sibilantes estende-se até às baixas frequências (para um estudo sobre as fricativas portuguesas /s, ʃ, f/, ver Lacerda, 1980; 1982). Também na percepção de fricativas se verifica o efeito do contexto. Por exemplo, Mann e Repp (1980) verificaram que o mesmo ruído é percebido como /s/ antes de /u/, e como /ʃ/ antes de /i/.

A característica mais saliente das nasais é a diminuição de intensidade relativamente aos sons vizinhos. Isto acontece porque o som é absorvido pelas paredes e fossas nasais e também porque se trata de um tubo acústico maior, com uma banda mais larga de frequências de resposta\*. Além disto, as nasais são reconhecíveis através dos chamados "murmúrios" nasais, ressonâncias ou formantes a  $\approx$  250, 2500 e 3250 Hz. A distinção entre as várias nasais depende do contexto (nomeadamente das vogais adjacentes, cf. Darwin, 1976).

A categoria das oclusivas é reconhecível através da interrupção do padrão acústico que ocorre no breve período da oclusão. Por isso, num dos primeiros artigos sobre a percepção das oclusivas se escrevia sobre "o som do silêncio". O silêncio é mais evidente no caso das consoantes surdas do que nas sonoras; nestas, devido à vibração das cordas vocais, permanece um

\* A atenuação é maior para ressoadores que respondem a uma maior gama de frequências, e menor em ressoadores cuja faixa de resposta é estreita.



som de baixa frequência e baixa amplitude. Logo após esta descontinuidade, dá-se uma explosão abrupta que se reveste de características ligeiramente diferentes conforme se trate de sonoras ou surdas. Nas sonoras há um rápido aumento de  $F_1$  e surge a estrutura de formantes; nas surdas, ocorre ruído por alguns milissegundos. Nos dois casos, seguem-se imediatamente as transições para os formantes vocálicos que acompanham sempre a emissão de oclusivas. A transição de  $F_2$  parece constituir o índice principal quanto ao lugar (e.g., Cooper, Delattre, Liberman, Borst e Gerstman, 1952; cf. mais à frente a discussão sobre invariância). As características de cada oclusiva dependem do lugar de articulação e do contexto (cf. Mann e Repp, 1981), nomeadamente da vogal seguinte. Por exemplo, um ruído centrado em torno de 1440 Hz e seguido de transições estáveis correspondentes às vogais [i, a, u] é percebido como [pi, ka, pu] (Liberman, Delattre e Cooper, 1952).

#### 3.4.4. Lugar de articulação

Muito grosseiramente, pode dizer-se que, quanto mais anterior for o lugar de articulação, mais baixas serão as frequências de ressonância. Concretamente, porém, a situação é mais complicada, principalmente no caso das oclusivas. Observemos a Figura 2 (p. 23-a) com espectrogramas de sílabas CV artificiais, ouvidas como /bi, ba, bu/, /di, da, du/ e /gi, ga, gu/. Como a inspeção daquela figura revela, consoante o contexto vocálico variam também quer o primeiro, quer o segundo formantes de cada sílaba. São sempre ascendentes os  $F_1$ , mas os  $F_2$  tanto são ascendentes como descendentes. Todavia, os vários  $F_2$  de cada oclusiva parecem convergir ou divergir de uma determinada frequência. Esta frequência de onde parecem divergir os formantes chama-se *locus*.

O *locus* é pois o ponto aparente de origem dos formantes correspondentes a um determinado lugar de articulação (Ladefoged, 1982; p. 182). Era usual pensar-se que havia um *locus* relativamente específico para cada lugar de articulação ( /d/ ≈ 1800 Hz, /b/ ≈ 720 Hz, /g/ ≈ 3000 Hz). Hoje sabe-se que esta versão é algo simplista, pois o ponto de origem dos vários formantes depende também da vogal adjacente: é que a parte da língua que não está envolvida na execução da oclusão consonântica está já em parte preparada para a vogal seguinte. Aquando da libertação da oclusão, as frequências dos formantes são devidas à forma global do tracto vocal, variando portanto em função da vogal adjacente. De qualquer dos modos, é possível estabelecer generalizações aproximadas quanto aos *loci* de  $F_2$  e  $F_3$  conforme o local de articulação. Para as bilabiais, ambos são relativamente baixos; para as alveolares, o *locus* de  $F_2$  é de ≈ 1700-1800 Hz; para as velares,  $F_2$  e  $F_3$  partilham o mesmo *locus*, geralmente elevado.

#### 4. Variação acústica dependente do contexto

##### 4.1. O problema da segmentação

A adaptação em geral e a coarticulação em particular têm consequências acústicas que desde a invenção do espectrógrafo intrigaram leigos e especialistas. A sobreposição dos movimentos articulatorios para segmentos fonéticos adjacentes tem efeitos claramente visíveis nos espectrogramas. Os dois efeitos principais foram descritos já em 1948 por Joos (*in* Nittrouer e Studdert-Kennedy, 1986).

O primeiro refere-se à dificuldade em dividir o padrão acústico em partes que digam respeito a um só segmento fonémico. Por exemplo, se na Figura 2 se ouvisse a onda sonora até ao segundo 0.1, que ingenuamente se esperaria corresponder ao /b/, ouvir-se-ia um barulho não característico da

fala, designado onomatopaicamente por "chirp"\*. Dos segundos 0.1 a 0.3, todavia, ouvir-se-ia distintamente /i/ precedido por um ténue /b/. O que é identificável no padrão acústico é, grosseiramente, como que o "centro" de um segmento (zonas que são ouvidas mais como esse segmento do que como um outro). Uma secção ao acaso daquele padrão dirá provavelmente respeito a dois ou mais segmentos adjacentes, ou seja, será ouvida como sílaba. Joos denominou esta questão como o problema da segmentação. A inexistência de segmentos acústicos directamente correspondentes aos fones e fonemas é também vulgarmente designada por transmissão paralela (Liberman, Cooper, Shankweiler e Studdert-Kennedy, 1967): a cada instante, nomeadamente nas transições de formantes, o sinal acústico contém "informação" sobre mais do que um segmento, ou seja, é redundante. Esta redundância ajuda à inteligibilidade da fala, mas põe aos investigadores problemas de difícil resolução.

#### 4.2. O problema da invariância

Outro efeito da adaptação e coarticulação é que as características acústicas correspondendo a um determinado fone, ou mesmo sequências de fones, variam em função do contexto. Esta variação acústica dependente do contexto é uma consequência directa da variação de produção discutida anteriormente. Repare-se, de novo na Figura 2, como as transições de formantes que especificam cada oclusiva variam com o contexto vocálico. Note-se também que a variação da zona estável dos formantes correspondentes às vogais (quase) não depende do contexto consonântico. Esta situação nem sempre se verifica, apesar de ser conhecido que o contexto vocálico afecta mais o consonântico que o contrário. A descoberta da

\* Cremos preferível usar aquele termo corrente na literatura anglo-saxónica, em que o fenómeno foi descoberto, a inventar outro em português.

dependência acústica do contexto ficou conhecida na literatura como o problema da invariância.

Vejamos dois exemplos da falta de invariância. Compare-se a articulação de [it], em "vi-te", e [ɔt], em "mote". Por questão de adaptação, as posições dos articuladores para um fone determinam os movimentos necessários para produzir os fones seguintes. No primeiro caso, é necessária uma mudança relativamente pequena do tracto oral para efectuar a oclusão dental (devido à posição da língua para realizar o [i], elevada e à frente); isto traduzir-se-á provavelmente numa pequena transição de  $F_2$ . No [ɔt] de "mote", para efectuar a mesma oclusão a mudança do tracto vocal é mais pronunciada: elevação da língua, antes baixa e atrás, e encurtamento do tracto vocal antes alongado. Estas alterações deverão evidenciar-se a nível acústico por uma marcada transição ascendente de  $F_2^*$ .

O segundo exemplo refere-se à fala normal rápida e ao esbatimento articulatorio das vogais. De facto, este esbatimento aparece claramente a nível acústico: Lindblom (1963) mostrou, através de análises espectrográficas de vogais, que à medida que o ritmo da fala se torna mais rápido os formantes das várias vogais vão perdendo as suas características distintivas em torno do padrão da vogal central neutra "schwa" [ə]. É o esbatimento acústico traduzindo o esbatimento articulatorio.

Estas variações articulatorias e acústicas são intrínsecas, derivam do próprio processo de produção da fala. Algumas são de tal modo subtis que nem o ouvido mais treinado as consegue detectar, sendo por isso referidas como variação alofónica intrínseca. Esta contrapõe-se à variação alofónica extrínseca, detectada por esforço de atenção: [l, ɫ, b, β], e tantas outras.

---

\* Trata-se de um exemplo hipotético, análogo a outros descritos por Borden e Harris (1982).

A incapacidade em encontrar índices acústicos invariantes para determinadas categorias fonéticas coloca problemas práticos e teóricos. Primeiro, devido a esta falta de invariância, as tarefas de simular a fala ou de a reverter em outro tipo de sinal acessível a surdos não são triviais. Segundo, do ponto de vista teórico, a questão é como se recupera informação fonética invariante (perceber o /d/ em /di/ e /da/) a partir da informação acústica variável. Veremos mais em pormenor esta questão no capítulo seguinte. Agora concentremo-nos na própria existência ou não de invariância acústica.

A descoberta da falta de invariância acústica nos Laboratórios Haskins foi seguida por inúmeros trabalhos experimentais que, em geral, a replicaram. Estes trabalhos recorrem à estratégia de manipular propriedades sonoras isoláveis com base nas representações espectrográficas. Porém, o que para o observador experiente é visível nos espectrogramas, ou noutra forma de representação dos sons da fala, pode não ser necessariamente o que é relevante para o sistema perceptivo. De facto, outros investigadores têm-se interrogado sobre a inexistência de invariantes acústicos (*e.g.*, Cole e Scott, 1974), e foram encetados trabalhos com o objectivo de encontrar propriedades acústicas globais que permitissem distinguir os vários fones.

Stevens e Blumstein (1978) tentaram identificar características independentes do contexto em oclusivas de diferentes lugares de articulação. Propuseram que os padrões espectrais logo a seguir à libertação da oclusão são de facto invariantes. As consoantes labiais apresentam uma série de picos, cujas amplitudes ou se mantêm semelhantes ou são mais fortes nas baixas frequências, o que os autores designam por padrão difuso, plano ou descendente, respectivamente. As consoantes alveolares caracterizam-se por um padrão também difuso, mas ascendente. As consoantes velares têm

um pico espectral saliente, em geral nas frequências intermédias; o seu padrão é compacto.

Numa série de experiências, Blumstein e Stevens (1979; 1980) tentaram demonstrar a validade desta descrição, quer em termos formais quer perceptivos. Mostraram primeiro que as características espectrais de uma série de sílabas Consoante-Vogal (CV) e Vogal-Consoante (VC) produzidas por locutores diferentes obedeciam, na zona de libertação de oclusão, aos padrões referidos (Blumstein e Stevens, 1979). Perante a apresentação de apenas 10 a 20 ms de sílabas sintéticas (da zona de libertação de oclusão), os sujeitos conseguiam identificar o lugar de articulação da consoante inicial (Blumstein e Stevens, 1980).

Na sequência deste e doutros estudos, tem-se vindo a desenvolver uma teoria da invariância acústica que faz duas asserções fundamentais: primeiro, existem padrões acústicos na fala, correspondentes a traços fonéticos, que são invariantes a despeito de variação do contexto, locutor e até língua; segundo, estes padrões invariantes são usados pelo sujeito humano ao processar os sons da fala, ou seja, são relevantes perceptivamente. Por exemplo, Stevens (1981; cf. também Blumstein e Stevens, 1981) sugere que segmentos fonéticos da mesma classe partilham certas propriedades acústicas a que o sistema auditivo reagiria de modo específico, permitindo assim a distinção entre essa e outras classes fonéticas. As propriedades invariantes a que o sistema auditivo seria sensível já muito precocemente constituiriam os índices primários usados quer pelas crianças, quer pelos adultos, ao identificar fones e fonemas. As propriedades acústicas dependentes do contexto, como as trajectórias ou a frequência inicial dos formantes, constituiriam índices secundários que acabam por ser usados na percepção em virtude de um processo de aprendizagem. Ou seja, na medida em que os índices primários ocorrem juntamente com os índices

secundários, ao longo do desenvolvimento ir-se-ia adquirindo a capacidade de fazer discriminações fonéticas com base em informações acústica variável e dependente do contexto para a qual não existe à partida uma sensibilidade auditiva específica.

Colocam-se porém sérias reservas a uma teoria naqueles termos, pelo menos no actual estado do conhecimento. A propósito da relevância perceptiva dos padrões acústicos invariantes, a evidência experimental é contraditória. Blumstein, Isaacs e Mertus (1982) fizeram uma comparação do efeito perceptivo dos índices primários, os invariantes acústicos espectrais, e dos índices secundários, as transições de formantes, quanto ao lugar de articulação de oclusivas, e verificaram que as respostas dos sujeitos, adultos, eram determinadas essencialmente pelas frequências iniciais dos formantes; a forma global do espectro exercia uma influência menor. Uma tentativa de averiguar directamente o papel destes índices primários e secundários em termos de desenvolvimento revelou também que, quer crianças de 5 anos, quer adultos, respondiam mais em função das transições de formante do que das características espectrais (Walley e Carrel, 1983). Estes resultados indicam que as propriedades invariantes caracterizando o lugar de articulação das oclusivas, pelo menos as usadas nestes estudos, e que consistem nos padrões espectrais dos 10 a 30 ms a seguir à libertação da oclusão, são menos relevantes perceptivamente do que os índices "clássicos" e dependentes do contexto, as transições de formantes.

A propósito da existência de padrões acústicos invariantes, o que se tem verificado é que estes têm sido caracterizados de modos diferentes conforme os investigadores (*e.g.*, Lahiri, Gewirth e Blumstein, 1984; Searle, Jacobson e Rayment, 1979). Uma questão que parece ser importante é se as propriedades invariantes se deveriam definir em termos espectrais e

estáticos, como propuseram Stevens e Blumstein (1978), ou em termos dinâmicos considerando a variação espectral ao longo do tempo.

Com base em técnicas que permitem representar a evolução temporal do espectro em três dimensões, Kewley-Port (1983) identificou propriedades invariantes espectro-temporais associadas ao lugar de articulação das oclusivas sonoras. Estas propriedades encontram-se na porção inicial (20 a 40 ms) de sílabas CV naturais; estímulos sintéticos preservando estas propriedades são correctamente identificados por auditores quanto ao seu lugar de articulação (Kewley-Port, Pisoni e Studdert-Kennedy, 1983). Comparando sílabas sintetizadas segundo o modelo de Kewley-Port (propriedades invariantes dinâmicas espectro-temporais) ou segundo o modelo de Stevens e Blumstein (propriedades invariantes estáticas, espectrais), verificou-se que os sujeitos identificavam melhor o lugar de articulação das primeiras (*ibidem*). Kewley-Port *et al.* (*ibidem*, p. 1793) sugerem que "the unitary, dynamic gestures of stop articulation are reflected in the coherent, time-varying properties of running spectra". Esta ideia de que um gesto articulatório unitário e dinâmico tem como consequência um padrão acústico coerente e também dinâmico é sem dúvida atraente. De certo modo, ela vem remeter para segundo plano a questão da invariância, e salientar a necessidade de se considerar a estrutura dinâmica do padrão acústico (cf. Schöner e Castro, submetido).

Em suma, a busca de invariantes acústicos para categorias fonéticas sofreu recentemente um ressurgimento graças ao recurso a novas técnicas de representação e análise do sinal acústico. É ainda muito cedo para extrair conclusões definitivas; as pesquisas feitas parecem no entanto apontar para os limites de um tal empreendimento se se considerarem apenas propriedades espectrais estáticas. Importante para resolver a questão da invariância será certamente o esclarecimento das sensibilidades

auditivas para tipos específicos de informação acústica. Um aspecto a considerar é que poderá eventualmente haver padrões acústicos invariantes associados a algumas categorias fonéticas e não a outras (cf. Lisker, 1985). Enfim, para já, o que sem dúvida se encontra solidamente estabelecido é que é possível perceber e distinguir fones com base em informação acústica dependente do contexto. Caberá à investigação futura mostrar até que ponto a percepção da fala poderá ser mediada por outros tipos de informação.

## CAPÍTULO II

## O SISTEMA DE PERCEPÇÃO DE FALA

No presente capítulo tentar-se-á esclarecer o funcionamento do sistema de percepção da fala em alguns dos seus aspectos. Começaremos por explicitar o que, com outros investigadores, entendemos por percepção da fala, e identificar as principais questões teoricamente relevantes. Examinaremos de seguida com algum pormenor o fenómeno conhecido por percepção categórica. Fá-lo-emos por duas razões: primeiro, o exame à investigação deste fenómeno fornece respostas e "insights" importantes para o problema mais geral da percepção de fala; segundo, uma das experiências que realizámos (Experiência VI) aparenta-se com este paradigma de investigação. Finalmente, retomaremos questões experimentais e teóricas mais gerais, privilegiando a abordagem característica da ciência da fala.

### 1. Extensão e limites do conceito e algumas questões sobre o processo

O fenómeno talvez mais saliente ao ouvir fala em situações quotidianas é que nós percebemos significados e não propriamente sons; como diz Darwin (1976), a nossa experiência consciente segue a intenção semântica do interlocutor. Na medida em que a fala é veículo de significações, fenomenologicamente perceber a fala é perceber significações. Há porém boas razões para crer que este acesso aparentemente directo ao significado é precedido por um conjunto complexo de transformações a partir da estimulação proximal constituída pelo padrão acústico; a título

meramente exemplificativo, refira-se que a percepção da fala não é instantânea, e os erros ocasionalmente observados só se compreendem postulando uma análise da estrutura interna dos estímulos. O ponto de vista correntemente aceite é que a percepção da fala envolve vários estágios de análise da estimulação (e.g., Elman e McClelland, 1988; Liberman e Studdert-Kennedy, 1978; Pisoni e Luce, 1987; Norris e Cutter, 1988; Sawusch e Nusbaum, 1983), uma posição que em certa medida reflecte o actual *Zeitgeist* marcado pelo paradigma cognitivo. Foi Studdert-Kennedy (1974) quem primeiro explicitou os possíveis estágios do processo perceptivo. Eles seriam pelo menos quatro: o primeiro, auditivo, o único directamente baseado no *input* físico; o segundo, fonético, em que o *output* das operações do primeiro estágio seria organizado em termos de traços fonéticos; o terceiro, fonológico, em que os segmentos fonéticos são convertidos numa representação fonológica que despreza a variação alofónica; e finalmente um quarto nível, que recebe os epítetos de lexical, sintáctico e semântico, que consistiria num conjunto complexo de processos interrelacionados. Este modelo genérico dos processos de tratamento da fala ainda hoje constitui um quadro de referência importante (cf., e.g., Frauenfelder e Tyler, 1987; Pisoni e Luce, 1987).

Um modelo como este levanta pelo menos dois problemas fundamentais. O primeiro é o seguinte: supondo que de facto existem estes quatro estágios de processamento, como estão eles organizados? Serão sucessivos e hierárquicos? Apesar de conceptualmente se poder considerar que há uma espécie de progressão de níveis inferiores para superiores, não será necessariamente isso que acontece em termos funcionais. Segundo Studdert-Kennedy, por exemplo (1974, p. 2350), estes vários estágios devem ser não só sucessivos, mas também simultâneos: o resultado das operações de um nível poderiam retroagir para um outro nível, não só para "resolver" eventuais contradições como também para ajudar à determinação, por

exemplo, das características fonéticas com base em regras fonológicas ou sintáticas. O problema da existência ou não de interações entre diferentes estágios de tratamento merece hoje considerável discussão, neste como noutros domínios; retomá-lo-emos mais tarde.

O outro problema, mais fundamental, é o da própria realidade psicológica do modelo: até que ponto é ele válido para dar conta dos processos envolvidos no processamento da fala? Ocorrerão aqueles vários estágios em todas as situações ou apenas nalgumas? De facto, quer na apresentação original de Studdert-Kennedy (*ibidem*), quer em reflexões recentes se nota por vezes uma ambiguidade quanto ao modo de encarar os vários estágios: se enquanto meios de clarificação conceptual sobre os processos perceptivos, se enquanto (estágios de) processos perceptivos propriamente ditos. Escreve Studdert-Kennedy (*ibidem*, p. 2350): "...it will be useful to lay out, for purposes of exposition, a rough model of the transformation from signal to message. The process entails, conceptually, at least these stages of analysis..." (sublinhado nosso). E Pisoni e Luce (1987), ao retomar o mesmo modelo, explicam que "although the exact nature of each and the interactions among the stages are still tentative, they are nevertheless theoretically justifiable on linguistic grounds" (p. 27, sublinhado nosso). Ou seja, a existência de um conjunto de diferentes tipos de operações e transformações do *input* acústico - os estágios auditivo, fonético, fonológico, lexical e outros - é plausível por razões de tipo linguístico e também psicológico; porém, estão ainda por especificar claramente quer a natureza dessas operações, quer a sua organização e o modo como se interrelacionam.

Enfim, aceitemos que perceber a fala envolveria uma série de processos e operações bem diferentes, das mais próximas do *input* físico até às mais relacionadas com o conhecimento e crenças sobre o mundo em geral e a linguagem em particular. Este problema da percepção da fala em sentido

lato é sem dúvida fascinante, mas, no estado actual do conhecimento, demasiado complexo para ser abordado cientificamente em toda a sua extensão e profundidade. Por isso, o termo percepção da fala é geralmente usado numa acepção mais restrita, que não engloba os processos mais gerais e talvez inespecíficos de compreensão e de processamento da linguagem propriamente ditos. Os trabalhos no domínio da percepção de fala têm-se centrado nos primeiros três estágios definidos por Studdert-Kennedy (*ibidem*), ou seja, no problema da extracção da mensagem fonética e fonológica do padrão acústico, relegando para fora do seu âmbito problemas do reconhecimento de palavras propriamente dito e da compreensão. De facto, por vezes a percepção da fala é tomada explicitamente como sinónimo da percepção dos sons da fala ou percepção fonética (*e.g.*, Borden e Harris 1982; Foss e Hakes, 1978). Isto obviamente não implica aceitar que perceber a fala se "reduz" a ouvir fones e fonemas; trata-se apenas de uma tática de longa tradição, a de dividir um grande problema em questões mais circunscritas, com o objectivo de as tornar cientificamente abordáveis.

Até há bem pouco tempo, distinguíam-se claramente na literatura dois campos: o da percepção da fala, muito marcado pelos trabalhos dos Laboratórios Haskins, e o do reconhecimento de palavras. A investigação psicológica sobre o reconhecimento de palavras, visando descrever e explicar como uma entidade significativa, a palavra, é reconhecida pelo sujeito, envolve por definição o nível lexical; o termo "processamento lexical" tem mesmo sido usado neste contexto (*cf.*, *e.g.*, Frauenfelder e Tyler, 1987). Os trabalhos neste domínio têm-se centrado no reconhecimento de palavras escritas, e só muito recentemente surgiram tentativas de lidar explicitamente com o reconhecimento de palavras faladas (*e.g.*, Cole e Jakimik, 1980; Elman e McClelland, 1984; Marslen-Wilson, 1987). Em parte resultado destas circunstâncias históricas, subsiste alguma ambiguidade

quanto à acepção das duas expressões genéricas, "percepção da fala" e "reconhecimento de palavras faladas". Veja-se, por exemplo, a posição de Frauenfelder e Tyler (1987, p. 6), que reservam o termo reconhecimento de palavras para o estágio final em que o auditor determina que entidade lexical foi de facto ouvida, e que é precedido pelo acesso lexical (o contacto com, ou a activação da representação mnésica da palavra, na sequência de uma estimulação visual ou auditiva); e a posição de Pisoni e Luce (1987), que usam indistintamente "percepção da fala" e "reconhecimento de palavras", para se referir aos processos responsáveis não só pela transformação da informação acustico-fonética da onda sonora num padrão, provavelmente de tipo fonológico, mas também pelo reconhecimento deste padrão, ou seja "matching this pattern to patterns previously stored in memory (*i.e.*, for words) or to patterns generated by rule (*i.e.*, for pseudowords)." (*ibidem*, p. 39).

A aproximação entre os dois campos tradicionalmente distintos, o da percepção da fala originário da ciência da fala, e o do reconhecimento de palavras originário da abordagem cognitiva da leitura, é talvez o desenvolvimento mais recente nestes domínios, e permitirá certamente avanços importantes na compreensão dos processos envolvidos no reconhecimento de palavras faladas e talvez também da própria percepção da fala. De facto, um dos problemas que recentemente tem vindo a ser debatido é o das possíveis influências lexicais na categorização fonética (cf., *e.g.*, Connine e Clifton, 1987; Elman e McClelland, 1988; Fox, 1984; Ganong, 1980).

Convém esclarecer, por questão de clareza terminológica e de delimitação deste trabalho, que o termo percepção da fala é por nós usado na acepção mais restrita, que não envolve necessariamente os processos de acesso lexical. Isto não significa que se tome a posição teórica segundo a

qual possíveis efeitos lexicais são pós-perceptivos. Outra limitação desta aceção de percepção da fala é que ela se centra na percepção de vocábulos isolados, e não no chamado "fluent speech", ou fala fluente. Enfim, centrar-nos-emos no presente capítulo nos problemas tradicionais da percepção da fala, nomeadamente a nível de percepção fonética.

Sem dúvida que o problema fundamental que uma teoria da percepção de fala tem de resolver é o da complexa relação entre sinal acústico e percepção, mesmo se esta consistir apenas numa sílaba ou numa palavra sem sentido (*e.g.*, Liberman, Cooper, Shankweiler e Studdert-Kennedy, 1967; Massaro e Cohen, 1983; Summerfield, 1982). A este nível, as duas questões mais debatidas na literatura são, por um lado, a relação entre os processos auditivos e fonéticos, e por outro a relação entre processos de percepção e de produção de fala. Estas duas questões estão relacionadas na medida em que, se se verificar que processos de tipo auditivo não chegam para explicar como se forma o percepto fala, uma opção atraente a considerar é que o sistema de produção de fala seja de algum modo útil na elaboração desse percepto.

Antes de discutirmos em geral aquelas questões, examinaremos de perto um fenómeno que as ilustra e que nos servirá de "analisador teórico"; ou seja, concentremo-nos na chamada percepção categórica, que durante algum tempo se pensou específica da fala, e vejamos através do desenrolar das investigações a ela devotadas o próprio desenvolvimento das concepções teóricas sobre a percepção de fala.

## 2. Percepção categórica da fala

### 2.1. Generalidades

Um dos fenómenos que mais atenção tem merecido por parte dos investigadores é conhecido por percepção categórica. Esta expressão refere-se, ao facto de a percepção operar de modo categórico, *i.e.*, se basear na utilização de categorias, não só para identificar os estímulos, mas também para fazer discriminações entre eles (*e.g.*, Liberman, Harris, Hoffman e Griffith, 1957; Studdert-Kennedy, Liberman, Harris e Cooper, 1970; Repp, 1984).

Como intuitivamente se poderia esperar, e a Psicofísica clássica analisou em detalhe, o ser humano é capaz de fazer discriminações finas entre atributos sensoriais como o brilho, a cor, o cheiro e tantos outros, sem necessariamente dispor de nomes específicos para cada concretização das diferenças percebidas: facilmente se nota que duas manchas de cor têm cambiantes distintos, partilhando embora a mesma categoria, por exemplo, azul (cf. Heider e Olivier, 1972). Além disto, esta capacidade discriminativa é "contínua", na medida em que uma dimensão do estímulo variando continuamente resulta numa série de perceptos variando também quantitativa e gradualmente. A determinação dos limiares diferenciais ou "jnds"\* em nada apela à identificação *per se* dos estímulos a comparar. Ora o que acontece na percepção da fala é que, pelo menos em certas condições, a capacidade de discriminar sons pertencentes a uma mesma categoria fonémica está severamente reduzida, sendo em compensação facilitada a discriminação inter-categórica.

A utilização de categorias reflecte provavelmente uma das facetas básicas do funcionamento cognitivo (cf. Harnad, 1987). Mesmo ao identificar

\* "Just noticeable difference", ou diferença mínima perceptível.



objectos e acontecimentos, estamos implícita ou explicitamente a categorizar: ignoramos variação irrelevante para nos centrarmos naquilo que "define" esse objecto ou acontecimento relativamente aos outros. Note-se porém que aqui a expressão "percepção categórica" não se refere apenas à acepção literal de "categórico" (uso de categorias ao interagir com o meio ambiente), mas também ao facto de o modo como se discriminam estímulos de um *continuum* depender do modo como são identificados.

## 2.2. A descoberta da percepção categórica

### 2.2.1. Requisitos técnicos

O fenómeno denominado percepção categórica foi descoberto nos Laboratórios Haskins por A. Liberman e colaboradores. O desenvolvimento neste laboratório do primeiro sintetizador de fala destinado à investigação, o famoso "Pattern Playback", permitiu abrir novas áreas na pesquisa (cf. Cooper *et al.*, 1984). Aquele aparelho consiste num sistema de reprodução acústica de padrões espectrográficos desenhados numa banda de acetato. Graças ao "Pattern Playback" foi possível construir séries de estímulos acústicos variando gradualmente de um extremo correspondente a um determinado fonema (*e.g.*, /ba/) ao extremo oposto, correspondente a uma outra categoria fonémica (*e.g.*, /ga/). Apesar de a variação de um extremo ao outro se fazer por pequenos passos, e nesse sentido não ser estritamente contínua, estas séries de estímulos são designadas pelo termo "*continuum*".

*Continua* daquele tipo podem ser realizados para vários contrastes fonémicos. Entre os primeiros, contam-se os três *continua* para lugar de articulação de oclusivas e nasais, elaborados por Liberman, Delattre, Cooper e Gerstman (1954), através da variação do ponto de partida da transição de  $F_2$  num contexto vocálico invariante (/b, d, g/, /p, t, k/ e /m, n, ŋ/). *Continua*

para contraste de vozeamento foram também dos primeiros efectuados, entre /d/ e /t/ em posição inicial por redução de  $F_1$  (Lieberman, Harris, Kinney e Lane, 1961), ou entre /b/ e /p/ em posição inter-vocálica por duração do silêncio (Lieberman, Harris, Eimas, Lisker e Bastian, 1961). Entretanto, em parte devido à cada vez maior sofisticação dos sintetizadores de fala e ao conhecimento mais detalhado de variáveis acústico-perceptivas, a construção dos *continua* alargou-se à distinção entre as aproximantes [r] e [l] (e.g., Miyawaki, Strange, Verbrugge, Liberman, Jenkins e Fujimura, 1975) e fricativas (entre [x] e [ʃ], por exemplo, cf. May, 1981).

### 2.2.2. Tarefas e funções de identificação e discriminação

Lieberman *et al.* (1957) utilizaram pela primeira vez um *continuum* sintético para examinar se perante diferenças acústicas semelhantes um auditor discrimina melhor os sons a que atribui diferentes nomes\*. Por neste trabalho não só estar descrita a descoberta básica, como ainda serem identificadas as questões que investigação posterior tentou resolver, vamos revê-lo de seguida.

O *continuum* usado abrangia as categorias /be, de, ge/, e era composto por dois formantes;  $F_1$  era constante e a frequência inicial de  $F_2$  variava, em passos iguais de 120 Hz. Havia ao todo 14 estímulos; a questão era averiguar se a diferença de 120 Hz seria detectada indiferentemente entre quaisquer destes 14 estímulos. Para resolver esta questão é necessário avaliar independentemente o modo como se percebe cada estímulo isolado, e o modo como são levadas a cabo comparações entre os vários estímulos. Trata-se de executar duas tarefas, uma de identificação (também conhecida

---

\* O título daquele artigo seminal é justamente "The discrimination of speech sounds within and across phoneme boundaries" (*ibidem*).

por teste de identificação ou nomeação\*\*) e outra de discriminação. Com base nos resultados da tarefa de identificação, é possível verificar a percentagem de respostas de determinada categoria para os vários estímulos, e estabelecer a função de identificação ("labeling function"). Um resultado típico é que a variação acústica contínua é dividida em categorias perceptivas claras, com transições relativamente abruptas entre elas.

A tarefa de discriminação pode ser levada a cabo de várias formas; todavia, seguindo o exemplo pioneiro de Liberman e colaboradores, a tarefa tradicionalmente usada nas investigações sobre a percepção categórica obedece ao paradigma ABX. Em cada ensaio, são apresentadas tríades de estímulos; A e B representam estímulos diferentes com certo grau de adjacência e X um terceiro estímulo idêntico a A ou a B (ABA, ABB). A tarefa consiste em responder se X é igual a A ou igual a B. A diferença entre A e B é especificada pela distância a que se situam no *continuum*. Por exemplo, Liberman *et al.* (*ibidem*) fizeram três testes de discriminação, consoante a separação entre A e B fosse de apenas um grau (estímulo E<sub>1</sub>, estímulo E<sub>2</sub>, X), dois graus (E<sub>1</sub>, E<sub>3</sub>, X) ou três (E<sub>1</sub>, E<sub>4</sub>, X).

Os dados da(s) tarefa(s) de discriminação são representados numa função de discriminação, que mostra a relação entre a posição no *continuum* e a percentagem de respostas correctas nas tríades ABX (ou, em geral, a discriminabilidade de pares de estímulos igualmente separados). Tipicamente, a função obtida não é homogénea, exibindo picos de alta discriminabilidade entre estímulos previamente identificados como diferentes, e vales de baixa discriminabilidade (incluindo o nível do acaso) entre estímulos percebidos sob o mesmo nome. A forma exacta desta função depende do intervalo entre os estímulos a discriminar: obviamente é mais

---

\*\* Optamos por traduzir "labeling" tal como é usado em "labeling test" e "labeling functions" por nomeação.

fácil distinguir maiores diferenças do que pequenas. Isto traduz-se em geral por funções relativamente mais homogêneas para discriminação de grau 1 e de grau 3, respectivamente mais próximas do nível do acaso ou da exactidão a 100% (cf. Liberman *et al.*, 1957; Pisoni, 1975).

### 2.2.3. A relação entre funções hipotéticas e observadas

Com o propósito de explorar teoricamente a relação entre identificação e discriminação, Liberman *et al.* (1957) calcularam funções de discriminação hipotéticas baseadas no pressuposto extremo de que o auditor só consegue ouvir diferenças entre os sons a que atribui diferentes nomes ou "etiquetas" fonémicas. Estas funções preditas são comparadas com as obtidas, sendo assim possível verificar até que ponto aquele modelo extremo de percepção categórica, conhecido como "modelo Haskins", dá conta dos resultados. O cálculo das funções de discriminação hipotéticas é feito com base nas respostas de identificação. Se dois estímulos tiverem uma probabilidade de 100% de lhes ser atribuído um nome diferente, o modelo prediz uma discriminação também de 100%; se outros dois estímulos forem inconsistentemente nomeados, cada um com 50% de respostas "A" e 50% respostas "B", então a discriminação entre eles estará, segundo o modelo, ao nível do acaso (cf. Liberman *et al.*, *ibidem*; Lane, 1965; Studdert-Kennedy *et al.*, 1970). Strange e Halwes (1977) desenvolveram um método ligeiramente diferente, apropriado para funções de identificação obtidas com menor número de resposta por estímulo, acompanhadas por julgamentos do respectivo grau de confiança.

As funções de discriminação hipotéticas prediziam razoavelmente as funções obtidas, nomeadamente nos picos, e para separações de dois e três graus. Nestes dois casos, a correlação entre função obtida e hipotética

rondava os  $+0.43^*$ . No caso da separação de um grau (diferença de 120 Hz), a correlação foi negligenciável e não significativa; é que o modelo previa respostas ao nível do acaso para quase todas as comparações, enquanto a *performance* de facto excedeu sistematicamente esse nível. Aliás, apesar do acordo genérico, mesmo para os outros pares de estímulos as funções obtidas excedem em geral as funções previstas: os sujeitos conseguem discriminar estímulos pertencentes à mesma categoria, embora com menos exactidão do que os estímulos de diferentes categorias.

Este problema é notado e discutido por Liberman *et al.* (*ibidem*); embora reconhecendo a impossibilidade de uma interpretação cabal, acham plausível que a discrepância entre funções obtidas e previstas possa ser devida pelo menos em parte a "true discrimination, *i.e.*, an ability of subjects to distinguish the speech sounds, not solely on the basis of phonemic labels, but also more directly by the essential acoustic differences among the patterns" (p. 366). Esta explicação prefigura interpretações mais modernas da percepção categórica, que concebem a acção de factores de tipo acústico e psicofísico ao lado de factores linguísticos (*e.g.*, Liberman e Mattingly, 1985; Pisoni, 1985; Repp, 1984).

Para um leitor desapaixonado dos anos 80, não deixa de ser curioso que grande parte da polémica em torno da natureza de percepção categórica em especial, e da percepção da fala em geral, tenha esquecido esta observação precoce. Parece que se tomou apenas o outro lado das descobertas de 1957: o do relativo acordo das funções obtidas com as previstas, segundo o pressuposto extremo de que os sujeitos apenas podiam discriminar sons da fala na medida em que os identificassem como diferentes.

---

\* Este valor é a medida de correlação não paramétrica *tan*; a sua probabilidade era inferior a .001.

## 2.3. Extensões da descoberta inicial

### 2.3.1. Replicações

Ao estudo pioneiro que vimos seguiu-se a eclosão de uma série de outros, visando explorar a generalidade do fenómeno. Contrastes de vozeamento foram dos primeiros a ser utilizados: foram obtidos resultados categóricos na percepção de /do/ vs /to/ em posição inicial (Liberman *et al.*, 1961-b), e de /b/ vs /p/ em posição medial em /rabid/ e /rapid/ (Liberman *et al.*, 1961-a). Também *continua* de maneira de articulação foram explorados: Bastian, Eimas e Liberman (1961) usaram um continuum entre /slit/ e /split/, em que se manipulou a duração do intervalo de silêncio entre sibilante e aproximante, para o qual obtiveram picos de discriminação nas fronteiras fonémicas.

Nestas experiências os sujeitos desempenhavam uma tarefa adicional de discriminação de um *continuum* de sons não-fala\* tão semelhante quanto possível ao *continuum* fala. Esta tarefa destina-se a permitir uma comparação directa entre a discriminabilidade de estímulos fala e não-fala, importante para esclarecer as origens do fenómeno: se puramente auditivo, ou "especial", fonético/linguístico.

### 2.3.2. Comparação da discriminabilidade de *continua* não-fala e fala com acontecimentos acústicos semelhantes

As primeiras tentativas de encontrar um *continuum* acústico não-fala equivalente ao *continuum* fala foram relativamente toscas: tratava-se

---

\* Usaremos "não-fala" como substantivo ou adjectivo oposto a fala, equivalente ao "nonspeech" vs "speech".

de versões invertidas dos mesmos padrões (Liberman *et al.*, 1961-a), ou intervalos de silêncio iguais aos dos estímulos-fala fora do contexto vocálico, entre dois surtos de ruído (Liberman *et al.*, 1961-b). Os resultados mostraram que a percepção destes acontecimentos acústicos não era categórica.

Mais tarde, Mattingly, Liberman, Syrdal e Halwes (1971) usaram um método que veio a tornar-se padrão em estudos subsequentes: apresentar isoladamente as transições do segundo formante, os chamados "chirps". Os seus estímulos fala consistiam em *continua* de lugar de articulação em oclusivas em posição inicial ou final, /bæ - gæ/ e /æb - æg/, sintetizados com apenas dois formantes: F<sub>1</sub> constante e F<sub>2</sub> com porção estável invariante e transições variáveis. Note-se que são as várias transições de F<sub>2</sub> que permitem a percepção de sílabas diferentes. O *continuum* não-fala consistia justamente nestas transições do segundo formante, mas isoladas quer da porção estável de F<sub>2</sub>, quer de F<sub>1</sub>; elas constituem um acontecimento acústico *per se*, sendo percebidas como ruídos. O objectivo da experiência era comparar a discriminação das transições de F<sub>2</sub> em sílabas e isoladamente. O resultado foi de acordo com o predito por um modelo de percepção categórica: os picos de discriminabilidade nas zonas de fronteira fonémica apareceram somente para as sílabas, e não para os "chirps". Estes eram discriminados de modo relativamente contínuo, e em geral melhor do que as sílabas (a discriminação de sílabas tinha picos e vales mais marcados).

Com base na observação de que as funções de resposta são diferentes consoante os estímulos veiculem, ou não, percepções da fala, Mattingly *et al.* (*ibidem*) arguem pela existência de dois modos de percepção: um característico da fala e outro da não-fala (veja-se o título do artigo "Discrimination in speech and nonspeech modes"). A percepção categórica seria típica do modo fala ("speech mode"), e a percepção contínua do modo

não-fala. Outro argumento em favor desta distinção entre modo fala e não-fala provinha dos estudos de lateralidade auditiva. Encontrava-se uma vantagem do ouvido direito, traduzindo uma superioridade do hemisfério esquerdo, com sons da fala percebidos categoricamente (Shankweiler e Studdert-Kennedy, 1967). Este resultado sugeria que para os perceber se recorria a um mecanismo especificamente linguístico, que se caracterizava por ser categórico.

### 2.3.3. Percepção não categórica para vogais?

Contrastando com as replicações de percepção categórica para vários tipos de contrastes fonémicos entre consoantes oclusivas (Liberman *et al.*, 1961-a,-b; Abramson e Lisker, 1970; Mattingly *et al.*, 1971), os primeiros estudos usando *continua* vocálicos obtiveram resultados menos categóricos. Fry, Abramson, Eimas e Liberman (1962), usando um *continuum* sintético /I-e-æ/, verificaram que a discriminação era semelhante intra e inter-categorias, sem exibir picos e vales. Num outro estudo empírico com vogais suecas e inglesas (Stevens, Liberman, Öhman e Studdert-Kennedy, 1969) observou-se uma percepção menos contínua do que a apresentada por Fry *et al.* (1962), mas ainda assim a discriminação parecia independente de identificação.

Apesar de posteriormente terem sido tecidas críticas ao modo de construção do *continuum* vocálico usado por Fry e colaboradores (Repp, 1984), um resumo justo das primeiras experiências sobre a relação entre identificação e discriminação de sons da fala é o da demonstração de percepção categórica para (consoantes) oclusivas e sugestão de percepção não-categórica para vogais.

#### 2.3.4. O problema dos critérios empíricos na determinação de percepção categórica; fenómeno absoluto ou gradual?

As investigações iniciais sobre percepção categórica, levadas a cabo na década de 60, saíram quase todas de Haskins: é que só a partir dos anos 70 outros laboratórios começaram a dispor de sintetizadores da fala. A única excepção a esta regra foi o trabalho de H. Lane e colaboradores da Universidade de Michigan.

Lane (1965) reviu as experiências de Haskins à luz de uma perspectiva psicofísica, perguntando-se se não poderiam ser observados resultados típicos de percepção categórica com estímulos não-fala submetidos a condições experimentais semelhantes. Infelizmente, não apresentou dados empíricos convincentes em favor daquela hipótese, pelo que o impacto que na altura exerceu foi medíocre (cf. Repp, 1984, p. 246).

Embora antecipando críticas actuais, nomeadamente a de que a percepção categórica não é específica nem exclusiva da fala, a consequência mais marcante do trabalho de Lane foi antes a afirmação extrema do ponto de vista de Haskins, no artigo resposta de Studdert-Kennedy *et al.* (1970). Nele são examinados os critérios empíricos que permitem inferir a existência de percepção categórica, a saber: (1) As respostas de identificação aos vários estímulos do *continuum* exibem transições abruptas, o que equivale a funções de identificação com zonas declive acentuado; o(s) ponto(s) de declive máximo define(m) a(s) fronteira(s) fonémica(s) (ou de categoria); este ponto é aquele em que são equiprováveis respostas de categorias adjacentes; (2) As funções de discriminação têm um pico na fronteira fonémica; (3) A discriminação dentro de cada categoria é próxima do nível do acaso; (4) Com base nas respostas de identificação podem prever-se as

funções de discriminação. Este critério tem como corolário que o pico de discriminação ocorre num local e magnitude determinados.

Estes critérios definem as quatro características semi-independentes de uma percepção categórica ideal. Na prática, raramente os dados se acomodam a uma configuração tão clara, sendo necessário atribuir mais ou menos peso a um ou outro critério. Assim, o primeiro é relativamente pouco fiável, pois obviamente o grau de declive depende do espaçamento dos estímulos ao longo do *continuum* (cf. Lane, 1965; Studdert-Kennedy *et al.*, 1970). O critério mais decisivo é o segundo, *i.e.*, que o pico de discriminação se encontra junto da fronteira fonémica. A esta característica foi dado o nome de efeito de fronteira fonémica por C. C. Wood em 1976 (*in* Repp, 1984, p. 253). A observação deste efeito é o principal índice de percepção categórica. Desvios moderados dos critérios 3 e 4 podem ser tolerados. A questão é qual o grau de desvio possível, e se a sua determinação pode ser impressionista ou requer algum critério estatístico. Se nas primeiras observações o recurso a critérios impressionistas era na prática justificado pela clara discrepância entre os resultados para as oclusivas por um lado, e para os sons não-fala e mesmo vogais isoladas por outro, o mesmo não se pode dizer de investigações posteriores, onde por manipulação de material e tarefa se obtiveram casos intermédios, mais ou menos contínuos, mais ou menos categóricos (cf. Repp, 1984, p. 288).

Estas observações colocam o problema de ser ou não lícito falar em grau(s) de percepção categórica, sem dúvida uma maneira evidente, talvez fácil e menos elegante, de dar conta dos factos observados. Na perspectiva clássica de Haskins (Studdert-Kennedy *et al.*, 1970), porém, a percepção categórica é um fenómeno absoluto, de tudo-ou-nada: "categorical perception refers to a mode by which stimuli are responded to, and can only be responded to, in absolute terms. Successive stimuli drawn from a physical

*continuum* are not perceived as forming a *continuum*, but as members of discrete categories" (p. 234). Os autores afirmam também que os vários estímulos "are identified absolutely, that is, independently of the context in which they occur" (*ibidem*, p. 234). Esta alusão à estabilidade na percepção da fala foi explicitada por Mattingly *et al.* (1971), que notam que factores como motivação, activação, e experiência com fala sintética pouco influenciam as respostas do sujeito desde que estes percebam os estímulos como fala.

## 2.4. Factores explicativos: o modo fala, a aprendizagem e a sensibilidade auditiva

### 2.4.1. Teoria motora e percepção categórica

A concepção da percepção categórica como fenómeno absoluto e específico de sons da fala está intimamente ligada aos desenvolvimentos teóricos dos anos 50 e 60 em ciência da fala, nomeadamente a teoria motora de Liberman, Mattingly e colaboradores (Liberman, Cooper, Harris e MacNeilage, 1962; Liberman, Cooper, Shankweiler e Studdert-Kennedy, 1967; Mattingly e Liberman, 1969). Para categorias fonéticas produzidas por gestos articulatorios discretos, como as oclusivas, a percepção é também discreta ou categórica. Quando as diferentes categorias fonéticas podem ser produzidas através da variação contínua da posição dos articuladores (como nas vogais), a percepção é contínua. A ideia fundamental é que a percepção de fala (categórica nuns casos, contínua noutros) é mediada por uma representação articulatória da estimulação. À luz desta teoria, a percepção categórica é uma consequência directa da organização articulatória (Liberman *et al.*, 1962). Como os dados disponíveis na altura sugeriam um claro paralelismo entre modo contínuo *vs.* categórico de percepção e produção, nas primeiras formulações da teoria a percepção categórica

figurava como um dos argumentos de apoio mais importantes (cf. Liberman *et al.*, 1967).

Em suma, a interpretação corrente para os primeiros resultados obtidos através deste paradigma era a seguinte: a percepção categórica é específica para a fala, mais precisamente para sons produzidos de modo descontínuo. Ela dever-se-á pois à existência de um modo especial para a percepção de fala caracterizado pela mediação articulatória. Desde então foram completados inúmeros estudos, que foram recentemente examinados em pormenor por Repp (1984; ver também Repp e Liberman, 1987) e Rosen e Howell (1987). Limitar-nos-emos aqui a discutir as principais vias de investigação e as linhas interpretativas que sugerem.

#### 2.4.2. O papel da aprendizagem e experiência linguística

Um dos primeiros pontos a ser examinado é o seguinte: até que ponto é a percepção categórica específica da fala? Os resultados obtidos com *continua* de padrões acústicos diferentes da fala foram primeiro um tanto equívocos: Cutting e Rosner (1974) apresentaram resultados categóricos para a identificação e discriminação de um *continuum* pseudo-musical, variando entre um extremo semelhante a um som dedilhado e outro extremo semelhante a um tom tocado com arco\* ; porém, estes resultados eram afinal devidos a uma descontinuidade física no *continuum* que não tinha sido controlada (Rosen e Howell, 1981). Outras experiências com estímulos musicais mostraram que pode existir percepção categórica com material deste tipo (Cutting, 1982). Ela parece porém depender da experiência: observam-se resultados categóricos quanto à discriminação de intervalos musicais em indivíduos com treino musical, mas não em sujeitos não-

---

\*"Plucks" e "bows", como lhes chamaram Cutting e Rosner (*ibidem*).

músicos (Burns e Ward, 1978; para outros exemplos, cf. Repp, 1984). Estes resultados sugerem que um factor explicativo da percepção categórica é a aprendizagem.

Se a aprendizagem é factor determinante, então deveria ser possível, também através da aprendizagem, induzir percepção não categórica para sons da fala em que tipicamente se observam resultados categóricos. De facto, é possível mediante o treino e o *feedback* melhorar substancialmente a discriminabilidade intra-categórica (e.g., Hanson, 1977; Pisoni, Aslin, Perey e Hennessy, 1982; Samuel, 1977). Entre os factores do procedimento experimental que parecem ser influentes, encontra-se o facto de se restringir a tarefa de discriminação, ou parte dela, a estímulos intra-categóricos. Quando se misturam pares inter-categóricos, além de possivelmente se induzir os sujeitos a usarem um critério fonético, introduz-se mais incerteza na situação, o que por sua vez aumenta os limiares psicofísicos (cf. Repp, 1984). Outra manipulação experimental sugere que a incerteza na situação determina em parte o perfil de discriminações observado: usando um tom de referência fixo, favorece-se a discriminação intra-categórica (cf. Pastore *et al.*, 1977; Repp, 1984).

Outro aspecto da influência da aprendizagem pode encontrar-se na própria experiência linguística. Aprendendo a dar nomes diferentes aos sons da fala através da experiência numa determinada comunidade linguística, aprender-se-ia também a discriminar melhor entre fonemas do que entre alofones.

Os resultados de várias comparações trans-linguísticas mostram que a localização das fronteiras fonémicas e dos picos de discriminabilidade é efectivamente afectada pela experiência linguística (cf. Repp, 1984; Rosen e Howell, 1987). Por exemplo, Williams (1977) usou um *continuum* de vozeamento que apresentou a espanhóis e ingleses. Para os primeiros, a

fronteira fonémica situava-se perto de 0 milissegundos de VOT, onde se observava também um pico de discriminabilidade; para os ingleses os resultados foram diferentes, com a fronteira e o pico perto dos 25 milissegundos de VOT. Resultados análogos foram observados por Miyawaki *et al.* (1975) com um *continuum* entre [ra] e [la] apresentado a sujeitos japoneses e americanos. Nenhum daqueles dois sons é usado em japonês, mas em inglês, como se sabe, são dois fonemas distintos (cf., por exemplo, "lent" e "rent"). Contrariamente aos americanos, que apresentaram resultados categóricos, os japoneses não. Particularmente esclarecedor é o facto de as respostas perante os terceiros formantes apresentados isoladamente serem idênticas nos dois grupos (são as características do terceiro formante que distinguem o [r] do [l] ingleses). Este último resultado indica claramente que o efeito da experiência linguística é específico: só ocorre quando se percebe estímulos fala, e não os índices acústicos isolados que lhes estão subjacentes.

Outras comparações inter-linguísticas mostram o efeito da experiência na percepção categórica (cf. Repp, 1984). Porém nalguns casos as funções de discriminação apresentam picos que não dependem da identificação; por exemplo, os sujeitos espanhóis observados por Williams (1977) exibiam um segundo pico de discriminação próximo do dos sujeitos americanos. Isto indica que era possível distinguir alofones (oclusivas com ou sem aspiração), certamente com base em sensibilidades auditivas. Temos pois um novo factor potencialmente relevante na determinação da percepção categórica: a sensibilidade auditiva.

### 2.4.3. Percepção categórica em bebês e animais não-humanos

Segundo alguns autores, a percepção categórica deve-se à existência de limiares psico-acústicos mais baixos coincidentes com as zonas de fronteira fonémica, não sendo necessário recorrer a factores explicativos como a categorização fonética (cf. Crowder, 1981; Tartter, 1982). Este tipo de interpretação foi denominada "common factor model" (Pastore *et al.*, 1977; p. 686). Dois tipos de resultados são particularmente relevantes para a avaliação deste modelo, os recolhidos em bebês e em espécies não-humanas.

A partir do estudo seminal de Eimas, Siqueland, Jusczyk e Vigorito (1971), têm-se multiplicado os trabalhos sobre capacidades discriminativas dos bebês relevantes para a linguagem. Usando a técnica de habituação, Eimas *et al.* (*ibidem*) observaram as frequências de sucção para sons de um *continuum* de vozeamento entre /ba/ e /pa/. Verificou-se que os bebês "discriminavam" sons a que os adultos dão nomes diferentes, mas não sons do mesmo *continuum* com idêntica diferença acústica (20 milisegundos) e pertencentes à mesma categoria perceptiva adulta (por exemplo, /ba/s com 20 e 0 milisegundos de VOT). Estes resultados foram replicados e observados com outros contrastes fonémicos (/ba/-/ga/, /ra/-/la/, /ba/-/wa/); deles se pode concluir que as capacidades perceptivas dos bebês de poucas semanas ou meses são muito semelhantes às dos adultos (cf. *e.g.*, Hillenbrand, 1984; Repp, 1984; Rosen e Howell, 1987). Convém porém referir que os bebês não fazem todas as distinções fonéticas usadas na língua adulta (Repp, 1984); isto atesta mais uma vez o papel da experiência da determinação do comportamento adulto.

As capacidades discriminativas dos bebês não se podem atribuir à experiência linguística, pois trata-se de sujeitos muito jovens que apenas

têm uma exposição passiva à linguagem. Corroborando esta suposição, alguns estudos trans-linguísticos revelaram que bebês de diferentes comunidades linguísticas discriminam *continua* da mesma maneira. (e.g., Aslin *et al.*, 1981; Eimas, 1975; Streeter, 1975; Lasky, Syrdal-Lasky e Klein, 1975). Aliás, parece que os bebês conseguem mesmo fazer discriminações que não são usadas na língua da sua comunidade; por exemplo, bebês americanos discriminam consoantes pré-vozeadas das não-vozeadas (Aslin *et al.*, 1981; cf. também Werker e Tees, 1984).

Os bebês podem exibir esta organização perceptiva por duas razões: ou porque ela corresponde a características da própria sensibilidade auditiva, sobre a qual se baseariam desenvolvimentos linguísticos posteriores; ou porque existiria muito precocemente um mecanismo específico para a linguagem. Se em espécies não-humanas se verificasse um padrão de resultados semelhante ao obtido com bebês, teríamos indicação de que a primeira alternativa poderia ser adequada.

A utilização de sujeitos não-humanos requer uma série de adaptações metodológicas que dificultam a exequibilidade destes estudos. Em consequência disto, o seu número é relativamente reduzido. As espécies mais estudadas são os macacos e chinchilas, cujo sistema auditivo é semelhante ao nosso. Foram feitas algumas demonstrações convincentes de capacidades discriminativas em não-humanos semelhantes às dos humanos, uma usando *continua* de vozeamento apresentados a chinchilas (Kuhl e Miller, 1978), e outras usando *continua* de lugar de articulação e de vozeamento apresentados a macacos (Kuhl e Padden, 1982; 1983). Apesar de haver outras experiências em que não se observaram estes efeitos, os resultados obtidos por Kuhl sugerem que há fronteiras psico-acústicas em espécies com um sistema auditivo semelhante ao humano, quanto ao

vozeamento na região da VOTs negativas, e quanto ao lugar (Repp, 1984; Kuhl, 1987).

A aplicação do paradigma da percepção categórica em sujeitos não-humanos veio pois apoiar indirectamente a ideia segundo a qual poderá haver sensibilidades auditivas específicas que estejam na base dos picos de discriminação inter-categórica observados em humanos. Convém porém acentuar que a demonstração de fronteiras psico-acústicas em animais não esclarece qual o papel de eventuais fronteiras análogas na percepção de fala por humanos (Repp, 1984); veja-se por exemplo os efeitos da experiência linguística.

## 2.5. Conclusão

Identificámos assim dois factores explicativos da percepção categórica, a aprendizagem e as sensibilidades auditivas, para além do inicialmente proposto pelos investigadores de Haskins, um mecanismo específico para a fala de natureza articulatória.

Os argumentos em favor do papel da experiência linguística e da aprendizagem são claros, e os resultados de percepção categórica em bebés e animais sugerem que sensibilidades auditivas poderiam ditar pelo menos algumas das descontinuidades observadas nos adultos. Será então que estes dois factores são suficientes para dar conta dos resultados? No estado actual do conhecimento, uma tal conclusão parece ser precipitada. Existe com efeito um resultado experimental que não é facilmente interpretável à luz daqueles factores: é o facto das fronteiras fonémicas variarem em função de factores como o contexto e o ritmo de fala, entre outros, até para o mesmo indivíduo. A flexibilidade das fronteiras fonémicas, para além da mera existência dos picos de discriminação, indica que hipotéticas sensibilidades

auditivas não explicam tudo (pois não seria de esperar essas hipotéticas sensibilidades fossem assim variáveis). Uma revisão recente sobre as condições em que se observa esta flexibilidade de fronteiras fonémicas leva Repp e Liberman (1987) a concluírem que o modo - variável - como se percebem as categorias fonéticas reflecte o conhecimento implícito das acções do trato vocal. Teremos ocasião mais à frente de debatermos este tipo de sugestão; por agora remetemos o leitor interessado para a revisão de Repp e Liberman (1987).

Em suma, as inúmeras observações disponíveis mostram que a chamada percepção categórica não é tão específica nem rígida (tipo tudo-ou-nada) como supunha o modelo inicial de Haskins; talvez a percepção categórica de sons da fala não se deva, pelo menos exclusivamente, à acção de um mecanismo especial para perceber fala; ela poderia assentar em sensibilidades auditivas até certo ponto inatas, alteráveis dentro de certos limites pela aprendizagem. O papel da mediação articulatória é incerto, mas parece prematuro bani-la completamente.

Em termos de processamento numa situação de percepção de fala, parece adequar-se melhor aos factos um modelo de processamento dual, que aceita a existência de processos auditivos ao lado de processos fonéticos (cf. Repp, 1984). Um modelo deste tipo foi proposto por Fujisaki e Kawashima já em 1969, e elaborado por Pisoni (*e.g.*, 1975). Contrariamente ao modelo de Haskins, que propõe um processo único, fonético, e ao modelo do factor comum, auditivo, este supõe que há dois modos perceptivos que determinam a *performance*: um auditivo, em que o sinal acústico é tratado enquanto som, por processos comuns aos da percepção auditiva; o outro, de categorização fonética. Exigências de tarefa, características da situação e estratégias do sujeito determinam se o desempenho depende mais de um ou outro processo. O facto de o sujeito poder mudar a sua *performance*,

tornando-a mais "auditiva" ou mais "fonética" numa mesma situação (e.g., Best, Morrongiello e Robson, 1981) é uma poderosa indicação que há de facto dois processos que podem mediar a resposta. Este modelo é plausível e útil; ele não invalida porém a necessidade de esclarecer a natureza dos dois processos, nomeadamente se há descontinuidades de sensibilidade psico-acústica, e até que ponto a categorização fonética usa a mediação articulatória, e/ou depende da aprendizagem.

### 3. O modo fonético

O exame da investigação consagrada à percepção categórica mostra-nos como um fenómeno que inicialmente se pensava traduzir o funcionamento de um mecanismo específico para perceber a fala pode afinal, em certas condições, ser interpretável à luz de outros factores, como as sensibilidades auditivas e a aprendizagem propriamente dita. Também o ressurgimento da procura de invariantes acústicos, veículos de informação fonética, aponta para a importância de sensibilidades auditivas gerais. Do ponto de vista teórico, a questão que se coloca é a da especificidade dos processos ou mecanismos que estão subjacentes à percepção de fala: até que ponto será ela explicada através de processos gerais comuns à modalidade auditiva, ou pelo contrário exigirá modos específicos de tratamento? Os argumentos em favor da existência de um modo especial, fonético, para perceber a fala foram recentemente revistos por Liberman e Mattingly (1985). Vamos examiná-los de seguida.

#### 3.1. Os efeitos de contexto

O percepto fonético não tem uma definição acústica clara. Isto acontece em resultado quer da articulação, quer da coarticulação, e é



demonstrado pelos factos seguintes. Primeiro, um mesmo acontecimento acústico em contextos diferentes pode dar azo a perceptos diferentes. A ilustração mais flagrante deste fenómeno encontra-se no famoso "sound of silence", ou seja, na percepção de oclusivas. Outros exemplos encontram-se no Capítulo I. Segundo, o mesmo percepto pode ser extraído de sons diferentes, em contextos diferentes. A variabilidade articulatória/acústica devida ao contexto é omnipresente: há efeitos de contexto mais ou menos marcados na concretização acústica de todas as categorias fonéticas até hoje estudadas.

Além dos dois factos anteriores, que constituem descobertas clássicas da ciência da fala, estudos recentes mostram que a segmentação perceptiva de uma determinada mensagem acústica não é equivalente a possíveis segmentações acústicas, e indicam que a extracção do percepto fonético tem em conta não só a parte do sinal onde se concentra a informação para um determinado fonema, mas também as regiões circundantes (Mann e Repp, 1980, 1981; Whalen, 1984; cf. também Jusczyk, Smith e Murphy, 1981).

### 3.2. Integração perceptiva de índices acústicos

Um poderoso argumento em favor da existência de processos específicos para tratar a fala encontra-se na multiplicidade e equivalência de índices. A existência de vários índices acústicos que podem ser manipulados independentemente (por exemplo, duração do silêncio e frequência inicial dos formantes) levou a uma descoberta interessante. Dentro de certos limites, a modificação do valor de um índice pode ser compensada pela modificação em sentido oposto do valor de um outro índice. Ou seja, há entre os vários índices acústicos relações de compensação ("trading relations", cf.

Repp, 1982), que traduzem uma equivalência perceptiva. Um exemplo de relações de compensação observa-se quanto à duração do silêncio e as frequências iniciais das transições, na distinção entre "split" e "slit": se estas frequências forem apropriadas a [p], a duração do silêncio necessário para que se ouça "split" é mais pequena (Fitch, Halwes, Erickson e Liberman, 1980).

Uma revisão aos vários fenómenos deste tipo foi feita por Repp (1982). O exame dos resultados experimentais sugere que pelo menos em algumas situações eventuais mecanismos de integração e interacção auditiva não explicam os resultados : só se observa equivalência perceptiva se os sujeitos ouvem os estímulos como fala (*e.g.*, Best, *et al.*, 1981; William, Verbrugge e Studdert-Kennedy, 1983). Vejamos um exemplo: Best *et al.* (1981) observaram uma relação de compensação entre a frequência inicial do primeiro formante ( $F_1$ ) e a duração do silêncio na percepção de "say" vs "stay". Depois substituíram os formantes por ondas sinusoidais com a mesma trajectória\*. Estes estímulos podem ser percebidos como fala, ou não (cf. Remez, Rubin, Pisoni e Carrel, 1981). Se os sujeitos ouvissem o padrão sinusoidal como fala, observavam-se relações de compensação: para ouvir [stei] era preciso menos silêncio se a frequência inicial da transição de  $F_1$  fosse apropriada à oclusiva. Pelo contrário, os sujeitos que não percebiam os estímulos como fala não integravam os dois índices, e baseavam as suas respostas ou nas propriedades espectrais, ou na duração do silêncio. Esta dissociação indica seguramente que a equivalência perceptiva observada não se deve apenas a interacções auditivas.

A multiplicidade e equivalência de índices acústicos tornam vão o empreendimento de tentar definir cada fone através dos respectivos índices: cada um é mais ou menos suficiente, mas nenhum é específico nem

---

\* Discutiremos já de seguida este procedimento.

realmente necessário, pois pode ser compensado por outros. É pois a própria noção de índice acústico, *stricto sensu*, que é posta em causa: não há uma característica acústica específica que seja necessária e suficiente para definir determinada categoria perceptiva.

Além de ser difícil definir características acústicas específicas de cada fone, parece também não ser fácil encontrar características que distingam acusticamente fala de não-fala. Dados empíricos em favor desta asserção provêm de investigações experimentais com análogos sinusoidais da fala. Trata-se de estímulos acústicos que destroem o típico padrão de formantes, substituindo-o por ondas sinusoidais que seguem as trajectórias dos formantes originais (Remez *et al.*, 1981). Apesar desta transformação, há sujeitos que percebem estes análogos sinusoidais como fala. Assim, apesar de fala natural e mesmo fala sintética corrente se caracterizarem por formantes em determinada banda de frequências, zonas de periodicidade em cada núcleo silábico, etc, estas propriedades não são estritamente necessárias para se perceber fala.

Por outro lado, quando se distorce um espectrograma não no sentido da simplificação, mas complicando-o, desviando os formantes normais em trajectórias contínuas, mas anormais, o percepto cinde-se em duas categorias distintas: fala, e um fundo de ruídos vários (Liberman e Studdert-Kennedy, 1978, p. 155).

### 3.3. Dissociação de perceptos fala e não-fala para o mesmo acontecimento acústico

Uma sílaba CV sintética é constituída por transições de formantes e formantes estáveis. Se isolarmos uma transição de formante, e a apresentarmos a um ouvido, o sujeito percebe um *glissando* ascendente ou

descendente conforme a trajectória da transição. Se apresentarmos a um ouvido o restante da sílaba, designada por base, e ao outro ouvido a transição isolada, devidamente alinhadas temporalmente, é possível ter dois perceptos, o *glissando* e a sílaba correspondente ao estímulo completo, como se este tivesse sido apresentado na íntegra aos dois ouvidos. Este efeito foi denominado "duplex perception" ou percepção dupla (Lieberman, Isenberg e Rakerd, 1981), e tem sido recentemente utilizado para demonstrar que um mesmo sinal acústico pode ter efeitos perceptivos qualitativamente diferentes: a transição de formante é percebida como *glissando* (pelo modo auditivo?) e é integrada com a base dando origem a um percepto de tipo fala.

Demonstrações mais fortes desta dissociação perceptiva têm sido levadas a cabo através da manipulação de certas características acústicas da transição. Por exemplo, as funções de discriminação são diferentes conforme os sujeitos respondam em função da sílaba ou do *glissando*. No primeiro caso, observam-se picos de discriminabilidade na fronteira fonémica; no segundo, a função de discriminação é homogénea (Mann e Liberman, 1983; cf. também Bentin e Mann, 1983; Repp, 1984; Repp e Bentin, 1984; Repp, Milburn e Ashkenas, 1983).

Em suma, a percepção dupla mostra que o mesmo acontecimento acústico pode ser tratado de duas maneiras, uma de tipo auditivo e outra de tipo fonético.

### 3.4. Integração audiovisual

Um outro fenómeno sugere que o percepto fonético não é puramente auditivo; trata-se do efeito de integração audiovisual, descoberto por McGurk e McDonald (1976).

Se se observa a imagem filmada de um locutor a pronunciar sílabas CV, e simultaneamente se ouvem sílabas CV diferentes, este conflito em geral passa despercebido e podem ouvir-se sílabas que não foram apresentadas nem visual nem auditivamente (cf. Liberman, 1982). Por exemplo, se a imagem corresponde a [gɛ], e o som sincronizado com a imagem é [bɛ], frequentemente os sujeitos percebem [dɛ]. Foi descrito um fenómeno análogo quanto à percepção de vogais: a imagem influencia o percepto vocálico tornando-o compatível com a informação visual (Summerfield e McGrath, 1984).

Apesar da integração audiovisual ser um tanto limitada (parece restringir-se à informação sobre o lugar de articulação, cf. Summerfield, 1979), ela sem dúvida sugere que perceber a fala não é só perceber sons.

Em suma, os factos experimentais que revimos sugerem que ao perceber a fala se recorre em certa medida, pelo menos, a formas de tratamento que não são apenas auditivas.

#### 4. Teoria motora e modularidade

A existência de um modo especial, fonético, para o processamento da fala foi defendida já nos anos 60 por Liberman e colaboradores. Esta proposta enquadra-se na teoria que desenvolveram para dar conta da percepção e produção da fala, vulgarmente designada por teoria motora.

A teoria motora tenta explicar como a partir da variabilidade acústica são extraídas unidades perceptivas invariantes. A utilização de índices acústicos dependentes do contexto só seria possível mediante o recurso a um processador especial para tratar a fala; este mecanismo

referiria a informação auditiva, variável e dependente do contexto, aos comandos motores necessários para a produzir.

Segundo esta teoria, o que é realmente percebido ao ouvir fala, ou o objecto da percepção, não se situa ao nível superficial da mensagem acústica; o percepto deve antes encontrar-se a nível de unidades invariantes de produção que, devido a sobreposições espaço-temporais dependentes do contexto, resultariam em sinais acústicos variáveis (Liberman, 1970; 1982; Liberman, Cooper, Harris e MacNeilage, 1962; Liberman, Cooper, Shankweiler e Studdert-Kennedy, 1967; Liberman e Mattingly, 1985; Liberman e Pisoni, 1977; Liberman e Studdert-Kennedy, 1978; Mattingly e Liberman, 1969; 1986).

#### 4.1. Aserções fundamentais (formulação inicial e teoria revista)

O artigo clássico onde a teoria é apresentada em detalhe, invariavelmente citado como fonte de inspiração mesmo pelos seus detractores, foi publicado com o título elucidativo "Perception of the Speech Code" (Liberman *et al.*, 1967). Os argumentos em favor do carácter *sui generis* da fala e do vínculo estreito entre percepção e produção aparecem com maior ou menor ênfase em outros escritos; alguns desses argumentos tornaram-se asserções clássicas sobre o funcionamento da percepção da fala (*e.g.*, Anisfeld, 1984, p. 199; Liberman, 1986; Miller, 1981, p. 64).

Na primeira formulação da teoria motora, foram defendidas as ideias que se seguem. (1) O ser humano não perceberia a fala da forma fácil e eficaz que conhecemos se cada fonema fosse veiculado por um som específico e imutável. Consegue-se sem esforço entender fala a um ritmo de 400 palavras por minuto, a que corresponde aproximadamente uma frequência de 30 fonemas por segundo. Todavia, a um tal ritmo,

acontecimentos acústicos discretos seriam percebidos sem individualidade, isto é, seriam fundidos num ruído contínuo (e.g., Warren, Obusek e Farmer, 1969). Deverá pois haver mecanismos de codificação e descodificação dos sons da fala que favorecem a possibilidade de os identificar com maior clareza do que os demais estímulos acústicos. (2) A informação fonética está presente na onda acústica de um modo altamente codificado; é justamente esta codificação que permite a eficiência de transmissão. (3) Para o auditor descodificar a mensagem, e recuperar informação fonémica, são necessários mecanismos ou processos específicos: "phoneme perception (...) requires a special decoder" (*ibidem*, p. 431). (4) Produção e percepção da fala submetem-se a um modelo hierárquico único, com níveis a que correspondem regras sintácticas, neuromotoras, miomotoras, articulatórias e acústicas. A codificação do fonema em sinal acústico dá-se após a emissão de comandos motores invariantes para os músculos, possivelmente a nível de regras articulatórias. A descodificação do sinal acústico em fonemas faz-se por referência a esses comandos invariantes: "speech is perceived by processes that are also involved in its production" (*ibidem*, p. 452). Em vez de dois processos independentes de produção e percepção, haveria um só processo codificador/descodificador, com as necessárias interligações entre componentes motores e sensoriais.

A teoria revista, tal como a apresentam Liberman e Mattingly (1985; cf. também Mattingly e Liberman, 1986), consiste essencialmente numa actualização com base nos resultados dos últimos 20 anos de pesquisa, em que são mantidas as proposições fundamentais: o vínculo percepção/produção, característico do modo de tratamento especial para a fala.

Talvez mais importante foi a descoberta de que existe considerável variabilidade nas trajectórias dos articuladores: a suposição de que a

invariância dos comandos neurais seria observável naquelas trajectórias, através de técnicas electro-miográficas, foi claramente infirmada. Esta suposição assentava numa visão simplificada da acção dos articuladores. Como notam Liberman e Mattingly (1985), pensava-se que cada gesto fonador seria realizado essencialmente por um só articulador, enquanto actualmente se aceita que cada gesto envolve em geral dois ou mais articuladores (por exemplo, no arredondamento dos lábios cooperam maxilar, lábios superior e inferior). As consequências da coarticulação e adaptação são assim ainda mais marcadas e insidiosas do que era pressuposto na versão inicial da teoria.

Importante parece ser também uma subtil alteração de perspectivas: em vez de se conceber em termos de posição dos articuladores (adoptando-se assim implicitamente uma perspectiva mais estática), concebe-se hoje mais em termos de sequências de gestos articulatorios, uma perspectiva mais dinâmica. Os invariantes, se os há, não podem ser encontrados a nível das contrações musculares que num determinado momento definem configurações espaciais específicas do tracto vocal: é necessário ter em conta a mudança das configurações gestuais ao longo do tempo, comparando diferentes gestos no mesmo contexto ou os mesmos gestos em diferentes contextos, na perspectiva de esclarecer as propriedades topológicas e dinâmicas típicas de cada gesto. A este propósito falam investigações recentes (e.g., Tuller e Kelso, 1984) que procuram determinar as invariâncias relacionais subjacentes à variabilidade dos actos motores.

Apesar de ser ainda imprescindível investigação básica sobre a invariância dos gestos motores, considerando-os numa perspectiva dinâmica (cf. Kelso, Saltzman e Tuller, 1986), a asserção de Liberman e Mattingly (1985, p. 23) é que "the gestures do have characteristic invariant properties, as the motor theory requires, though these must be seen, not as



peripheral movements, but as the more remote structures that control the movements. These structures correspond to the speaker's intentions." Estas estruturas remotas que controlam os movimentos e correspondem às intenções do locutor constituem, no dizer de Liberman e Mattingly (*ibidem*, p. 2), os objectos da percepção da fala. Ou seja, ao ouvir fala os perceptos não são de natureza acústica; são antes de natureza motora e talvez nem só: são os gestos intentados pelo locutor. Como funciona exactamente esta interacção não é explicado; é apenas proposta a arquitectura geral de um tal sistema: uma arquitectura modular.

A informação fonética seria pois percebida por um sistema específico, um módulo especializado em detectar os gestos intentados pelo locutor, que Liberman e Mattingly chamam o módulo fonético (*ibidem*). A noção de modularidade está actualmente no cerne da problemática em psicologia cognitiva. Abordá-la-emos já de seguida.

#### 4.2. Modularidade e percepção da fala

Recentemente, Fodor (1983; 1985) desenvolveu uma teoria sobre o sistema cognitivo cuja influência tem sido determinante no debate de ideias entre os investigadores da mente e comportamento humanos. Segundo esta teoria, o sistema cognitivo em sentido lato é constituído por dois tipos de processos, os sistemas centrais e os sistemas de *input* ou módulos. Os sistemas centrais são responsáveis pelos conhecimentos e crenças, pelo pensamento e planeamento da acção. Os módulos são sistemas altamente especializados, cada um com uma função específica e biologicamente relevante. Eles são os intermediários entre os transdutores sensoriais e os processos centrais; estão desenhados para a partir do *input* proveniente dos transdutores, provavelmente uma representação da estimulação proximal,

extrair propriedades ecologicamente relevantes, de modo a fornecerem aos sistemas centrais uma representação do objecto distal.

Os módulos caracterizam-se por um conjunto de propriedades que, no conjunto ou em parte, os distinguem dos sistemas centrais. Fodor propôs, entre outras, as seguintes: (1) especificidade de domínio: as operações de cada módulo aplicam-se somente a certas classes de acontecimentos; (2) especificidade de arquitectura neural: o suporte biológico do módulo não serve outras funções que não as suas; (3) estrutura especificada de modo inato, não sendo adquirida por aprendizagem; (4) desenvolvimento por maturação: o módulo desenvolve-se através de estágios em que é negligenciável a influência do meio; (5) autonomia: o módulo não partilha com outros sistemas os seus recursos atencionais, mnésicos ou outros; (6) funcionamento rápido e obrigatório: o módulo não está submetido à vontade ou controlo estratégico.

A relação entre os módulos e os sistemas centrais obedece a duas características. Primeiro, o acesso limitado: só os produtos finais do funcionamento do módulo ficam disponíveis para os sistemas centrais. O acesso dos sistemas centrais às representações intermediárias produzidas pelo módulo é muito reduzido. A segunda característica é designada por "informational encapsulation" (Fodor, 1983; p. 71), ou "encapsulamento" cognitivo. Este termo idiossincrático refere-se ao carácter estanque do módulo relativamente às representações dos sistemas centrais: o funcionamento do módulo não é afectado pela informação disponível a outros níveis. Segundo Fodor (*ibidem*), esta é a característica essencial do funcionamento modular: os sistemas de *input*, para cumprirem a sua função biológica, fornecem aos sistemas centrais representações do mundo como ele "é", e não como se quer ou espera que ele seja. Para expressar esta ideia de que os sistemas perceptivos são largamente insensíveis ao que o

sujeito deseja ou presume, Pylyshyn (1980) usa a expressão "impenetrabilidade cognitiva", que usaremos para nos referirmos a esta característica do módulo.

A utilização destes conceitos no contexto da teoria motora é discutida em pormenor por Liberman e Mattingly (1985) e Mattingly e Liberman (1986). Não é difícil ver como pelo menos algumas propriedades do módulo se aplicariam ao modo fala descrito por Liberman e colaboradores: suporte neural no hemisfério esquerdo, operação automática, rápida, de limitado acesso à consciência e cognitivamente impenetrável. Aqueles autores sustentam que o módulo fala é uma especialização com propriedades comuns a outras especializações que tratam acontecimentos ecologicamente significativos, quer em humanos (como a localização sonora através do tempo de chegada aos dois ouvidos, ou a estereoscopia através da disparidade retínica), quer em outros animais (localização sonora na coruja, eco-localização no morcego, canto nas aves). A percepção da fala é especial, como são também especiais outros sistemas biológicos adaptados a certos acontecimentos ecologicamente relevantes. A especificidade do módulo fala, tal como é concebido por Liberman e Mattingly, reside no vínculo íntimo entre percepção e produção.

Para o nosso trabalho, porém, é importante discutir a modularidade fora do contexto da teoria motora. A concepção modular da arquitectura cognitiva contrasta com o ponto de vista segundo o qual haveria um fluxo de informação mais livre entre sistemas de *input* e centrais. Ou seja, não haveria só um fluxo informativo do *input* para processos mais centrais, no sentido ascendente ("bottom-up"), haveria também troca de informações e influências mútuas no sentido descendente. Este último ponto de vista é denominado "interactivo", e desde a abordagem New Look dos anos 50 tem exercido uma influência considerável no desenvolvimento de modelos

em psicologia cognitiva. Ao propor uma dicotomia irreduzível entre módulos e sistemas centrais, Fodor (1983; 1985) veio obrigar a reconsiderar explicitamente a questão da interactividade entre processos de tratamento.

A ideia de que há processos de análise precoce da informação que são cognitivamente impenetráveis tem sido expressa repetidamente em vários domínios. Por exemplo, Neisser (1967) distinguiu entre um nível de processamento precoce, denominado pré-atentivo, que opera automaticamente e em paralelo, e um nível mais tardio, susceptível de controlo atencional. A distinção entre processos automáticos, rigidamente determinados e inacessíveis à consciência, e processos opcionais, flexíveis e mais ou menos conscientes é sem dúvida inatacável. A questão está em se a organização cognitiva humana se deixaria explicar pela existência de apenas dois tipos de processos, os "módulos" e os "centrais" e, se sim, onde acabariam uns e começariam os outros.

Mesmo numa situação de identificação de estímulo, em que é elaborado um percepto sem que sobre ele sejam efectuadas manipulações conscientes, estarão envolvidos apenas "módulos" ou estarão também implicados "processos centrais"? Segundo Fodor (*ibidem*), a própria percepção sofre influência de processos centrais; ele considerou mesmo a possibilidade de chamar aos módulos "sistemas perceptivos", mas abandonou esta nomenclatura por achar que a percepção é essencialmente um mecanismo de fixação de crença, e enquanto tal sofrer a influência de processos centrais.

A ideia de que existe uma dicotomia entre módulos e sistemas centrais é sem dúvida conceptualmente atraente; mas não simplificará ela a realidade a tal ponto que o que se ganhou em definição elegante de conceitos se perde em ambiguidades na análise concreta das situações? Por exemplo, o termo "análise pós-perceptiva" é usado por vezes como equivalente a

"análise pós-modular", estabelecendo assim, uma certa confusão entre modular e perceptivo. Talvez, como sugere Shallice (1984), a dicotomia entre módulos e sistemas centrais seja demasiado rígida.

Apesar destes problemas, ou talvez justamente por os suscitar, o quadro conceptual fornecido por Fodor pode desempenhar um papel central no esclarecimento de questões importantes em percepção e cognição, nomeadamente pelo seu valor heurístico. Como sugeriu Kolinsky (1988), as propriedades definidoras do módulo podem ser usadas para esclarecer relações hierárquicas entre níveis de processamento.

Ao longo deste trabalho usaremos tentativamente o conceito de modularidade no contexto dos fenómenos de percepção da fala sobre que nos debruçamos. De facto, a distinção entre processos obrigatórios e opcionais está subjacente a qualquer tentativa em esclarecer o papel de variáveis como a alfabetização no tratamento da fala. Haverá certamente processos perceptivos precoces e obrigatórios que não sofrem a influência da alfabetização, que poderíamos designar por modulares. A influência da alfabetização na percepção da fala dar-se-ia apenas ao nível de processos opcionais, mais ou menos acessíveis à consciência, mais ou menos impenetráveis cognitivamente, que poderíamos designar como não-modulares.

## CAPÍTULO III

## FUSÃO FONOLÓGICA: EXPLORAÇÕES EXPERIMENTAIS E TEÓRICAS

Após termos apreciado alguns problemas gerais sobre a percepção de fala e discutido com algum pormenor aspectos mais relacionados com o nível fonético, passamos agora para um fenómeno que se situa a nível fonológico.

Neste capítulo, discutiremos aprofundadamente a fusão fonológica. Há duas razões para o fazermos: primeiro, trata-se de um facto experimental sobre que se colocam uma série de problemas interessantes e que, tanto quanto sabemos, não foi objecto de uma revisão crítica na literatura; segundo, uma das nossas experiências visa averiguar aspectos específicos da fusão fonológica (Experiência VII).

Começaremos por apresentar as características gerais deste fenómeno, e por rever os resultados experimentais conhecidos, analisando o papel de variáveis de tipo acústico-auditivo e de tipo linguístico. Passaremos então a uma discussão centrada no problema da interpretação e explicação da fusão fonológica. Situa-la-emos no contexto de outros fenómenos de integração de informação auditiva, e examinaremos com pormenor uma questão polémica, a das diferenças individuais quanto à tendência em fundir informação fonológica. Finalmente, tentar-se-á com base nos dados disponíveis na literatura esclarecer se a fusão fonológica provirá de um processo puramente automático, ou será susceptível de controlo atencional.

## 1. Características gerais da fusão fonológica

Suponhamos que a palavra "cara" é apresentada, por meio de auscultadores, a um ouvido, enquanto ao mesmo tempo e no outro ouvido se apresenta a palavra "lara". Se pedirmos ao sujeito que nos relate o que ouve, é possível que ele responda "clara". Assim, numa situação de audição dicótica, ocasionalmente o percepto pode ser descrito como uma combinação dos fonemas iniciais dos estímulos em competição.

Entre as primeiras observações conhecidas sobre este fenómeno, contam-se as de Day (1968), que obteve, por exemplo, respostas "blanket" aos estímulos dicóticos "banket" e "lanket", ou "glorigin" perante o par "gorigin + lorigin". A inexistência de respostas erradas de tipo "lbanket" ou "lgorigin", bem como a inexistência de fenómenos deste tipo quando os estímulos em competição são mais dissimilares ou não obedecem a determinadas regras (*e.g.*, perante "four" + "six" não se ouve "fsix" ou "fsour") indica que a combinação dos fonemas não se dá arbitrariamente; parece antes depender de regras fonológicas da língua, no sentido de que só se "fundem" fonemas que constituem encontros consonantais aceitáveis num determinado sistema fonológico. Por este motivo, denominou-se fusão fonológica esta combinação de fonemas que aparece a nível de percepto (*ibidem*; Cutting, 1975). Simplificando, diz-se que há fusão fonológica quando dois estímulos separados e em competição, cada um com  $n$  fonemas, produzem um percepto linguisticamente mais complexo, com  $n+1$  fonemas (Cutting, 1976; Poltrock e Hunt, 1977; Sexton e Geffen, 1981).

Logo os primeiros casos observados por R. Day sugeriam que a fusão fonológica não é devida exclusivamente a uma tendência em

responder com palavras em vez de não-palavras\*. Não parece ser por mero viés de resposta que se ouve *e.g.*, "product" em vez de /pahduct/ ou /rahduct/, pois os sujeitos identificavam correctamente estas pseudo-palavras em apresentação monaural (Day, 1968). Além disto, ocorrem fusões de tipo pseudo-palavra com base em duas pseudo-palavras (por exemplo "glorigin", em resposta a "gorigin + lorigin") e há também fusões quando estímulos são eles próprios palavras: por exemplo, "back + lack" são por vezes ouvidos como "black" (Day, 1968), "pato + rato" como "prato" (cf. Experiência VII). A fusão fonológica não é portanto um fenómeno "trivial" de selecção de resposta; justifica-se que sejam empreendidos trabalhos empíricos para esclarecer a sua natureza.

## 2. Explorações experimentais

### 2.1. Introdução: fala natural e fala sintética

A tarefa a que se propuseram J. Cutting e R. Day foi a de esclarecer as condições que afectam a ocorrência de fusões fonológicas. Uma questão teoricamente relevante é se este fenómeno é mais ou menos sensível a variáveis de tipo estritamente auditivo-acústico, ou se nele se faz sentir com maior ou menor impacto o efeito de factores fonéticos e linguísticos. Embora sem perfilhar a irredutibilidade desta distinção, e por questão de conveniência, utilizá-la-emos ao rever os resultados experimentais. Antes, porém, vejamos uma questão prévia.

Quando se pretende averiguar com pormenor o efeito de factores de tipo acústico, é indispensável poder manipulá-los com uma precisão e

---

\* Usaremos "não-palavra" na acepção genérica, que engloba as pseudo-palavras - sequências fonológicas que obedecem às regras fono-tácticas da língua, por exemplo /bosa/, e as não-palavras propriamente ditas - sequências fonéticas inadmissíveis na língua, por exemplo /mfid/.

independência que se encontram mais facilmente no uso de sintetizadores do que na própria voz humana. Assim, simultaneamente uma condição para investigação subsequente, e já um factor cujo efeito é interessante averiguar, temos a comparação da incidência de fusões fonológicas com fala natural e com fala sintética.

Esta comparação foi efectuada com duas séries análogas (versões naturais e sintéticas) de estímulos constituídos por palavras, cuja combinação dos primeiros fonemas resultava também numa palavra (Cutting, 1975; Experiência 1). Trata-se de quatro conjuntos, o de "pay" constituído por "pay, lay, ray", o de "bed", constituído por "bed, led, red", o de "cam", constituído por "cam, lamb, ram" e o de "go", constituído por "go, low, row". A metodologia seguida foi a mais corrente nos estudos deste fenómeno: o sujeito desempenha duas tarefas, uma de identificação dos estímulos individuais, que visa garantir a inteligibilidade de cada estímulo, e outra tarefa de identificação dos pares fundíveis\*.

A incidência de fusões foi duas vezes maior com fala sintética, (61%), do que com fala natural (apenas 31%); esta razão de 2:1 foi altamente significativa. Por outro lado, a configuração das respostas era semelhante para os pares sintéticos e os naturais: algumas palavras fundiam melhor que outras e, em geral, as fusões eram mais frequentes com a líquida /l/ do que com /r/. Não se verificou efeito de ouvido, nem de ordem (começar pela fala sintética ou fala natural), nem de assincronia do início dos estímulos (nas condições de apresentação síncrona, ou com 50 milissegundos de atraso da líquida ou da oclusiva). Assim, a diferença entre fala sintética e fala natural é meramente quantitativa (Cutting, 1975). Parece aceitável usar fala

---

\*Na discussão de todos os trabalhos que segue, salvo em casos explicitamente referidos, fica subentendido que a identificação de cada estímulo se fez sem dificuldades e obedeceu ao padrão esperado.

sintética como meio de favorecer a ocorrência de fusões, e generalizar os resultados obtidos à fala natural.

## 2.2. Variáveis de tipo acústico-auditivo

### 2.2.1. Altura tonal, intensidade e tracto vocal

Dentro de certos limites, factores como a altura tonal e a intensidade não afectam a ocorrência de fusões fonológicas. Construindo pares de estímulos em que um fosse agudo e outro grave (respectivamente, com uma frequência fundamental de 140 Hz ou de 120 Hz) encontrou-se exactamente a mesma proporção de fusões, 50%, do que quando ambos os estímulos do par tinham a mesma altura tonal (Cutting, 1975, Experiência 4). Este resultado, originalmente observado com palavras ("pay", "lay", "ray" e "kick", "lick", "rick"), foi replicado e expandido a diferenças de 2, 20 e 80 Hz na frequência fundamental de sílabas apresentadas dicoticamente (/ba-la/ e /da-ra/); independentemente da altura tonal, a percentagem de fusões mantinha-se a cerca de 75% (Cutting, 1976).

Quanto à intensidade, em experiências análogas às anteriores, verificou-se que pares de palavras em que um elemento fosse 15 dB mais forte que o outro fundiam tanto como pares de intensidade idêntica, cerca de 36% (Cutting, 1975); usando sílabas, e diferenças de intensidade entre 0 e 40 dB, em passos de 5 dB, verificou-se que só a partir de uma diferença de 30 dB havia uma redução no número de fusões fonológicas (Cutting, 1976). Estas são pois relativamente pouco afectadas por esta variável física.

Outra manipulação considerada foi o tamanho do tracto vocal. Sintetizando estímulos como se fossem produzidos por um tracto vocal adulto masculino, ou por um locutor cujo tamanho do tracto vocal fosse

apenas 5/6 do anterior, foi possível emparelhar palavras diferentes, como se fossem pronunciadas por um adulto e por um anão. Comparando a percentagem de fusões em pares de palavras de tractos vocais diferentes ou do mesmo tracto vocal, verificou-se ser nos dois casos semelhante, 59% (Cutting, 1975).

Num teste mais exigente quanto à relativa independência das fusões fonológicas daqueles factores físicos, Cutting (1975) constituiu pares de estímulos em que nenhum dos valores de altura tonal, intensidade ou tracto vocal fosse idêntico (Experiência 7). Corroborando os resultados anteriores, aquelas diferenças tiveram um efeito desprezível na percentagem de fusões: esta foi de 47% nos pares diferentes, e de 44% nos pares idênticos.

### 2.2.2. Assincronia: precedência da oclusiva ou da líquida

Potencialmente relevante será o facto de os membros de cada par serem apresentados com uma ligeira assincronia, tal que, por exemplo, o início da oclusiva preceda em alguns milisegundos o início da líquida. Os primeiros trabalhos sobre fusão fonológica indicavam que, numa faixa de 150 milisegundos, ms, de assincronia, mesmo quando a líquida precedia a oclusiva, não havia diferença digna de nota na percentagem de fusão (Day, 1968; Cutting e Day, 1975). Todavia, estudos posteriores vieram mostrar que o quadro não é assim tão simples.

Com assincronias entre 0 e 800 milisegundos\*, as fusões são mais frequentes quando a oclusiva precede a líquida em 50 e 100 ms (cerca de 60%); com assincronias maiores, e quando é a líquida a preceder a oclusiva,

\*Assim intervaladas: 0, ±50, ±100, ±200, ±400 e ±800 (Cutting, 1975, p. 109).

a percentagem de fusões desce para 40 a 10% (Cutting, 1975, Experiência 3). Usando sílabas (/da-la/ e /da-ra/) e assincronias entre 0 e 160 ms,\* Cutting (1976) observou o mesmo padrão de resultados: mais fusões com ligeiras assincronias de 20 e 40 ms, nomeadamente se a oclusiva precede a líquida, e menos fusões com assincronias de 80 e 160 ms. Finalmente, usando apenas assincronias de 25, 50 e 75 ms, Poltrock e Hunt (1977) confirmaram que há mais fusões quando a oclusiva precede a líquida; 74% *vs* 68% se a líquida precede a oclusiva (trata-se de uma diferença pequena, mas significativa). Corroborando resultados anteriores, naquela faixa relativamente estreita não se observaram efeitos da magnitude da assincronia propriamente dita, independentemente da sua direcção. A fusão fonológica apresenta-se pois relativamente resistente aos efeitos de assincronias temporais até cerca de 100 milisegundos, e é favorecida se a oclusiva precede a líquida.

Estes resultados são particularmente reveladores (ou talvez intrigantes), na medida em que parecem sobrepôr-se ao efeito de máscara que tipicamente ocorre quando o intervalo inter-estímulo é muito pequeno. Por exemplo, se o estímulo "B" é apresentado 50 ms depois do estímulo "A", aquela apresentação vai interferir com o processamento do primeiro estímulo, um fenómeno conhecido por máscara retroactiva (cf., *e.g.*, Pisoni, 1973; Massaro, 1974). Ora o padrão que se observa quanto à ocorrência de fusões fonológicas mostra uma curiosa interacção. Se a oclusiva precede a líquida até 100 ms, praticamente não se observam efeitos de máscara retroactiva: o percepto não é o elemento iniciado por líquida; pelo contrário, o segundo estímulo é frequentemente combinado com o primeiro, gerando um percepto de tipo fusão. Todavia, se é a líquida a preceder a oclusiva, observa-se o fenómeno de máscara retroactiva: o percepto é a palavra iniciada por

\* Assim intervaladas: 0,  $\pm 50$ ,  $\pm 10$ ,  $\pm 20$ ,  $\pm 40$ ,  $\pm 80$  e  $\pm 160$  (Cutting, 1975, p. 127)

oclusiva cerca de metade das vezes; na outra metade há fusões (Cutting, 1975).

Parece pois que quando a informação chegada em segundo lugar é compatível fonologicamente com a informação que está a ser processada, ambas vão ser fundidas gerando um percepto único. A compatibilidade fonológica inibiria a máscara retroactiva.

### 2.3. Variáveis de tipo linguístico

Abordaremos nesta rubrica um conjunto mais ou menos heterogéneo de factores, seguindo um percurso que vai dos mais molares para os mais moleculares.

#### 2.3.1. Nível semântico: estatuto lexical, contexto de frase e frequência da palavra

Situemo-nos a nível semântico e consideremos o estatuto lexical dos estímulos. Apesar de fusões fonológicas ocorrerem também com pseudo-palavras, elas são particularmente frequentes quando dois estímulos não-palavra formam uma palavra (Day, 1968; 1970). Este facto coloca a questão do possível enviesamento para responder palavras, que poderia favorecer a ocorrência de fusões. Retomaremos já a seguir este aspecto.

Ainda a nível semântico, averiguou-se (Cutting, 1975, Experiência 8) se estímulos inseridos numa frase seriam tão ou mais fundíveis do que os mesmos estímulos apresentados isoladamente. Por exemplo, apresentando a um ouvido a frase "the coals are (go)ing again" e ao outro "the coals are (low)ing again", há a possibilidade de o sujeito responder com "glowing".

Usaram-se dois conjuntos de estímulos, "pay, lay, ray" e "go, low, row"\*. Ora os pares de palavras suscitaram 63% de fusões se apresentados isoladamente, e 89% se inseridos na frase-contexto. Serão estes resultados interpretáveis à luz de factores estritamente semânticos? É que respostas de fusão tornam a frase mais plausível do que respostas de não fusão\*\*. Ora as fusões de tipo oclusiva + /l/ foram sempre mais frequentes, mesmo em contexto semântico desapropriado, como "the trees are glowing again" - as árvores brilham de novo (83% das respostas; só 4% das respostas foram com a fusão apropriada, "growing"), e "the minister plays for us" - o sacerdote toca música para nós (74% das respostas, *vs* 16% com "prays"). Em contrapartida, não houve fusões semanticamente deslocadas que envolvessem o /r/: nenhum sujeito percebeu, *e.g.*, que "the trumpeter prays for us"... O facto de se manter a preponderância de fusões com /l/, apesar de semanticamente inapropriadas, é um indício forte de que o que está em jogo não é responder o que se julga mais correcto (com uma palavra em vez de uma não-palavra, com uma palavra congruente em vez de uma palavra semanticamente desapropriada).

Uma observação comum é que a incidência de fusões varia conforme o par de palavras. Cutting (1975) avançou a sugestão que isto se poderia dever, em parte, a um efeito de frequência de palavras: se a fusão resultasse numa palavra frequente, ela ocorreria mais vezes do que se resultasse numa palavra pouco frequente. De facto aquele autor observou resultados deste tipo: foram mais frequentes as fusões "play, pray, glow, grow" do que "bled, bread, clam, cram", grupos de palavras que são, respectivamente, mais e menos usados em publicações de carácter geral.

\* O conjunto "go, low, row" usava-se com as frases "the coals are ...ing again" e "the trees are ...ing again"; com o conjunto "play, lay, ray," usavam-se as frases "the minister ...s for us" e "the trumpeter ...s for us".

\*\* Faz mais sentido ouvir "o carvão brilha de novo" ("is glowing") do que "o carvão vai de novo" ("is going"), ou "cresce de novo"; é igualmente mais plausível que o sacerdote reze por nós ("the minister prays ...") do que pague por nós ("the minister pays ...")

Todavia, este resultado não foi replicado por Cutting e Day (1975), nem o factor mereceu a atenção de outras investigações.

### 2.3.2. Nível da palavra: posição do *cluster*, semelhança fonológica e variação alofónica

Até aqui temos apreciado fusões ocorrendo na primeira sílaba; todavia, em observações piloto, Day (1968) tinha verificado que podiam também verificar-se fusões na última sílaba (e.g., "gif/git → gift), ou em sílabas intermédias (e.g., "demonstration/demonstration → demonstration"). Num exame mais sistemático à eventual influência da posição dos *clusters*, Cutting e Day (1975) observaram um efeito pronunciado daquela variável: para o conjunto de estímulos usados, a percentagem de fusões foi de 51% em posição inicial e 8% em posição final. Uma observação digna de nota é que mesmo com os estímulos que supostamente só deveriam permitir fusões na posição final, houve algumas respostas (cerca de 5%) em que a posição da líquida foi alterada, produzindo-se fusões em posição inicial; dados, e.g., os estímulos "paid", [peid] e "pail", [peil], resposta de "played", [pleid]. Estas respostas, envolvendo troca de posição da líquida, foram quase tão frequentes como as fusões em posição final, como "paled".

A interpretação oferecida pelos autores apela para a noção de variação dependente do contexto. Os estímulos sintéticos usados tinham a mesma informação acústica nos dois primeiros fonemas, o que não acontece nas produções normais das palavras em que estão inseridos; por outro lado, as vogais naturais das palavras com encontro consonantal em posição final têm uma qualidade diferente dos /i/ sintetizados. Sem dúvida, o carácter "estranho" da fala sintética, e a sua própria constituição acústica, terão alguma influência no percepto. O facto de os fonemas e/ou sílabas não serem exactamente os mesmos, acústica e perceptivamente, consoante o contexto,

dificulta (impedirá até?) que se possa avaliar o efeito da posição *per se*. Uma dificuldade adicional provém das próprias regras fonológicas da língua: numa mesma língua, não se encontram necessariamente todos os padrões fonológicos requeridos para uma comparação controlada do efeito de posição. Em inglês, por exemplo, /ld/ é permissível em posição final, mas não inicial, e vice-versa para /dl/.

O facto de, mesmo em reduzido número, haver fusões que não respeitam integralmente a sequência fonológica dos estímulos ([peid] + [peil] → [pleid], por exemplo) levanta a questão de até que ponto a semelhança fonémica das palavras é suficiente para produzir fusões. Observações preliminares de Day (1968) sugeriam que era possível haver fusões com estímulos "parecidos", em vez de idênticos excepto num fonema. Isto parecia ocorrer mais facilmente quando fosse possível um encontro consonantal no início da palavra, *e.g.*, "catch + lift" → "clatch"), embora outros casos também fossem observados ("squirrel + turtle" → "squirtle").

Numa experiência expressamente desenhada para averiguar o efeito da semelhança dos pares de palavras, comparou-se a incidência de fusões em pares idênticos, do tipo "bed/led/red", e dissimilares do tipo "bed/low/row"; não se encontraram diferenças significativas quanto à percentagem de fusões num caso ou noutro, respectivamente 40% *vs* 45% (Cutting e Day, 1975; Experiência 4). Analisando mais de perto as respostas dos sujeitos aos pares dissimilares, verifica-se que há mais respostas duplas, em que além da fusão há um outro percepto correspondente a um dos estímulos (por exemplo, "blow" e "low"). Portanto, dentro de condições comparáveis em termos de permissibilidade de encontros consonantais, a dissemelhança fonémica afecta mais o número de respostas do que a incidência de fusões. Este resultado indica que para haver fusão fonológica não é necessário que os estímulos sejam foneticamente semelhantes

(idênticos excepto num fonema); pelo contrário, é necessário, e nalguns casos, suficiente, que os estímulos sejam compatíveis em termos fonológicos.

Apoiando a conclusão acima formulada, verificou-se que a variação alofónica não afectava a incidência de fusões (Cutting e Day, 1975; Experiência 2). Estes autores constituíram pares de palavras com "pay, lay ray" e "kick, lick, rick", em que /l/ e /r/ foram sintetizados em duas versões, uma normal, como líquidas inglesas, e outra alofónica, sob forma de trilos\*. A percentagem de fusões foi de 43% nos ensaios com as líquidas normais, e de 50% nas versões alofónicas. A qualidade dos perceptos nem sempre era idêntica nos dois casos: frequentemente os sujeitos respondiam "prray" ( a resposta era dada por escrito) quando o estímulo era o trilo, o que mostra que a manipulação efectuada era perceptivamente saliente, apesar de não exercer efeito sobre a incidência de fusões.

### 2.3.3. Nível de fonema: o efeito de substituição por /l/, fusibilidade de diferentes fonemas

Os trabalhos até agora revistos têm usado encontros consonantais que envolvem a líquida. Um dos resultados mais consistentes é que a fusão ocorre mais facilmente com /l/ do que com /r/. Além disto, as respostas dos sujeitos frequentemente são de oclusiva + /l/ mesmo perante estímulos de tipo oclusiva + /r/. Este fenómeno é tão reprodutível e saliente que foi baptizado com um nome próprio, o efeito de substituição por /l/ ("l/-substitution effect", cf. Day, 1968; Cutting, 1975; Cutting e Day, 1975). Parece ocorrer menos com palavras multi-silábicas, ou diferindo em duração,

\*O /r/ "trilado" não é uma líquida, aproxima-se do /r/ francês e é inexistente no inglês; o /l/ "trilado" é um som artificial, talvez impossível de ser produzido mesmo com muita "ginástica articulatória", como dizem Cutting e Day (1975, p. 103).

timbre vocálico e contorno; é pelo contrário favorecido pela semelhança acústico-fonética dos estímulos (Day, 1968; Cutting, 1975).

Como uma proporção substancial das respostas de fusão é constituída por casos em que ocorre a substituição do /r/ por /l/ (cf., e.g., Cutting e Day, 1975), é importante esclarecer se este fenómeno poderá ser atribuído exclusivamente a factores de resposta: poderia acontecer que indecisos entre "pray" e "play" os sujeitos considerem a segunda mais apropriada, ou mais clara que a primeira. Por outras palavras, talvez os sujeitos consigam distinguir os dois perceptos e escolham um deles para a sua resposta. Cutting e Day (1975; Experiência 5) indicam que não é isto que se passa. Propondo aos sujeitos, além das habituais tarefas de identificação em monaural e em dicótico, duas tarefas de discriminação, uma dos pares dicóticos e outra das líquidas, verificou-se que os sujeitos discriminam o /l/ do /r/, mas ficam-se pelo nível do acaso na discriminação dos pares dicóticos. Com um paradigma de tipo ABX, em que as tríades seriam, por exemplo, "pay/lay" - "pay/ray" - "pay-ray", o nível de desempenho rondava os 54%.

Talvez as raízes do efeito de substituição por /l/ se encontrem a nível acústico: ao examinar espectrogramas de emissões naturais de "play" e "pray", nota-se uma descontinuidade na intensidade e na estrutura de formantes antes do início das transições ascendentes apropriadas ao /l/. Esta descontinuidade não aparece em "pray" (Cutting e Day, 1975). A interpretação avançada por aqueles autores é que a apresentação dicótica poderia introduzir descontinuidades análogas, e apresentam em seu apoio o facto de a apresentação binaural de pares de estímulos misturados electricamente reduzir o número de substituições deste tipo: a percentagem de fusões provenientes de substituições por /l/ desceu de 77% em apresentação dicótica para 43% em apresentação binaural (Cutting, 1975;

Experiência 2). Porém, outra interpretação possível será o facto de o estímulo funcional para a ocorrência de fusão fonológica ser o segundo formante (ver mais à frente), que é semelhante em /l/ e /r/.

O facto de, independentemente do efeito de substituição por /l/, ocorrerem mais fusões em resposta a pares com oclusiva e com /l/ coloca a questão do grau de "fusibilidade" dos diferentes fonemas, que foi abordada em pormenor em duas experiências de Cutting (1975; Experiência 9 e 10). Comparando a percentagem de fusões em pares de oclusivas com líquidas, ou fricativas com líquidas, foi possível verificar que a percentagem de fusões difere, sendo muito maior no primeiro caso. A diferença encontrada foi de 3:1 (57% e 18%), e todos os sujeitos exibiam a diferença neste sentido.

Uma dúvida que se nos pode colocar é se este resultado não estaria confundido com outros factores, nomeadamente a frequência dos próprios encontros consonantais com fricativas, ou das palavras que os têm relativamente às palavras com oclusiva/líquida. De facto, estas variáveis não dão conta do padrão observado: em inglês, encontros consonantais com fricativa e líquida em posição inicial não são menos frequentes que os de oclusiva e líquida e, nos estímulos usados, as fusões possíveis com fricativas eram palavras ainda mais frequentes do que as fusões possíveis com oclusiva - *e.g.*, "fled" e "flow" vs "bled" e "glow" (cf. Cutting, 1975, p. 114). Um aspecto importante, e particularmente revelador neste contexto, é que as ditas fusões com fricativas consistiam maioritariamente em respostas de tipo oclusiva/líquida; ou seja, sendo os estímulos, *e.g.*, "fed" de um lado, "red" do outro, o sujeito ouvia "bred". De facto, cerca de 70% das fusões perante fricativa eram deste tipo.

Outra comparação feita por Cutting (*ibidem*; experiência 10) envolveu líquidas e semivogais, numa tentativa de esclarecer aspectos mais específicos do fenómeno da substituição por /l/. Com dois grupos de

estímulos, "kick, lick, rick, wick", e "coo, lieu, rue, you", houve cerca de 56% de fusões, cuja esmagadora maioria (89%) eram as palavras "click" e "clue". Assim, quer o /r/ quer as semivogais foram substituídas pelo /l/; mais uma vez, isto não se deve à frequência das palavras, pois, por exemplo, "quick" e "crew" são mais comuns que "click" e "clue" (Cutting, 1975).

Estes resultados são consistentes com os revistos anteriormente e que mostraram que a fusão fonológica é relativamente insensível à variação fonética dos estímulos, prevalecendo certas características fonológicas das respostas: estas consistem geralmente em encontro consonantal de oclusiva e líquida.

#### 2.3.4. Nível fonético-acústico: o papel das transições de formante e "chirps"

Poder-se-á explicar a maior "fusibilidade" das oclusivas com as líquidas através da sua estrutura acústica? Uma primeira observação é que estímulos com transições menos pronunciadas, como as semivogais, deram também origem a fusões em que o percepto era de tipo oclusiva e líquida. Talvez então as transições de formante não sejam relevantes na determinação do percepto. Esta hipótese pode testar-se comparando a incidência de fusões sob as condições habituais (*e.g.*, "pay + lay") ou variando o declive das transições de formante até ao extremo de inexistência de transição (*e.g.*, "pay + "ay"). Esta experiência foi realizada (Cutting, 1975; Experiência 10) e os resultados mostraram claramente que as respostas de tipo fusão ocorrem tanto mais quanto mais pronunciada for a transição, entre 52% com os estímulos que haviam sido previamente identificados como /l/, e 6% para os restantes estímulos tipo vogal. Como estes estímulos tinham sido previamente ouvidos como /l/ 8% das vezes, torna-se evidente que o grau

de fusão depende dos estímulos envolvidos suscitarem isoladamente o percepto de líquida, ou não.

Assim, as transições de formante são um aspecto da estimulação importante para suscitar fusões. Uma nova experiência destinada a elucidar que transição de formante, ou combinação de transições, é mais pertinente, mostrou claramente que no caso do /l/ é a transição do segundo formante, e no caso do /r/, é a transição do terceiro formante que mais peso têm na determinação das fusões fonológicas. Apresentando dicoticamente uma oclusiva e apenas o segundo formante de /l/, a percentagem de fusões era de 60% (vs 19% com o primeiro formante, e 16% com o terceiro correspondente a /l/); emparelhando a oclusiva com o terceiro formante, a percentagem de fusões era de 64%\*. São pois as transições de formante situadas entre 1000 e 2000 Hz o aspecto da estimulação que desencadeia o fenómeno da fusão fonológica.

Tendo analisado quais os constituintes acústicos relevantes para suscitar respostas de fusão, cabe perguntar até que ponto o fenómeno assenta em interacções auditivo-acústicas propriamente ditas, ou pelo "contrário", envolve estes factores enquanto aspectos do estímulo fala. Embora esta distinção esteja longe de ser simples, é possível abordá-la tentativa e pragmaticamente usando os "chirps", ou as transições de formante isoladas, que são percebidas como ruído e não como fala. Cutting (1975; Experiência 13) comparou a percentagem de fusões em pares de oclusivas com líquidas, e pares de oclusivas com "chirps"; enquanto para os primeiros houve 47% de respostas de fusão, para os segundos aquela percentagem desceu para 8%. Vendo separadamente a percentagem de fusão para cada tipo de "chirp", nota-se que ela é ligeiramente maior nos

---

\* Lembremos que a distinção entre as líquidas inglesas /l/ e /r/ se faz essencialmente, e se pode fazer exclusivamente em fala sintética, através do terceiro formante.

"chirps" ascendentes (como são as transições de formante da palavra total); principalmente para o "chirp" correspondente ao segundo formante (entre 1.2 e 1.8 KHz), onde atingem 24%. Assim, nem a transição mais apropriada ao fonema líquido "consegue" suscitar um nível de fusões semelhante ao obtido com o estímulo íntegro.

### 3. Explorações teóricas

#### 3.1. A fusão fonológica no contexto de outros fenómenos de fusão de sons

O retrato que acabámos de traçar apresenta-nos a fusão fonológica como um fenómeno relativamente insensível a aspectos não-linguísticos, como a altura tonal, intensidade, timbre e assincronia de apresentação; em contrapartida, características fonéticas, como a presença ou ausência de transições de formante, características fonético-fonémicas, como o grau de fusibilidade dos fonemas ou a semelhança fonológica, e características semânticas, como o estatuto lexical, o contexto e a frequência das palavras, ora favorecem, ora inibem a ocorrência daquele fenómeno. Uma visão sucinta e simplificada é que factores linguísticos e não-linguísticos exercem um efeito diferencial na fusão fonológica. Na medida em que ela é menos afectada por factores não-linguísticos, e sensível aos linguísticos, Cutting (1976) adiantou que os processos envolvidos são de natureza central.

Aquela asserção torna-se mais clara se compararmos a fusão fonológica com outros fenómenos que, de um ou outro modo, combinam características da estimulação acústica: a localização da fonte sonora, as fusões psico-acústicas, espectral e espectro-temporal, e a fusão de traços fonéticos ou *blending*. De facto, o termo fusão é aplicado a qualquer destes fenómenos, e uma comparação sistemática entre eles indubitavelmente esclarece quer a natureza da "fusão", quer as peculiaridades de cada uma

das "fusões". Para ser sistemática (e reveladora), uma tal comparação deverá incidir sobre o efeito que uma determinada variável exerce nos vários tipos de fusão; se esse efeito for diferencial (influindo numas fusões e não noutras, por exemplo), então poder-se-á inferir sobre a diferença nos processos invocados e eventualmente conceber as fusões no seio de uma organização percepto-cognitiva de diferentes níveis.

Usando vários modos de apresentação auditiva, Cutting (1976) comparou a robustez da resposta de fusão perceptiva face a variações de três tipos: temporal - assincronia da apresentação dos estímulos aos dois ouvidos, espectral - diferença entre as frequências fundamentais dos estímulos apresentados a um e ao outro ouvido, e de intensidade - *idem, mutatis mutandi*. Vejamos os resultados desse trabalho; começemos por caracterizar sucintamente cada uma das fusões consideradas.

Consoante a localização da fonte sonora, o som por ela emitido pode chegar aos nossos ouvidos com ligeiras diferenças temporais e de intensidade. Compare-se, de modo meramente esquemático, a propagação de um som a partir de uma fonte situada à nossa frente com o de outra fonte, situada exactamente à nossa esquerda. Neste segundo caso, o som tem de percorrer uma distância maior para chegar ao ouvido direito e tem de ultrapassar o obstáculo causado pela própria cabeça do auditor: assim, o som chega ao ouvido direito ligeiramente mais tarde, e com uma intensidade relativamente menor (principalmente nas frequências elevadas). Estas diferenças mínimas são "aproveitadas" pelo nosso sistema auditivo, que em vez de nos levar a perceber diferenças no tempo de chegada ou no timbre de dois sons, nos fornece um percepto único cuja origem espacial podemos localizar. *I. e.*, as mensagens acústicas que chegam ao ouvido direito e esquerdo são integradas num percepto único com informação espacial (cf., *e.g.*, Lindsay e Norman, 1979, pp. 178-184).

Experiência 2). Porém, outra interpretação possível será o facto de o estímulo funcional para a ocorrência de fusão fonológica ser o segundo formante (ver mais à frente), que é semelhante em /l/ e /r/.

O facto de, independentemente do efeito de substituição por /l/, ocorrerem mais fusões em resposta a pares com oclusiva e com /l/ coloca a questão do grau de "fusibilidade" dos diferentes fonemas, que foi abordada em pormenor em duas experiências de Cutting (1975; Experiência 9 e 10). Comparando a percentagem de fusões em pares de oclusivas com líquidas, ou fricativas com líquidas, foi possível verificar que a percentagem de fusões difere, sendo muito maior no primeiro caso. A diferença encontrada foi de 3:1 (57% e 18%), e todos os sujeitos exibiam a diferença neste sentido.

Uma dúvida que se nos pode colocar é se este resultado não estaria confundido com outros factores, nomeadamente a frequência dos próprios encontros consonantais com fricativas, ou das palavras que os têm relativamente às palavras com oclusiva/líquida. De facto, estas variáveis não dão conta do padrão observado: em inglês, encontros consonantais com fricativa e líquida em posição inicial não são menos frequentes que os de oclusiva e líquida e, nos estímulos usados, as fusões possíveis com fricativas eram palavras ainda mais frequentes do que as fusões possíveis com oclusiva - e.g., "fled" e "flow" vs "bled" e "glow" (cf. Cutting, 1975, p. 114). Um aspecto importante, e particularmente revelador neste contexto, é que as ditas fusões com fricativas consistiam maioritariamente em respostas de tipo oclusiva/líquida; ou seja, sendo os estímulos, e.g., "fed" de um lado, "red" do outro, o sujeito ouvia "bred". De facto, cerca de 70% das fusões perante fricativa eram deste tipo.

Outra comparação feita por Cutting (*ibidem*; experiência 10) envolveu líquidas e semivogais, numa tentativa de esclarecer aspectos mais específicos do fenómeno da substituição por /l/. Com dois grupos de

estímulos, "kick, lick, rick, wick", e "coo, lieu, rue, you", houve cerca de 56% de fusões, cuja esmagadora maioria (89%) eram as palavras "click" e "clue". Assim, quer o /r/ quer as semivogais foram substituídas pelo /l/; mais uma vez, isto não se deve à frequência das palavras, pois, por exemplo, "quick" e "crew" são mais comuns que "click" e "clue" (Cutting, 1975).

Estes resultados são consistentes com os revistos anteriormente e que mostraram que a fusão fonológica é relativamente insensível à variação fonética dos estímulos, prevalecendo certas características fonológicas das respostas: estas consistem geralmente em encontro consonantal de oclusiva e líquida.

#### 2.3.4. Nível fonético-acústico: o papel das transições de formante e "chirps"

Poder-se-á explicar a maior "fusibilidade" das oclusivas com as líquidas através da sua estrutura acústica? Uma primeira observação é que estímulos com transições menos pronunciadas, como as semivogais, deram também origem a fusões em que o percepto era de tipo oclusiva e líquida. Talvez então as transições de formante não sejam relevantes na determinação do percepto. Esta hipótese pode testar-se comparando a incidência de fusões sob as condições habituais (*e.g.*, "pay + lay") ou variando o declive das transições de formante até ao extremo de inexistência de transição (*e.g.*, "pay + "ay"). Esta experiência foi realizada (Cutting, 1975; Experiência 10) e os resultados mostraram claramente que as respostas de tipo fusão ocorrem tanto mais quanto mais pronunciada for a transição, entre 52% com os estímulos que haviam sido previamente identificados como /l/, e 6% para os restantes estímulos tipo vogal. Como estes estímulos tinham sido previamente ouvidos como /l/ 8% das vezes, torna-se evidente que o grau

de fusão depende dos estímulos envolvidos suscitarem isoladamente o percepto de líquida, ou não.

Assim, as transições de formante são um aspecto da estimulação importante para suscitar fusões. Uma nova experiência destinada a elucidar que transição de formante, ou combinação de transições, é mais pertinente, mostrou claramente que no caso do /l/ é a transição do segundo formante, e no caso do /r/, é a transição do terceiro formante que mais peso têm na determinação das fusões fonológicas. Apresentando dicoticamente uma oclusiva e apenas o segundo formante de /l/, a percentagem de fusões era de 60% (vs 19% com o primeiro formante, e 16% com o terceiro correspondente a /l/); emparelhando a oclusiva com o terceiro formante, a percentagem de fusões era de 64%\*. São pois as transições de formante situadas entre 1000 e 2000 Hz o aspecto da estimulação que desencadeia o fenómeno da fusão fonológica.

Tendo analisado quais os constituintes acústicos relevantes para suscitar respostas de fusão, cabe perguntar até que ponto o fenómeno assenta em interacções auditivo-acústicas propriamente ditas, ou pelo "contrário", envolve estes factores enquanto aspectos do estímulo fala. Embora esta distinção esteja longe de ser simples, é possível abordá-la tentativa e pragmaticamente usando os "chirps", ou as transições de formante isoladas, que são percebidas como ruído e não como fala. Cutting (1975; Experiência 13) comparou a percentagem de fusões em pares de oclusivas com líquidas, e pares de oclusivas com "chirps"; enquanto para os primeiros houve 47% de respostas de fusão, para os segundos aquela percentagem desceu para 8%. Vendo separadamente a percentagem de fusão para cada tipo de "chirp", nota-se que ela é ligeiramente maior nos

\* Lembremos que a distinção entre as líquidas inglesas /l/ e /r/ se faz essencialmente, e se pode fazer exclusivamente em fala sintética, através do terceiro formante.

"chirps" ascendentes (como são as transições de formante da palavra total), principalmente para o "chirp" correspondente ao segundo formante (entre 1.2 e 1.8 KHz), onde atingem 24%. Assim, nem a transição mais apropriada ao fonema líquido "consegue" suscitar um nível de fusões semelhante ao obtido com o estímulo íntegro.

### 3. Explorações teóricas

#### 3.1. A fusão fonológica no contexto de outros fenómenos de fusão de sons

O retrato que acabámos de traçar apresenta-nos a fusão fonológica como um fenómeno relativamente insensível a aspectos não-linguísticos, como a altura tonal, intensidade, timbre e assincronia de apresentação; em contrapartida, características fonéticas, como a presença ou ausência de transições de formante, características fonético-fonémicas, como o grau de fusibilidade dos fonemas ou a semelhança fonológica, e características semânticas, como o estatuto lexical, o contexto e a frequência das palavras, ora favorecem, ora inibem a ocorrência daquele fenómeno. Uma visão sucinta e simplificada é que factores linguísticos e não-linguísticos exercem um efeito diferencial na fusão fonológica. Na medida em que ela é menos afectada por factores não-linguísticos, e sensível aos linguísticos, Cutting (1976) adiantou que os processos envolvidos são de natureza central.

Aquela asserção torna-se mais clara se compararmos a fusão fonológica com outros fenómenos que, de um ou outro modo, combinam características da estimulação acústica: a localização da fonte sonora, as fusões psico-acústicas, espectral e espectro-temporal, e a fusão de traços fonéticos ou *blending*. De facto, o termo fusão é aplicado a qualquer destes fenómenos, e uma comparação sistemática entre eles indubitavelmente esclarece quer a natureza da "fusão", quer as peculiaridades de cada uma

das "fusões". Para ser sistemática (e reveladora), uma tal comparação deverá incidir sobre o efeito que uma determinada variável exerce nos vários tipos de fusão; se esse efeito for diferencial (influindo numas fusões e não noutras, por exemplo), então poder-se-á inferir sobre a diferença nos processos invocados e eventualmente conceber as fusões no seio de uma organização percepto-cognitiva de diferentes níveis.

Usando vários modos de apresentação auditiva, Cutting (1976) comparou a robustez da resposta de fusão perceptiva face a variações de três tipos: temporal - assincronia da apresentação dos estímulos aos dois ouvidos, espectral - diferença entre as frequências fundamentais dos estímulos apresentados a um e ao outro ouvido, e de intensidade - *idem, mutatis mutandi*. Vejamos os resultados desse trabalho; comecemos por caracterizar sucintamente cada uma das fusões consideradas.

Consoante a localização da fonte sonora, o som por ela emitido pode chegar aos nossos ouvidos com ligeiras diferenças temporais e de intensidade. Compare-se, de modo meramente esquemático, a propagação de um som a partir de uma fonte situada à nossa frente com o de outra fonte, situada exactamente à nossa esquerda. Neste segundo caso, o som tem de percorrer uma distância maior para chegar ao ouvido direito e tem de ultrapassar o obstáculo causado pela própria cabeça do auditor: assim, o som chega ao ouvido direito ligeiramente mais tarde, e com uma intensidade relativamente menor (principalmente nas frequências elevadas). Estas diferenças mínimas são "aproveitadas" pelo nosso sistema auditivo, que em vez de nos levar a perceber diferenças no tempo de chegada ou no timbre de dois sons, nos fornece um percepto único cuja origem espacial podemos localizar. *I. e.*, as mensagens acústicas que chegam ao ouvido direito e esquerdo são integradas num percepto único com informação espacial (cf., *e.g.*, Lindsay e Norman, 1971, pp. 178-184).

O termo "fusão psico-acústica" foi usado por Cutting (1976) para dar conta do fenómeno seguinte: perante apresentação dicótica de estímulos cujas características acústicas diferem numa determinada dimensão, é possível o percepto envolver a combinação das características divergentes numa espécie de valor médio (e.g., Halwes, 1969). Concretizando, tomemos o par dicótico /ba-ga/ em fala sintética. Aqueles estímulos diferem apenas quanto à transição do segundo formante, ascendente em /ba/ e descendente em /ga/. Ora um dos erros comuns consiste em ouvir /da/, que acusticamente é como que intermédio entre /ba/ e /ga/: a frequência de início da transição do segundo formante fica mais ou menos a meio entre as propriadas àquelas sílabas "extremas".

As duas fusões seguintes, a espectral e a espectro-temporal, ocorrem quando o sistema auditivo junta o que artificialmente foi fragmentado. Por meios artificiais, é possível separar um determinado sinal, acústico em alguns dos seus constituintes, por exemplo, em altas e baixas frequências. Se estes componentes forem apresentados cada um a um ouvido, normalmente o sujeito ouve como se se tratasse do sinal original (cf. Cutting, 1976); ocorreu portanto uma fusão espectral.

A fusão espectro-temporal está na base da percepção duplex (e.g., Rand, 1974; Mattingly, Liberman, Syrdal e Halwes, 1971): fornecendo a um ouvido a transição de segundo formante excisada do estímulo integral, e ao outro ouvido os formantes estáveis e transições restantes, o sujeito tem, em regra, dois perceptos. Um é resultante da integração (espectro-temporal) da informação acústica apresentada aos dois ouvidos, e consiste na audição do estímulo original, integral; outro é de tipo ruído, correspondente à transição isolada (cf. pp. 69 - 70).

Finalmente, a fusão de traços fonéticos consiste em, a partir de uma estimulação contendo traços fonéticos numa determinada combinação,

ser gerado um percepto em que os mesmos traços fonéticos estão presentes, mas segundo uma combinação diferente. Concretamente, num par dicótico contrastando vozeamento e lugar, em que um elemento é labial sonoro e outro velar surdo, uma resposta de tipo fusão de traços fonéticos ou *blending* consistirá numa velar sonora.

As manipulações temporais, espectrais e de intensidade tiveram efeitos diferenciais nestes seis tipos de fusão, sendo possível destacar certas configurações de resultados (Cutting, 1976).

Temos assim uma primeira configuração, específica da localização sonora. Esta foi a única fusão que revelou uma acentuada sensibilidade à assincronia e à diferença de frequência entre os estímulos (foram usadas neste estudo sílabas sintéticas de tipo CV). Assincronias tão reduzidas como 4 e 5 milisegundos são suficientes para impedir a fusão, enquanto nos outros casos - da fusão espectral à fonológica - são necessárias diferenças de no mínimo 20 a 40 milisegundos para que o percepto de fusão seja abalado. Trata-se de uma importante diferença de sensibilidade a factores temporais entre a localização auditiva e as restantes fusões. Observa-se o mesmo tipo de efeito com a frequência: basta uma diferença de 2 Hz para que desapareça esta fusão, enquanto as restantes continuam a ocorrer com diferenças de 80 Hz.

Além do padrão de resultados ser distinto em termos quantitativos, a localização sonora é também o único tipo de fusão cuja existência se baseia no número de perceptos, e não na própria identidade do percepto. Por estas razões, é claro que os processos envolvidos são qualitativamente diferentes das outras fusões. Na linha de trabalhos de Allport (1968) e Efron (1970) sobre momentos perceptivos, Cutting (*ibidem*) sugere que esta fusão ocorre a um nível relativamente baixo ("lower level"), correspondendo a um primeiro momento perceptivo de 2-5 milisegundos, intervalo durante o qual a

estimulação é integrada. Este processo de integração de baixo nível, característico de um intervalo temporal com aquela magnitude, lembra a distinção feita por Turvey (1973, cit. também por Cutting, *ibidem*) a propósito do tratamento de informação visual, entre processos periféricos e centrais: os primeiros desenrolam-se em intervalos temporais reduzidos e caracterizam-se por elevada sensibilidade a variáveis físicas e por fenómenos de integração, enquanto os segundos abarcam "momentos perceptivos" ou intervalos temporais mais alargadas onde não ocorrem nem aquela sensibilidade, nem aquela integração. Concebendo as operações do sistema perceptivo como organizadas em diferentes níveis, uns mais próximos da estimulação (dos "dados") do que outros\*, a localização auditiva será então um tipo de fusão que se situaria a um nível "baixo" e periférico; no contexto da literatura sobre percepção da fala, este nível corresponderá ao auditivo, por oposição ao nível globalmente referido como linguístico.

Relativamente às restantes fusões, e seguindo a mesma estratégia de encontrar configurações distintivas, a análise do conjunto de resultados obtidos por Cutting permite distinguir de um lado a fusão psico-acústica, a espectral e a espectro-temporal, e do outro os *blendings* e as fusões fonológicas. Esta distinção centra-se no efeito diferencial da assincronia e de uma outra variável, o modo de apresentação\*\*.

\*Retomamos aqui, se bem que num sentido algo diferente, a distinção feita por Norman e Bobrow (1975) entre processos "data-driven" e "conceptually driven".

\*\* Quanto às outras manipulações efectuadas por Cutting (1976), é de referir o que segue. As manipulações da frequência tiveram um efeito moderado nas cinco restantes fusões, e as de intensidade mostraram essencialmente uma dissociação entre a fusão fonológica e os *blendings*, relativamente insensíveis (só diferenças a partir de 30 dB começavam a ser influentes) e a fusão psico-acústica, que era reduzida de modo significativo com "apenas" 10 dB de diferença. Na medida em que estes dados não contradizem os restantes e trazem algum esclarecimento sobretudo quanto às especificidades da fusão psico-acústica relativamente às fusões espectrais e espectro-temporais, que não interessa discutir neste contexto, optamos por não os analisar.

Quanto à assincronia, os *blendings* e as fusões fonológicas são os únicos casos em que há um aumento de perceptos tipo fusão quando existe uma ligeira diferença, nomeadamente de 20 milisegundos, no tempo de chegada dos estímulos (cf. também Poltrock e Hunt, 1977). Nos três casos restantes, a falta de sincronia determinava uma redução de perceptos fundidos; esta redução é significativa com 20 milisegundos de assincronia para a fusão psico-acustica e a espectro-temporal, e com 40 milisegundos para a fusão espectral.

Voltemo-nos agora para a influência do modo de apresentação. Em vez de apresentar um estímulo a um ouvido, e outro estímulo a outro, como é de regra na audição dicótica, é possível "integrar" previamente os estímulos, misturando-os electricamente e gravando-os num mesmo canal, que o sujeito recebe nos dois ouvidos; este é o modo de apresentação binaural. Se a incidência de fusões\* for semelhante para os estímulos apresentados em condição dicótica e em condição binaural, ter-se-á encontrado um argumento em favor da ideia de que os processos que levam ao percepto fusionado são relativamente precoces, e próximos da estimulação. Ora o número de respostas "fusão" só foi semelhante nos dois casos para as fusões psico-acústica, espectral e espectro-temporal. A apresentação binaural reduzia fortemente quer o número de *blendings*, quer o de fusões fonológicas (Cutting, 1976; cf. também Cutting, 1975; Experiência 2). O facto daquelas respostas serem menos frequentes com a apresentação binaural sugere que elas são dependentes de processos centrais de combinação da informação.

O padrão de resultados que acabámos de rever levou Cutting (1976) a propor a existência de, além do nível a que se daria a localização auditiva, pelo menos mais dois níveis de processamento: um em que se enquadrariam

---

\* Entende-se por fusão sob apresentação binaural uma resposta qualitativamente idêntica à fusão sobre apresentação dicótica.

as fusões psico-acústicas, espectrais e espectro-temporais, que se caracterizaria por uma integração de traços acústicos; e outro, em que se dariam os *blendings* e as fusões fonológicas, caracterizado pela análise ("disruption") e recombinação de traços linguísticos. No primeiro destes níveis, a integração da informação dá-se em intervalos que chegam a atingir os 40 milissegundos (ou seja, mesmo com assincronias daquela grandeza, mantém-se o percepto fusionado), o que por si só bastaria para distinguir este nível (Nível II), correspondendo a um segundo "momento perceptivo", do da localização auditiva (Nível I).

Apesar de Cutting ter sugerido, com reservas, que a este nível II fosse aplicado o epíteto de pré-fonético, trabalhos mais recentes sobre fusão espectro-temporal indicam que talvez isso não seja correcto. Aliás, mesmo no conjunto das cinco experiências de Cutting em 1976, há indícios de que a fusão psico-acústica é mais afectada por variáveis como a intensidade, frequência e assincronia dos estímulos. Isto levanta a suposição de que a fusão psico-acústica é relativamente mais periférica ou "lower-level" do que as fusões espectrais e espectro-temporal. Em contrapartida, não se nos colocam dúvidas ao usar o termo "linguístico" para dar conta do nível a que se situam os *blendings* e fusões fonológicas. Com efeito, as experiências revistas mostram que não é só em termos de descrição formal que é necessário recorrer a conceitos de fonética e fonologia: é também a nível de interpretação de resultados experimentais e dos processos que lhe dão origem.

No contexto da comparação com o nível anterior, vale a pena reapreciar o efeito da assincronia dos estímulos. Quer para os *blendings*, quer para as fusões fonológicas, a função que relaciona o grau de assincronia com a incidência de respostas daquele tipo não é linear: há uma descontinuidade cerca dos 20 a 40 milissegundos, que separa uma relação

crescente (entre 0 e 20 ms de intervalo aumentam as fusões) de uma relação decrescente (com aumento da assincronia reduzem-se as respostas de fusão). Ora isto sugere que a fusão ocorre mais facilmente quando o estímulo que chega primeiro já foi parcialmente processado, tendo sido "extraídas" algumas das suas características; este processamento sofre então a interferência provocada pelo segundo estímulo, de que são extraídas outras características... e ocasionalmente as características são erradamente recombinadas. *Blending* e fusão fonológica envolvem pois processos de um nível "superior" e central, entre os quais se contam pelo menos processos fonéticos.

Os processos do terceiro nível proposto por Cutting envolveriam pois, pelo menos, processos fonéticos. É que como os resultados revistos sugerem, e o próprio Cutting (1976) se interroga, talvez os *blendings* e as fusões fonológicas requeiram a distinção entre um terceiro e quarto níveis. A questão é se o mesmo processo básico de análise e recombinação de traços linguísticos está subjacente à formação do percepto, e.g., /ta/ perante /da-ka/, e /dla/ perante /da-la/, ou se pelo contrário este último envolve já a constituição de uma outra representação intermediária. Claro que em termos lógicos é diferente combinar entidades como "lugar de articulação" e "vozeamento", ou fonema /p/ com fonema /l/. A questão aqui, todavia, não é de lógica formal, ou de caracterização do fenómeno, mas sim uma questão psicológica sobre a própria formação do percepto.

Para estabelecer dois níveis diferentes correspondentes um ao *blending* e outro à fusão fonológica, sem contrariar a regra da parcimónia, será importante encontrar dissociações em termos do efeito de uma ou mais variáveis nos dois fenómenos. Tanto quanto sabemos a única comparação possível provém de uma das experiências de Cutting (1976, Experiência 2): apesar do efeito da assincronia dos estímulos ser do mesmo tipo quer para as

fusões fonológicas, quer para os *blendings*, só nas primeiras se verificou uma assimetria: há ligeiramente mais respostas /dra/ quando o /da/ precede o /ra/ do que na ordem de apresentação contrária\*. Ora nos *blendings* não acontece nenhum fenómeno análogo. Outro argumento, se bem que indirecto, em favor da distinção dos processos envolvidos nos *blendings* e nas fusões fonológicas, seria a alegada existência de marcadas diferenças individuais quanto à tendência em ouvir fusões fonológicas, que nunca foi relatada a propósito dos *blendings* (cf. Studdert-Kennedy e Shankweiler, 1970; Cutting, 1976; Repp, 1977; Tartter e Blumstein, 1980). No ponto seguinte apreciaremos esta questão das diferenças individuais.

Por agora, em jeito de conclusão, digamos que é possível enquadrar a fusão fonológica num conjunto de outros fenómenos de combinação de informação acústica e acústico-fonética num percepto único; nesta perspectiva, a fusão fonológica partilha algumas características com a fusão de traços fonéticos, pois ambas parecem situar-se a um nível de processamento em que foi já extraída informação linguística, fonética (no caso dos *blendings*) ou possivelmente fonético-fonémica (no caso das fusões fonológicas).

### 3.2. A questão das diferenças individuais

#### 3.2.1. A distinção entre "high" e "low fusers": primeiros trabalhos

As observações pioneiras de Day (1970-a e-b; 1973) revelaram um facto curioso. A distribuição dos sujeitos consoante o número de fusões fonológicas ouvidas apresentava características peculiares: era bimodal. Em

---

\* Este efeito não foi significativo nesta experiência; foi-o todavia noutras, como as de Cutting (1975; Experiência 3), e de Poltrock e Hunt (1977).

vez de estarem normalmente distribuídos, os sujeitos dividiam-se por dois grupos, um com uma elevada frequência de fusões e outro em que aquelas respostas eram relativamente raras. Os indivíduos pertencentes ao primeiro grupo foram classificados de "high fusers", e os do segundo de "low fusers".

As diferenças entre os "high" e os "low fusers" não se restringiam apenas à tendência em ouvir fusões. Usando os mesmos estímulos numa tarefa diferente, de julgamento de ordem temporal, o padrão de respostas daqueles sujeitos era distinto: mesmo quando a líquida precedia a oclusiva, os "high fusers" relatavam que ouviam primeiro a oclusiva; os "low fusers" desempenhavam a tarefa a um nível semelhante, quer quando a líquida precedia a oclusiva, quer na situação contrária. Os "high fusers" evidenciaram pois, no julgamento de ordem, um viés para a oclusiva (usemos esta expressão para nos referirmos àquele resultado), ausente no "low fusers". Parece então que mesmo numa tarefa não necessariamente linguística, de julgamento de ordem temporal, as respostas dos "high fusers" se conformam a regras de tipo fonológico; por isso, Day (1970-a) denominou estes indivíduos "language bound", contrapondo-os aos "language optional" ou "stimulus bound", que conforme as exigências da tarefa recorrem ou não a um processamento linguístico. Para explicar a relação entre a tendência em perceber fusões fonológicas, por um lado, e o viés para oclusivas, por outro, Day fornece assim uma interpretação: nos indivíduos que frequentemente ouvem fusões, haveria uma dependência perceptiva de regras linguísticas. Esta dependência perceptiva de regras linguísticas permitiria distinguir dois grupos extremos de sujeitos, os dependentes dos independentes (termos nossos). De facto, buscando evidência convergente para esta hipótese, Day (1973-b, referido por Poltrock e Hunt, 1977) verificou que indivíduos com viés para oclusiva tinham mais dificuldade em usar a linguagem de modo não convencional: demoravam mais tempo em trocar os fonemas /l/ por /r/ e vice-versa, num jogo em que o

experimentador pronunciava a palavra que os sujeitos tinham então de transformar; e eram mais lentos numa tarefa de *feedback* auditivo atrasado\*.

Resumindo, as primeiras observações sobre a distribuição dos sujeitos quanto à incidência de fusões fonológicas revelaram, facto pouco comum, dois grupos extremos de "high" e "low fusers". A interpretação proposta apontava para a dependência, ou não, do processo perceptivo de regras linguísticas.

As implicações teóricas e mesmo práticas da existência de uma dicotomia entre "low" e "high fusers" são importantes. Será a fusão fonológica um processo central, fonte de marcadas diferenças individuais? Como se relacionará com outras competências de linguagem falada e escrita? Importa pois averiguar até que ponto esta dicotomia é robusta. Na sequência da hipótese de Day foram efectuados alguns estudos que procuraram (i) replicar a observação de uma distribuição bimodal; (ii) explorar a relação entre diferenças individuais quanto à incidência de fusão fonológica e outras variáveis.

### 3.2.2. Variabilidade inter-individual

Cutting (1975) e Cutting e Day (1975), no conjunto das experiências efectuadas, obtiveram distribuições dos dados individuais em geral consistentes com a distinção entre "high" e "low fusers". Num total de 68 sujeitos, escrevem Cutting e Day (*ibidem*, p. 112) que se distinguem dois grupos, um com a percentagem de fusões em torno de 39% (entre 2 e 60%), e outro com a percentagem de fusões em torno de 88% (entre 70 e 100%). No

\* "Delayed auditory feedback", ou DAF, em que por meios artificiais se introduz um atraso entre a própria pronúncia e o *feedback* auditivo.

estudo de Cutting (*ibidem*), a distribuição dos 140 sujeitos vai também no sentido de uma bimodalidade, em que a maioria tem uma percentagem de fusões entre 41 a 60% (54 sujeitos) e 21 a 40% (39 sujeitos), seguindo-se uma proporção razoável (26 sujeitos) com cerca de 81 a 100% de fusões; há menos indivíduos cuja percentagem de fusões se situa na faixa intermédia dos 61-80% (14 sujeitos). Aqui, a bimodalidade não é tão flagrante, o que se poderia eventualmente atribuir ao facto dos sujeitos terem participado em diferentes experiências, com importantes variações no material e número de ensaios.

O primeiro trabalho a abordar sistematicamente a questão das diferenças individuais na incidência de fusões fonológicas foi efectuado por Poltrock e Hunt (1977), que procuraram estabelecer relações entre a variabilidade das respostas de fusão e de outras respostas, nomeadamente de julgamento de ordem temporal. Numa primeira experiência, usando como estímulos não-palavras cuja combinação resultava em palavras, aqueles autores observaram 60 sujeitos; a percentagem média de fusão rondou os 70%, e a distribuição dos sujeitos por frequência de fusão era claramente unimodal em torno daquele valor.

Uma das razões da discrepância entre estes resultados e os observados por Day poder-se-ia encontrar no seguinte: segundo aquela autora\*, para alguns sujeitos a percentagem de fusão decresce ao longo dos primeiros 60 ensaios, e a bimodalidade só aparece após cerca de 180 ensaios. Isto implica que a bimodalidade da distribuição é mais nítida após um elevado número de ensaios. Reanalizando os seus dados nesta perspectiva, Poltrock e Hunt confirmaram um decréscimo significativo do número de fusões entre os primeiros e os últimos 20 ensaios (num total de 60), e um aumento significativo da respectiva variância; porém, a distribuição dos

---

\* Referido por Poltrock e Hunt (*ibidem*, p. 65), reportando-se a um *memorandum* não publicado "Distribution of fusion scores as a function of number of trials" de R. S. Day.

sujeitos por frequência de fusão nos últimos 41 a 60 ensaios mantinha-se unimodal, se bem que mais dispersa.

Uma segunda experiência, usando material e procedimento diferentes, e um maior número de sujeitos (106), replicou estes resultados: não se vislumbrava a distinção entre "high" e "low fusers", a distribuição dos sujeitos quanto à incidência de fusão mantinha-se unimodal. Isto verificava-se tanto com a medida "habitual" de identificação, em que o sujeito escreve aquilo que ouve, como com uma medida de reconhecimento, em que o sujeito dispõe de uma folha com três possibilidades de resposta - estímulo "a", "b", fusão - em que assinala o seu percepto.

O carácter unimodal da distribuição dos sujeitos por frequência de fusão foi de novo observado em amostras de 36 e 49 sujeitos, por Keele e Lyon (1977) e Sexton e Geffen (1981). Ambos os trabalhos usaram metodologias diferentes, como se discutirá mais à frente, o que contribui para salientar a robustez deste achado empírico: a distribuição dos sujeitos por frequência de fusão não é bimodal. Mesmo tendo em consideração um elevado número de ensaios, como fizeram Sexton e Geffen (1981), mantem-se a unimodalidade. Aqueles autores retestaram 25 sujeitos da amostra inicial, numa situação com 400 ensaios de fusão, com o objectivo de apreciar melhor as diferenças individuais; além da distribuição unimodal, outro resultado digno de nota foi a elevada fidelidade teste-reteste quanto à frequência de fusão dos sujeitos, apesar do intervalo de três meses entre a primeira e a segunda passagem. Isto mostra claramente a estabilidade das diferenças individuais quanto à frequência de fusão; e que esta reflecte características individuais estáveis quanto aos processos de tratamento da informação envolvidos (cf. Sexton e Geffen, *ibidem*).

Que fazer então das primeiras observações sobre "high" e "low fusers"? Segundo Poltrock e Hunt (1977), a explicação mais plausível será o

reduzido número de sujeitos inicialmente observado por Day: em 16 indivíduos não é difícil acontecer que se encontrem distribuições peculiares, particularmente quando se observam fenómenos susceptíveis de grande variação individual. E, considerando o conjunto das observações feitas, uma constante é sem dúvida a variabilidade intersujeitos: uns indivíduos fundem mais, outros menos, e, contrariamente ao sugerido inicialmente por Day, há outros que fundem "mais ou menos". A questão relevante deixa de ser a de identificar os sujeitos extremos (confirmar/infirmar a bimodalidade), para se concentrar na exploração dos mecanismos subjacentes às diferenças individuais.

### 3.2.3. Factores explicativos da variabilidade inter-individual

Uma das razões invocadas para dar conta das diferenças individuais foi, como vimos, a maior ou menor dependência do processo perceptivo face a regras linguísticas. Esta hipótese assenta na existência de uma correlação entre fusões e viés para a oclusiva numa tarefa de julgamento de ordem temporal. Esta correlação, observada inicialmente numa amostra relativamente pequena e composta por sujeitos que apresentavam frequências extremas de fusão (Day, 1968; 1970-b), mereceu o exame detalhado por parte de alguns investigadores, e foi replicada com amostras mais numerosas e diversificadas.

Pollock e Hunt (1977; Experiência 1) obtiveram correlações de .42 ( $r$  de Pearson) entre a frequência de fusões e a exactidão no julgamento de ordem temporal se a oclusiva precedesse a líquida, e de -.45 se fosse a líquida a preceder a oclusiva. Aqueles autores calcularam ainda uma medida de viés para a oclusiva em que fosse minimizado o papel da sensibilidade à

ordem temporal\* e de novo foi replicada a correlação positiva entre o viés e a frequência de fusões ( $r=.48$ ). É importante acrescentar que não se observou relação significativa entre a frequência de fusões e a exactidão propriamente dita (considerando todos os ensaios, em que a líquida precedia a oclusiva, ou vice-versa). Estes resultados foram replicados pelos mesmos autores numa segunda experiência com certos afinamentos metodológicos: para eliminar, ou pelo menos minimizar, um eventual viés em relatar palavras, os estímulos foram palavras em vez de não-palavras como na primeira experiência, e os sujeitos foram previamente familiarizados com eles. Não se trata portanto de que sujeitos menos discriminativos quanto à ordem temporal sejam aqueles que ouvem fusões.

Esta conclusão é corroborada por uma série de resultados obtidos em tarefas semelhantes de julgamento de ordem temporal, examinadas por Keele e Lyon (1982). Sucintamente, estes autores encontraram correlações positivas elevadas e significativas em várias tarefas de julgamento de ordem temporal, inclusivé com estímulos não-verbais, como tons (julgar qual foi o primeiro estímulo ouvido, ou qual o primeiro fonema apresentado ao ouvido direito, ou que ouvido foi estimulado em primeiro lugar, se o direito, se o esquerdo), mas não entre a exactidão nalgumas daquelas tarefas e uma medida de fusão.

Uma sugestão quanto ao mecanismo responsável pela associação entre frequência de fusões e viés do julgamento de ordem temporal para as oclusivas foi fornecida por Poltrock e Hunt (*ibidem*). Trata-se de um modelo de confusão dicótica. Quando se apresentam diferentes estímulos aos dois

---

\* O baixo desempenho ao julgar que estímulo surgiu primeiro nos ensaios em que a líquida precede a oclusiva poderia dever-se a menor capacidade de discriminação de ordem temporal, *tout court*. Por isso, é conveniente ter em conta o desempenho global, independentemente da consoante que aparece primeiro. A medida adoptada por Poltrock e Hunt (*ibidem*) foi a diferença entre a percentagem de respostas correctas quando a oclusiva precede e a percentagem de respostas correctas quando a líquida precede.

ouvidos, é possível que seja extraída informação sobre os fonemas apresentados, mas que se perca, total ou parcialmente, informação sobre a respectiva origem, se à direita, se à esquerda; um processo deste tipo daria origem a erros quer de fusão (confundindo-se a localização, por exemplo, do fonema /l/ com a do fonema /p/), quer de separação. Se a um ouvido for apresentado "black" e a outro "lack", há a possibilidade de confundir a origem do fonema /l/ e responder com "back", uma resposta de separação. Se os erros de fusão se devessem a um mecanismo deste tipo, então os sujeitos que fundem mais seriam também os que separam mais.

Claro que subjacente a este modelo está a ideia de que os julgamentos de ordem temporal revelam (também?) a maior ou menor capacidade de separação espacial: quando a tarefa do sujeito é dizer que fonema ouviu primeiro, entre dois fonemas que têm diferentes origens espaciais, a sua exactidão dependerá não só da discriminação temporal propriamente dita, como também da localização espacial.

Trabalhos empíricos efectuados para testar este modelo, dito de confusão dicótica, foram claros em refutá-lo. Usando procedimentos apropriados à observação de respostas de separação, em que metade dos ensaios fornecem oportunidade de fusão, e outra metade oportunidade de separação (apresentando o estímulo total, *e.g.*, "great", e o parcial, *e.g.*, "gate"), não se observaram resultados consistentes com aquela hipótese: além da correlação entre frequência de respostas de fusão e de separação ser pequena e nalguns casos negativa; o efeito de outras variáveis, como o tipo de tarefa, identificação propriamente dita *vs* reconhecimento, não obedecia ao mesmo padrão para aqueles dois tipos de resposta (Pollock e Hunt, 1977; *cf.* também Keele e Lyon, 1982).

Podemos pois concluir que não é correcta uma interpretação das fusões em termos de dificuldade em localizar a fonte dos estímulos. Parece-

-nos que as diferenças individuais observadas quer a nível de frequência de fusões, quer a nível de julgamento de ordem temporal, tal como ele tem sido entendido neste contexto, se deverão atribuir ao recurso diferencial a "regras" linguísticas, que estabelecem uma predisposição\* em favor de perceptos fonologicamente aceitáveis. A conclusão de Poltrock e Hunt (1977) é semelhante à nossa, na medida em que aceita o papel de regras linguísticas na determinação da frequência de fusão, e lhes atribui também a origem das diferenças individuais.

Não deixa de ser desconcertante a conclusão contrária, extraída por Keele e Lyon (1982, p. 442), praticamente com base nos mesmos resultados experimentais: "the results of our experiments were, on the whole, rather discouraging for the hypothesis that individual differences in fusion are caused by differences in the influence of linguistic rules on perception (...) a prototypical 'fuser' would be an individual who tends to have difficulty discriminating the temporal order of the word components and, also, has a bias to decide that poorly resolved wordlike stimuli were probably words". Esta conclusão obriga-nos a ver mais de perto as experiências realizadas por aqueles autores, e seus resultados.

O trabalho em questão pretende testar uma versão forte da hipótese da dependência de regras linguísticas (a hipótese da "language boundedness", como lhe chamam Keele e Lyon) como determinante da fusão fonológica. Haverá formas diferentes de conceber o que é realmente uma fusão; a mais extrema será que, quando um indivíduo ouve uma fusão, ele não tem acesso à informação relativa a cada uma das palavras componentes. Ou seja, a fusão impede que o sujeito seja capaz de emitir julgamentos sobre aspectos perceptivos específicos dos estímulos não fundidos. Esta ideia é sem dúvida plausível; é consistente com o que temos

---

\* No sentido de "priming".

vindo a examinar, e que coloca o fenómeno da fusão mais a nível perceptivo do que a nível de viés de resposta. Segundo Keele e Lyon (*ibidem*), foi justamente esta concepção de fusão que permitiu a R. Day estabelecer a relação teórica entre os julgamentos de ordem temporal (que implicam separar a informação relativa a um ou outro componente) e a fusão fonológica.

Para avaliar a validade desta hipótese, Keele e Lyon (*ibidem*) elaboram uma série de tarefas com estímulos apresentados dicoticamente, e que envolvem o acesso à informação espácio-temporal característica deste ou daquele componente. Nas duas experiências que apresentam em pormenor, apenas uma tarefa permite extrair uma medida (indirecta) de fusão: há ensaios de dois tipos, um em que se apresentam pares de não-palavras fundíveis, como "dagon + ragon", e outro em que se apresenta aos dois ouvidos a palavra correspondente, como "dragon"; o sujeito deve responder se o que ouve é uma palavra ou não. Ora se o sujeito responde "palavra" quando só os componentes foram apresentados, deve ter ocorrido uma fusão. Calculou-se a correlação entre os erros nesta tarefa de julgamento de palavras (indicadores de fusão) e erros em duas tarefas de julgamento, ditas de ordem temporal e de ouvido. Na primeira, o sujeito devia assinalar, numa folha de resposta contendo as alternativas possíveis, qual a não-palavra que tinha ouvido primeiro; na segunda, devia também assinalar na folha de resposta qual o componente que tinha sido apresentado no ouvido direito. Ora, se as fusões fonológicas se devessem a uma incapacidade global em discriminar a ordem temporal dos componentes, esperar-se-ia, no mínimo e pondo de lado o eventual papel de factores atencionais, que o desempenho nas três tarefas estivesse intercorrelacionado. Não foi isso que aconteceu: as correlações entre os erros na tarefa de julgamento de palavra/não-palavra e qualquer das outras duas rondavam os .23 ( $r$  de Pearson) e não eram significativas. Os autores imaginaram então uma

outra tarefa de julgamento de ordem temporal (que ouvido foi estimulado primeiro), cujos resultados se correlacionaram com o desempenho numa tarefa de decisão lexical análoga à anterior\*. Replicou-se a inexistência de correlação. Outro aspecto dos resultados é que se observaram sempre correlações relativamente elevadas (entre .38 e .60) e significativas entre os vários julgamentos de ordem temporal\*\*.

Em geral, os resultados mostram que os sujeitos conseguem efectivamente emitir juízos sobre os pares de componentes. Este facto leva os autores a rejeitarem a ideia segundo a qual haveria indivíduos tão dependentes de regras linguísticas (ou "language-bound") que lhes estaria definitivamente vedado o acesso aos componentes de uma palavra. Procuram então explorar que mecanismos de processamento da informação estariam subjacentes à fusão e não conseguem encontrar relações entre várias tarefas envolvendo estímulos verbais\*\*\* e a de julgamento de ordem temporal. Concluem que a fusão não parece estar dependente de factores linguísticos e numa derradeira experiência exploram a relação entre julgamentos de ordem de fonema e de tons. Como de facto encontram uma correlação de .62 entre o desempenho naquelas duas tarefas, acham apoiada

---

\* Em metade dos ensaios só aparecia a palavra num ouvido, e no outro a não-palavra correspondente. A outra metade dos ensaios era constituída por pares de não-palavras fundíveis.

\*\* A existência de intercorrelações entre as várias tarefas de julgamento de ordem temporal leva os autores a afirmar que o baixo desempenho em avaliar que fonema é apresentado em primeiro lugar é determinado em larga medida pela baixa discriminabilidade de ordem temporal e não pela fusão (*ibidem*, p. 438). Ora a fusão só "determinaria" o viés para oclusivas e não a exactidão, *grosso modo*, daqueles julgamentos (cf. Poltrock e Hunt, 1977).

\*\*\* Trata-se de testar a hipótese segundo a qual os indivíduos que têm dificuldade no julgamento de ordem temporal dos componentes (*ibidem*, p.439), o que aparentemente os autores assimilam com 'responder fusão', são aqueles que têm dificuldades em lidar com códigos diferentes, o fonémico e o de palavra; eles não conseguiriam manter separados estes códigos semelhantes. Entre as tarefas usadas para operacionalizar estas ideias, contam-se as de julgamento igual/diferente de letras com o mesmo nome, mas forma física diferente (Aa vs AT ou AA).

a sua ideia de que os julgamentos de ordem temporal de fonemas reflectem a "temporal order ability regardless of language content" (*ibidem*, p. 441), e proclamam a conclusão de que as diferenças individuais na frequência de fusão não se devem à dependência de regras linguísticas, mas sim à capacidade de fazer discriminações temporais.

A razão fundamental por que rejeitamos esta conclusão é que ela assenta no equívoco de praticamente assimilar o julgamento de ordem temporal de fonema com o fenómeno de fusão. Isto torna-se particularmente evidente quando os autores apresentam as suas investigações teóricas, dizendo que elas visam explicar por que há indivíduos com dificuldades no julgamento de ordem temporal dos fonemas, e interpretam os resultados em termos da dependência da fusão de regras linguísticas. Ora a associação que tem vindo a ser experimentalmente confirmada é entre frequência de fusão e viés para a oclusiva em julgamentos de ordem temporal de fonema. As análises de Keele e Lyon, porém, nunca tomaram em consideração a possível assimetria de desempenho para quando é a oclusiva a preceder a líquida ou vice-versa. O facto de medidas de *performance* em várias tarefas de julgamento de ordem temporal (de fonemas, ouvido, ou tom) estarem correlacionadas só nos mostra que há em todas elas possivelmente um factor comum, não necessariamente linguístico, claro; não permite concluir que o desempenho numa outra tarefa, que se supõe estar relacionada - apesar da evidência experimental o desmentir -, se deverá atribuir a esse hipotético factor comum, não linguístico.

Convém referir a este propósito que outros autores (Pollock e Hunt, 1977) exploraram uma situação análoga de julgamento de ordem temporal na modalidade visual, apresentando asteriscos nos campos visuais esquerdo e direito os sujeitos deveriam relatar. A percentagem de respostas correctas nesta situação estava de facto significativamente correlacionada

com o desempenho na sua análoga auditiva, uma correlação positiva com a exactidão no julgamento da ordem de fonemas ( $r = .40$ ), e negativa com o viés para as oclusivas ( $r = -.26$ ). Não houve porém nenhuma correlação com a frequência de fusões ( $r = -.08$ ). Estes resultados indicam que alguma variância do viés para oclusivas deverá depender da sensibilidade à ordem temporal e ser independente da tendência em ouvir fusões. Relembremos que a correlação entre o viés para oclusivas e a frequência de fusão era de .55; assim, o hipotético factor comum dará conta de apenas cerca de 30% da variância. Como notam Poltrock e Hunt (*ibidem*), mais de metade da variância do viés para oclusivas fica por explicar, mesmo considerando o factor comum com as fusões ( $\approx 30\%$ ) e a sensibilidade à ordem temporal ( $\approx 7\%$ ).

Um outro aspecto merece discussão. Não sendo a fusão fonológica uma mera escolha consciente de resposta tipo palavra, é plausível que o sujeito que a ouve não consiga relatar características específicas de cada componente, como supunham Keele e Lyon (*ibidem*). Estes autores pretenderam testar aquela hipótese avaliando o desempenho em tarefas que à partida eram apresentadas ao sujeito como de discriminação de ordem temporal. Ora quando o sujeito consegue relatar que fonema ouviu primeiro, no contexto de uma tarefa em que isso lhe é pedido, poder-se-á transpor este conhecimento para uma situação em que ele apenas repete o que vai ouvindo, ou em que decide se o que ouviu é ou não uma palavra? Só se supusermos um isomorfismo entre as duas situações: a ambas estariam subjacentes um ou vários processos independentes de estratégias atencionais. A questão é se a fusão fonológica resultará de um processo puramente automático e imune à influência de estratégias de processamento. Se sim, então mesmo no desempenho de tarefas com diferentes exigências se deveriam vislumbrar os efeitos daquele processo.

Parece-nos que as análises feitas por Keele e Lyon assentam implicitamente no pressuposto de que a fusão fonológica é pouco sensível a variáveis atencionais, e rigidamente determinada (esta suposição é consistente com a ideia de o atribuir a maior ou menor capacidade de discriminação de ordem temporal). Ora isso é justamente uma questão teórica importante, que só pode ser resolvida empiricamente. Em vez de pressupor que a fusão fonológica é um fenómeno rigidamente determinado, podemos interrogar-nos se ela será sensível a factores como a estratégia atencional. Se assim fosse, as diferenças individuais encontradas poderiam depender do modo como o sujeito atende à situação, tanto ou mais do que de diferenças de mecanismos de processamento. Para clarificar estas questões, pelo menos duas vias são atraentes: manipular explicitamente a atenção do sujeito sem mudar a natureza da tarefa, e fazer uma análise das respostas de fusão usando a teoria da detecção do sinal. Esta última sugestão tinha sido formulada por Poltrock e Hunt (1977) e foi levada à prática por Sexton e Geffen (1981), num estudo que passamos a examinar.

Suponhamos uma tarefa de detecção de alvo, em que o sujeito tem de, o mais rapidamente possível, pressionar uma tecla quando ouve uma palavra previamente estabelecida, *e.g.*, "black". Os estímulos são-lhe apresentados através de auscultadores, de tal modo que nalguns ensaios aparece a palavra-alvo num dos ouvidos e no outro ouvido uma palavra diferente, por exemplo "pen", ou similar, por exemplo, "lack"; noutros ensaios apresentam-se pares fundíveis, por exemplo "back + lack". Numa situação deste tipo, é possível distinguir, por um lado, as respostas correctas propriamente ditas em que o indivíduo detecta o alvo que efectivamente foi apresentado, e por outro as fusões e os outros erros, em que o indivíduo julga ouvir o alvo quando apenas foram apresentados os respectivos componentes ou palavras semelhantes; em termos da teoria de detecção do sinal, concebemos aquelas respostas respectivamente como êxitos e falsos alarmes.

A relação entre êxitos e falsos alarmes permite, mediante certos pressupostos (cf., e.g., Tanner e Swets, 1954; Swets, 1964), extrair uma medida de sensibilidade,  $d'$ , correspondente à capacidade discriminativa entre os dois acontecimentos, e outra de critério, independente da primeira, correspondendo a um viés de resposta.

Sexton e Geffen (*ibidem*) manipularam a estratégia atencional dos sujeitos, pedindo a um grupo que efectuasse a tarefa atendendo só ao ouvido direito, e ignorasse tudo o que aparecesse no ouvido esquerdo; ao outro grupo eram dadas instruções para que prestasse atenção igualmente aos dois ouvidos\*. O primeiro é o grupo de atenção focada e o segundo de atenção dividida. Ao comparar-se a incidência de fusões e outras medidas nestes dois grupos, pode avaliar-se se há alguma influência do esforço voluntário em orientar selectivamente a atenção. A percentagem de fusões foi de 37 no grupo de atenção dividida, e de 23 no grupo de atenção focada. Este decréscimo na incidência de fusões não chegou a atingir a significância, talvez porque, como notam os autores, houve grande variabilidade interindividual. Nos dois casos, a percentagem de fusões foi significativamente superior à de outros erros (responder 'sim' a estímulos distractivos semelhantes ao alvo - por exemplo, a "crate" ou "drain" se o alvo fosse "crane"), que rondava os 15%. Isto indica que a fusão não provém apenas de um viés em responder 'sim' perante estímulos parecidos com o alvo, ou de uma fraca sensibilidade discriminativa perante fonemas semelhantes: se assim fosse a sua frequência deveria ser próxima daqueles erros. Outro aspecto digno de realce é que nas duas condições de atenção houve significativamente menos fusões do que respostas correctas (a percentagem de respostas correctas foi de 64 com atenção dividida e 77 com

\* Todos os estímulos eram palavras, tanto os alvos ("crane", "prod", "black", "great", "plane", "track", "brace", "dread", "glow" e "clap") como os componentes ("cane + rane", etc), como os estímulos distractivos semelhantes ao alvo ("drain", "crate", "plod", "prop", etc).

atenção focada), o que indica que não há equivalência perceptiva entre uma situação em que o alvo realmente está presente e outra em que apenas existem os componentes respectivos.

A mera existência de fusões fonológicas na situação de atenção focada mostra, obviamente, que a informação do canal não-atendido não foi eficaz e totalmente bloqueada; o que é mais revelador é que naquela condição não houve outras respostas devidas à interferência da informação do canal rejeitado (como detecção do alvo apresentado no ouvido a que não se presta atenção). A influência da informação do canal não-atendido só acontece quando ela é fonologicamente compatível com a do canal atendido. A fusão fonológica parece pois dever-se a um processo pelo menos parcialmente involuntário.

Vejamos se uma análise em termos de detecção do sinal permite esclarecer alguns destes pontos. É possível calcular a discriminabilidade entre alvos e pares fundíveis (usando a percentagem de detecções correctas e a de fusões),  $d'_1$ , e a discriminabilidade entre alvos e estímulos distractivos (comparando detecções correctas e outros erros),  $d'_2$ . Os resultados revelaram que na situação de atenção dividida há uma diferença significativa entre aqueles dois  $d'$ , mais um argumento em favor da ideia que não há equivalência perceptiva nas duas situações (ouvir fusão ou confundir erroneamente um estímulo semelhante com o alvo). Na condição de atenção focada, verifica-se um aumento significativo na discriminabilidade dos alvos relativamente aos pares fundíveis ( $d'_1$ ) e relativamente aos distractores ( $d'_2$ )\*. Como notam Sexton e Geffen, a estratégia de atenção foi suficiente para aumentar a discriminabilidade

---

\* Os valores de  $d'_1$  e  $d'_2$  são respectivamente .67 e 1.38 (atenção dividida). Na condição de atenção focada os dados apontam no mesmo sentido,  $d'_1$  de 1.56 e  $d'_2$  de 2.01, mas os autores não referem nenhum teste estatístico à significância daquela diferença (Sexton e Geffen, *ibidem*, p. 427). Quanto ao viés de resposta, os autores só referem o calculado relativamente aos estímulos distractivos, que não sofreu nenhum efeito da estratégia atencional.

entre alvos e pares fundíveis em atenção focada a um nível ligeiramente superior ao da discriminabilidade entre alvos e estímulos distractivos em atenção dividida (respectivamente, 1.56 e 1.38).

Em suma, a análise em termos de detecção do sinal mostra que a estratégia de atenção selectiva leva a que se discrimine melhor entre uma palavra com encontro consonantal apresentada num ouvido e o respectivo par de palavras componentes, apresentadas uma a um ouvido e outra a outro. Este aumento de discriminabilidade traduz-se na menor percentagem de fusões quando a atenção está dirigida a um só ouvido, indicando que os processos responsáveis pelo percepto fusão são susceptíveis de algum controlo atencional. Por outro lado, mesmo na situação de atenção selectiva ocorrem mais fusões do que "confusões" (resposta positiva a estímulo semelhante ao alvo), o que sugere uma componente involuntária nos processos que levam à fusão.

A propósito do carácter parcialmente involuntário dos processos responsáveis pela fusão fonológica, vale a pena mencionar os dados relativos à latência de resposta, que foram também analisados por Sexton e Geffen (*ibidem*). Sendo conhecida a existência de uma vantagem de ouvido direito em tarefas de audição dicótica, pode prever-se que a detecção correcta do alvo seja mais rápida quando ele aparece no ouvido direito do que no esquerdo. Como a fusão envolve informação dos dois ouvidos, ela necessariamente terá uma latência de resposta maior do que a "simples" detecção de alvo no ouvido direito. Por outro lado, se a fusão fonológica fosse devida a uma escolha mais ou menos consciente entre uma resposta aceitável linguisticamente e um percepto ambíguo, poder-se-ia esperar que a sua latência de resposta fosse maior do que a de uma detecção correcta do alvo. Ora o que se verificou foi que os tempos de reacção médios foram de facto significativamente mais baixos para a detecção correcta do alvo no ouvido

direito do que a detecção correcta no ouvido esquerdo ou a fusão. Entre estas duas respostas não havia diferença significativa.

O facto de se demorar tanto tempo a detectar um alvo apresentado no ouvido esquerdo como a perceber uma fusão é consistente com a proposta de Cutting (1976) de que haveria um processamento paralelo dos dois elementos do par fundível (referido também por Sexton e Geffen, *ibidem*). Se isto for verdade, o efeito da atenção na fusão fonológica não ocorrerá numa fase relativamente precoce do tratamento da informação, mas sim numa fase mais tardia. Algumas observações fenomenológicas relatadas por Sexton e Geffen (*ibidem*) sugerem que, pelo menos para alguns indivíduos, poderá ser isso que acontece. Nalguns ensaios, certos sujeitos (e também o primeiro autor daquele trabalho) vivenciavam dois perceptos perante a apresentação do par fundível: um primeiro percepto, imediato e claro, de que o alvo tinha aparecido (uma fusão, portanto), logo seguido pela impressão de que tinham aparecido de facto as duas palavras componentes.

#### 4. Conclusão

Após termos descrito as características da fusão fonológica e revisto os principais resultados experimentais e interpretações teóricas propostas, podemos concluir que a fusão fonológica é um fenómeno que envolve a combinação de informação fonologicamente compatível, apresentada separadamente aos dois ouvidos, num percepto único; essa combinação é mais afectada por factores de tipo linguístico do que por factores fisico-acústicos (Cutting, 1975, 1976; Cutting e Day, 1975). Não se trata simplesmente da "confusão" de mensagens, pois não há equivalência perceptiva entre fusão e outras respostas interpretáveis à luz da similaridade dos estímulos (Pollock e Hunt, 1977; Sexton e Geffen, 1981).

Envolve antes a integração da informação proveniente dos dois canais segundo regras fonológicas, que provavelmente se dá a nível central após processamento, paralelo e em certo grau automático, da informação dos dois canais. O facto de a fusão fonológica ser moderadamente afectada pela orientação selectiva da atenção leva a crer que não se trata de um fenómeno puramente "ilusório" (de tipo ilusão), mas passível de um certo grau de controlo estratégico.

## PARTE II - CONTRIBUIÇÃO EXPERIMENTAL PARA O ESTUDO DOS EFEITOS DA ALFABETIZAÇÃO NA PERCEPÇÃO DA FALA

### CAPÍTULO IV

#### LATERALIZAÇÃO DAS FUNÇÕES LINGUÍSTICAS

No presente capítulo damos início à nossa contribuição experimental para o esclarecimento do papel da alfabetização em certos aspectos do processamento da linguagem falada. O problema que examinaremos, o da relação entre alfabetização e lateralização cerebral para a linguagem, foi directamente motivado pela controvérsia existente na literatura a este propósito. Assim, o objectivo principal das duas experiências que realizámos foi o de tentar resolver as discrepâncias apontadas em estudos anteriores sobre o mesmo problema. Porém, estas experiências são também úteis para a compreensão em geral dos processos envolvidos na percepção da fala, nomeadamente quanto ao grau de rigidez do seu alicerce biológico.

Começaremos por apresentar os dados e sugestões disponíveis a propósito da influência de variáveis experienciais na lateralização cerebral, em animais e em humanos. Discutiremos mais detalhadamente o papel da aprendizagem da leitura e da escrita, principalmente através do exame aos dados provenientes da neuropsicologia clínica sobre o efeito de lesões cerebrais unilaterais em adultos letrados e iletrados.

Consideraremos então os resultados relevantes obtidos com sujeitos normais em tarefas de audição dicótica, que inspiraram a realização das Experiências I e II. O restante deste capítulo é-lhes devotado.

Quer alguma literatura que discutiremos, quer as nossas duas experiências, recorrem à técnica de audição dicótica. Ela consiste em apresentar simultaneamente uma mensagem a um ouvido e uma mensagem diferente ao outro ouvido. A sua utilização no contexto da investigação em lateralidade cerebral foi já por nós discutida em outro lugar (Castro, 1983; pp. 37-47), pelo que não nos deteremos aqui sobre este aspecto. Claro que uma vantagem de ouvido não reflectirá simples e linearmente a superioridade do hemisfério contralateral (para um exame crítico veja-se, *e.g.*, Morais, 1981; Teng, 1981; cf. também Lauter, 1982). Porém, em condições metodológicas devidamente cuidadas, principalmente quanto ao papel da atenção e do material apresentado a um ou outro ouvido, a interpretação mais plausível e, de facto, aceite, é que uma vantagem do ouvido direito obtida com material verbal traduz uma superioridade do hemisfério esquerdo para o processamento da linguagem (cf. *e.g.*, Porter e Hughes, 1983).

### 1. Introdução Geral

É hoje um facto solidamente apoiado em provas clínicas e experimentais que o hemisfério cerebral esquerdo desempenha um papel preponderante no processamento da linguagem (cf., entre outros, Beaumont, 1983; Bradshaw e Nettleton, 1981, 1983; Gazzaniga e LeDoux, 1978; Goodglass e Geschwind, 1976; Hellige, 1983; Lenneberg, 1967; Luria, 1970; Poeck, 1982). O ponto de vista geralmente aceite e bem documentado empiricamente é que aquela lateralização cerebral existe muito precocemente (*e.g.*, Best, Hoffman e Glanville, 1982; Lewrowicz e Turkewitz 1982; Molfese e Molfese, 1979; Segalowitz e Chapman, 1980; Trevarthen, 1983; Witelson e Pallie, 1973), estando já estabelecida antes da idade escolar e do início da aprendizagem da leitura e da escrita (*e.g.*, Eling,

Marshall e van Galen, 1981; Kinsbourne, 1975; Kinsbourne e Hiscock, 1977). De facto, tem-se mesmo procurado esclarecer até que ponto a superioridade do hemisfério esquerdo para a linguagem é condição de sucesso na aprendizagem da leitura (para revisões recentes, veja-se Obrzut, Hynd e Boliek-Uphoff, no prelo; Pirozzolo, Rayner e Hynd, 1983).

Por outro lado, vários autores desde Weber (1904) têm defendido a ideia de que a própria experiência com a linguagem escrita pode, em certa medida, favorecer o estabelecimento da superioridade verbal do hemisfério esquerdo. Mais à frente faremos a revisão dos trabalhos relevantes quanto à elucidação desta hipótese. Porém, na medida em que ela levanta a questão do papel de variáveis experienciais na determinação - pelo menos parcial - da especialização hemisférica, vejamos rapidamente o que tem sido discutido na literatura a esse propósito.

### 1.1. Variáveis experienciais e lateralização

Um domínio onde tem sido investigado o efeito da experiência na lateralidade cerebral - no sentido genérico de uma especialização funcional, independentemente da sua natureza - é o da organização cerebral de alguns animais como as aves canoras, os roedores e os primatas não-humanos (LeMay e Geschwind, 1975; Notebohm, 1970; Rogers e Anson, 1979; Sherman Garbanati, Rosen, Yutzey e Denenberg, 1980, entre outros; para uma revisão ver Denenberg, 1981). A experiência precoce, como o manuseamento\* e um meio enriquecido, pode induzir a lateralização num cérebro à partida simétrico ou favorecer o seu desenvolvimento num cérebro já algo lateralizado; isto foi demonstrado com ratos (Denenberg, 1981). O exame das

---

\* Manuseamento no sentido de "handling". Na experiência de Denenberg (*ibidem*), manuseamento ("handling") desde o nascimento até 20 dias de idade.

investigações neste domínio leva Denenberg a concluir que "one major function of early experience is to provide stimulation during development, which acts to enhance the growth and the development of the corpus callosum, thereby giving rise to a more specialized brain" (*ibidem*, p. 1).

As implicações dos resultados obtidos com animais para a lateralização cerebral humana estão evidentemente abertas a discussão; entre outras razões, não é evidente aceitar a continuidade evolutiva da especialização hemisférica, nomeadamente se encararmos a lateralização de funções linguísticas - classicamente consideradas como especificamente humanas (para uma discussão mais aprofundada, vejam-se os comentários ao artigo referido, e ainda Corballis e Morgan, 1978; Geschwind, 1984). No contexto do presente trabalho, as observações sobre o efeito da experiência precoce na lateralização cerebral em animais constituem apenas um argumento indirecto em favor da ideia de que não é de pôr de lado *a priori* a hipótese segundo a qual a lateralização em seres humanos pode ser afectada pela experiência.

Será então a dominância do hemisfério esquerdo para a linguagem determinada por uma complexa interacção entre *nature* e *nurture*? Esta possibilidade foi examinada por investigadores confrontados com resultados empíricos algo inesperados e aparentemente interpretáveis à luz de variáveis experienciais.

Kimura (1963), por exemplo, observou uma vantagem significativa do ouvido direito numa tarefa de audição dicótica de dígitos em crianças de 4 anos; porém, não replicou este resultado com crianças masculinas de 5 anos, provenientes de meios socio-económicos inferiores (Kimura, 1967). Outros estudos abordaram explicitamente a questão da influência do nível socio-económico na lateralidade auditiva (e, supostamente, na organização funcional do cérebro). Nuns observou-se que os efeitos de lateralidade

apareciam mais precocemente (aos 4 ou 5 anos, em vez de aos 7), ou mais pronunciados, em crianças de níveis sociais superiores (Geffner e Hochberg, 1971; Pizzamiglio e Cecchini, 1971; Borowy e Goebel, 1976). Outros, todavia, não conseguiram replicar aqueles resultados (Dorman e Geffner, 1974; Geffner e Dorman, 1975).

Foi sugerido que o desenvolvimento da assimetria cerebral poderia ser facilitado por experiências de "enriquecimento intelectual" (Bever, 1970, referido também por Starck, Genesee, Lambert e Seitz, 1977). Todavia, um estudo desenhado para testar expressamente aquela hipótese foi inconcludente. Com base em estudos prévios sobre os efeitos de experiências multilingues, Starck *et al.* (*ibidem*) argumentaram que estas constituem uma forma poderosa de enriquecimento, e compararam a *performance* de crianças monolinguas e trilingues de 6 a 8 anos, emparelhadas por classe social e Q.I., numa tarefa de audição dicótica de dígitos. Numa primeira experiência, obteve-se uma interacção significativa entre ouvido e experiência linguística, em que os trilingues tendiam mais do que os monolinguas a ter vantagem do ouvido direito. Porém, numa segunda experiência, não se conseguiu replicar aqueles resultados. Mesmo que os resultados fossem consistentes, estudos deste tipo enfermam de dificuldades de interpretação: o nível socio-económico e o "enriquecimento intelectual" são condições genéricas e vagas, em que jogam de modo variável e desconhecido factores mais específicos, como o uso quantitativo e qualitativo da linguagem, a motivação para o sucesso em ambientes académicos, a valorização de certas competências em detrimento de outras, para nomear apenas alguns.

Deparam-se-nos dificuldades interpretativas análogas ao considerar os efeitos da privação extrema\*. É conhecido o caso de Genie, que desde os 20 meses até 13 anos e 9 meses de idade viveu em isolamento (Fromkin, Krashen, Curtiss, Rigler e Rigler, 1974). Depois de ter sido encontrada, Genie desenvolveu gradualmente competências linguísticas. Dois anos mais tarde foi testada por audição dicótica, e evidenciou uma superioridade do ouvido esquerdo, quer para sons ambientais, quer para palavras. Ambas eram bem marcadas, principalmente a verbal: com sons ambientais, 95% de respostas correctas à esquerda *vs* 67% à direita,\*\* e com palavras 100% no ouvido esquerdo *vs* 16%, e numa segunda sessão 0%, no direito. Poderá isto indicar que devido à ausência de interacções linguísticas - e outras - o hemisfério esquerdo não desenvolveu a sua típica superioridade para competências verbais? Obviamente Genie poderia ser um caso atípico por outras razões.

Uma via mais atraente para elucidar a questão *nature/nurture* neste contexto é tomar factores mais especificamente relacionados com a experiência linguística. Bever (1971) encontrou uma relação entre a magnitude dos efeitos de lateralidade auditiva e a tendência em adoptar uma determinada estratégia para a compreensão de frases. Vejamos em que consiste. Segundo Bever (*ibidem*, p. 245), na idade de aproximadamente 4 anos, a criança adquire uma estratégia perceptiva, segundo a qual em qualquer sequência substantivo-verbo-substantivo (S-V-S), o primeiro é tomado como actor, o segundo como acção e o terceiro como objecto (actor-

---

\*O papel da exposição à linguagem falada na ontogenia cerebral tem sido averiguado através de observações de sujeitos surdos. Observou-se que nem sempre exibem efeitos de lateralidade tão marcados para tarefas linguísticas como sujeitos normais (e.g., McKeever, Hoeman, Florian e Van Deventer, 1976; Ashton e Beasley, 1982). Todavia, além de serem possíveis explicações alternativas à da especialização hemisférica, parece-nos discutível "reduzir" a surdez a uma variável só experiencial.

\*\* Apresentando dois pares de estímulos em cada ensaio, a vantagem era mais marcada: 96% de respostas correctas *vs* 54% para ouvidos esquerdo e direito, respectivamente.

acção-objecto). Esta estratégia é particularmente apropriada para compreender frases activas, mas provoca erros nas passivas reversíveis, como "o cão é perseguido pela raposa". Usando estes dois tipos de frases, Bever (*ibidem*) avaliou até que ponto crianças de 3 a 5 anos tendem a usar aquela estratégia, analisando a percentagem de respostas correctas e o tipo de erros na compreensão de frases activas e passivas reversíveis. As crianças que davam menos erros nestas passivas, e que supostamente aderiam menos à estratégia S-V-S → actor-acção-objecto, tendiam a não exibir efeitos de lateralidade numa tarefa verbal de audição dicótica; pelo contrário, as crianças que consistentemente recorriam àquela estratégia mostravam vantagens laterais, concretamente uma superioridade significativa do ouvido direito e uma superioridade não significativa do ouvido esquerdo. Assim, o uso da estratégia perceptiva S/actor-V/acção-S/objecto parece estar associado a vantagens laterais e, segundo Bever, "since the development of these behavioral strategies appears to be responsive to actual experience, their close relation to dominance suggests that cerebral dominance develops (at least in part) in response to external experiences" (*ibidem*, p. 231).

É sabido que a experiência linguística é necessária para a aquisição da linguagem; ora havendo lateralização cerebral para funções linguísticas parece razoável supor que aquela lateralização envolva não só capacidades básicas inatas, mas também os "learned processes of utilization of language" (*ibidem*, p. 234). Nesta perspectiva, o que é relevante será estabelecer que aspectos da especialização hemisférica seriam afectados pela experiência e, reciprocamente, que experiências críticas exerceriam tal influência.

Outro argumento em favor da influência da experiência linguística nos efeitos de lateralidade encontra-se num estudo de Tsunoda

(1975). Usando o paradigma de interferência entre audição e batimento\*, ele observou uma superioridade do ouvido direito para vogais em sujeitos japoneses, mas não em sujeitos ocidentais ou japoneses com línguas maternas ocidentais. Este resultado apareceu consistentemente para os vários estímulos utilizados (em velocidade normal e rápida, por exemplo) e foi parcialmente replicado por Cooper e O'Malley (1975), que usando os mesmos estímulos e um procedimento experimental semelhante não obtiveram assimetrias de ouvido em sujeitos ocidentais. Note-se que estes sujeitos não perceberam os sons como vogais, nem sequer como sons da fala; o mesmo tinha acontecido com os japoneses quando ouviam os estímulos a uma velocidade rápida. Estes resultados indicam que a experiência precoce na língua materna pode promover o estabelecimento de modos específicos de processamento de características acústicas associadas aos sons da fala no sentido de uma diferente lateralização.

## 1.2. O papel da aprendizagem da leitura e da escrita

A possibilidade de a especialização do hemisfério esquerdo para a linguagem ser afectada pela aquisição das competências de escrita e de leitura foi recentemente discutida em pormenor por Lecours e Mehler, que a submeteram a exame sistemático. Segundo Lecours (1980), mesmo aceitando as raízes biológicas da superioridade do hemisfério esquerdo para a linguagem, permanece em aberto a possibilidade de interacção entre factores biológicos e ambientais na determinação específica daquela especialização. Mais concretamente, a aquisição da linguagem escrita teria um papel preponderante na actualização da tendência inata e universal para a superioridade do hemisfério esquerdo em competências verbais. Em

---

\*Traduziremos "tapping" por batimento; "tapping task" será pois tarefa de batimento.

abono desta suposição são apontados dados oriundos da neuropsicologia clínica infantil e adulta. Vejamos uns e outros.

### 1.2.1. Dados da neuropsicologia infantil

O argumento é como segue: crianças dexas poderiam tornar-se afásicas na sequência de lesões cerebrais à esquerda e também à direita até à idade de 5 ou 6 anos (com recuperação possível independentemente do lado da lesão); a partir desta idade, porém, só as lesões esquerdas provocariam afasia (sendo as exceções atribuíveis ao sinistrismo familiar). As lesões hemisféricas começariam pois a produzir efeitos diferenciais aproximadamente na altura em que se aprende a ler e escrever. Apesar de esta associação ser passível de várias interpretações, ela sem dúvida nos alerta para um factor experiencial, a aquisição da linguagem escrita (Lecours, 1980).

Este argumento é obviamente filiado na teoria do desenvolvimento linguístico de Lenneberg (1967): após um primeiro período de equipotencialidade seguir-se-ia uma fase crítica para o estabelecimento da lateralização. Levantam-se todavia severas reservas a um tal cenário, pois em anos recentes a hipótese de equipotencialidade precoce não tem sido confirmada. Há evidências em favor de uma especialização precoce do hemisfério esquerdo para competências relevantes para a linguagem (cf., e.g., Bradshaw e Nettleton, 1983). Revisões recentes sobre as consequências de lesões cerebrais precoces coincidem na conclusão de que mesmo na primeira infância a afasia é mais frequente após lesão esquerda (*ibidem*; Denis e Whitaker, 1977; Hécaen, 1976; Woods, 1980). A incidência de afasia por lesão do hemisfério direito é provavelmente mais frequente na primeira infância do que na idade adulta, mas, mesmo assim, reduzida (Bradshaw e

Nettleton, 1983). Outros indícios de lateralidade precoce provêm de estudos comportamentais com crianças normais com menos de cinco anos, usando tarefas de batimento (e.g., Hiscock e Kinsbourne, 1977) ou de audição dicótica (Young, 1981; Lokker e Morais, 1985).

### 1.2.2. Efeito de lesões cerebrais em letrados e iletrados

Examinemos agora o efeito de lesões cerebrais unilaterais em pacientes adultos letrados e iletrados. Segundo Lecours (1980), haveria razões para crer que nos sujeitos iletrados a afasia ocorre em consequência de lesões quer à esquerda, quer à direita, e que o prognóstico da afasia é mais favorável para adultos iletrados do que para os letrados. Essas razões provêm de relatos clínicos, alguns deles bem antigos. Por exemplo, von Mundy (1957, *cit. in* Lecours, *ibidem*) observou relativamente menos casos de afasia, ou casos mais ligeiros e transientes, em iletrados lesionados no hemisfério esquerdo e com hemiplegia direita. Mais recentemente, Cameron, Currier e Haerer (1971) registaram a incidência de afasia em pacientes hemiplégicos com lesão no hemicortex esquerdo, aparentemente muito díspare consoante a escolarização: 78% nos letrados (29 em 37 casos observados), 64% nos semiletrados (9 em 14) e apenas 36% nos iletrados (5 em 14). Os autores deste estudo sugerem que a "literacy" acentua a dominância hemisférica para a fala. Outro argumento nesta linha, sugerido por Lecours (*ibidem*), é que em fins do século XIX eram relatados mais casos negativos - não existência de afasia na sequência de lesão no hemisfério esquerdo - mesmo na área de Broca (Marie, 1926; Moutier, 1908, *in* Lecours, 1980). Tais casos ter-se-ão tornado cada vez menos frequentes possivelmente devido ao aumento gradual do nível educativo das populações.

Resultados recentes sobre o efeito de lesões no hemisfério direito em pacientes com diferentes níveis de escolarização são, neste contexto, particularmente reveladores. Joannette, Lecours, Lepage e Lamoureux (1983) observaram perturbações moderadas de linguagem em 33 de 42 pacientes com lesão à direita; um estudo sobre os factores capazes de predizer a ocorrência destas perturbações linguísticas revelou que um dos mais eficazes era justamente o grau de escolarização.

Apesar deste quadro, assim globalmente traçado, ser coerente, ao analisarmos-lo mais de perto a sua consistência esboroa-se. Vejamos primeiro a sugestão de que a associação entre lateralidade cerebral e escolarização seria responsável pelo progressivo decréscimo de casos negativos. Uma outra possibilidade para explicar esta diminuição histórica é que tais casos negativos assumissem na altura um valor especial, sendo por isso sobrevalorizados e mais relatados, como argumento contra a ideia relativamente nova e não unanimemente aceite da dominância hemisférica. Do modo análogo, factores sociais poderão enviesar de modo incontrollável os registos da incidência de determinado tipo de doença ou perturbação, na medida em que diversas camadas sociais podem usar - ou ter usado - diferentes critérios para recorrer ao hospital.

Quanto ao estudo de Cameron *et al.* (1971) convém notar que o número de sujeitos observados foi reduzido, o que limita o grau de generalidade dos efeitos obtidos. Além disto, a definição de "iletrado" e "semiletrado" assenta em critérios que não eliminam possíveis fontes de confusão, como dificuldades de aprendizagem ou intelectuais. Com efeito, os "semiletrados" eram sujeitos que conseguiam assinar o nome ou escrever uma ou duas palavras, e cuja leitura habitual era um versículo da Bíblia uma vez por semana (*ibidem*, p. 162); estes indivíduos tinham frequentado a

escola 5.6 anos, em média. Sobre os "iletrados" não é fornecida qualquer explicação, excepto a de terem andado na escola em média 2.5 anos.

Apreciemos agora com mais pormenor o trabalho recente de Joannette *et al.* (1983). O objectivo foi estudar o efeito de lesões hemisféricas direitas na fala e outras *performances* linguísticas de indivíduos dextros não-afásicos. Para tal, usaram uma bateria composta por vinte testes abordando aspectos diferentes da função linguística: narração, repetição, evocação lexical, nomeação (de desenhos), produção de frases, antónimos, leitura de texto e de palavras-ratoeira\*, anagramas, palavras-escondidas\*\*, completamento de frases, compreensão de texto, raciocínio verbal, ditado, paráfrases, cópia e leitura vertical. Esta extensa bateria foi passada a 20 sujeitos saudáveis (ou controlo) e 42 lesionados no hemisfério direito, dextros e não-afásicos. A maioria dos testes revelou uma diferença mais quantitativa que qualitativa entre os dois grupos de sujeitos; é o caso das provas envolvendo mais os níveis lexical (nomeação, evocação lexical), fonémico propriamente dito (repetição), fonémico/grafémico (ditado) e sintáctico (produção de frases, completamento). O único tipo de tarefas em que nenhum dos sujeitos controlo deu erros situa-se a nível fonético e gráfico. Porém, sete dos sujeitos lesionados tiveram disartria, disprosodia e desordens gráficas. Em geral, 33 (79%) sujeitos lesionados no hemisfério direito apresentaram um comportamento linguístico perturbado relativamente aos sujeitos controlo.

Para esclarecer mais precisamente a existência de subgrupos distintos no conjunto dos sujeitos observados, foi usado o método de Chernoff.

\* Trata-se de palavras cujas primeira e última partes constituem elas próprias palavras, por exemplo "AMIDON", amido, divisível em "ami" e "don" (*ibidem*, p. 223).

\*\* É o equivalente verbal das conhecidas figuras escondidas: uma palavra encontra-se englobada num contexto que a "disfarça". Por exemplo, "pont" em "brapontimo" (*ibidem*, p. 225).

Trata-se de um procedimento de estatística descritiva, que usa representações pictóricas cujas características são determinadas consoante os resultados do sujeito em cada teste específico. Usa-se uma face humana em que a largura da face, o comprimento do nariz, a curvatura da boca, etc refletem os *scores* dos vários testes; por exemplo, neste estudo o comprimento das sobrancelhas reflectia o *score* obtido na leitura vertical. As faces resultantes são então avaliadas por juízes segundo certos critérios; neste caso, uma vez consoante a intensidade geral das características faciais, e outra vez consoante a semelhança global das faces. Em geral, os agrupamentos feitos dividiram os sujeitos em pelo menos três "famílias" ou grupos, correspondentes a graus de gravidade das perturbações linguísticas.

A questão interessante é se existem características anatómicas, genéticas ou sociais, que permitam destringer os subgrupos com maiores ou menores perturbações. Efectivamente, a observação dos dados sugeria, e a análise discriminante confirmou, que os pacientes com perturbações mais graves são os que têm uma história familiar de ambidextrismo ou sinistrismo e menor nível educacional. Sujeitos controlo e pacientes relativamente menos perturbados tinham frequentado a escola por aproximadamente 10 anos, enquanto os restantes o tinham feito por apenas 7 anos. Este facto sugere, segundo Joannette *et al.* (*ibidem*), "a relatively more ambicerebral representation of language" (p. 239) em dextros menos escolarizados. O problema que se nos coloca face a esta interpretação é que, sendo o nível educacional dos sujeitos controlo também superior, não se pode eliminar a possibilidade de que a deficiência linguística relativamente mais acentuada nos sujeitos pouco escolarizados seja já, pelo menos em parte, anterior à própria lesão. Poderemos estar perante um efeito da escolarização não na organização cerebral, mas apenas no nível pré-mórbido de competência linguística (para uma demonstração dos efeitos de

escolarização no desempenho de sujeitos em testes de avaliação da afasia, veja-se Lecours, Mehler, Parente e colaboradores, 1987-a).

Contrariamente aos trabalhos clínicos que acabamos de rever, outros do mesmo tipo sugerem inequivocamente que "brain specialization for language does not depend on literacy" (Damásio, Castro-Caldas, Grosso e Ferro, 1976). Damásio e colaboradores observaram uma amostra de 209 letrados e 38 iletrados com lesão cerebral focal, e encontraram uma incidência de afasia por lesões no hemisfério esquerdo muito semelhante nos dois tipos de pacientes: 54% nos letrados e 55% nos iletrados. De entre os 21 afásicos iletrados, 20 tinham sofrido uma lesão no lado esquerdo; o único afásico com lesão no hemisfério direito era esquerdino. Além de não haver diferenças quantitativas entre os letrados e iletrados, os autores acrescentam que as afasias respectivas não se distinguem qualitativamente. Estes resultados foram replicados num segundo estudo por Ferro, Castro-Caldas, Martins e Salgado (não publicado), que registaram a taxa de afasia por lesão vascular esquerda: 78% nos letrados, 82% nos semiletrados e 86% nos iletrados. No entanto, estes autores notam que aparentemente a afasia de Broca aparecia com mais frequência nos letrados do que nos iletrados.

A aparente inconsistência dos resultados provenientes da neuropsicologia clínica levaram Lecours e Mehler a efectuar um estudo sistemático e cuidadosamente preparado sobre este tema (cf. Lecours, 1980), cujos resultados foram recentemente publicados.

Lecours, Mehler, Parente e colaboradores (1987-a, -b; 1988) compararam as sequelas provocadas por lesão cerebral unilateral em sujeitos letrados (cerca de 8 anos de escolarização) e iletrados (que por razões socio-económicas não frequentaram a escola). Foi usada uma adaptação em português (os dados foram recolhidos no Brasil e em Portugal) de uma

bateria para avaliação da afasia\*, composta por entrevista, tarefas de repetição, nomeação e identificação de figuras ("matching task"), que primeiramente foi passada a sujeitos saudáveis, 57 iletrados e 43 letrados. Os resultados destes sujeitos controlo mostraram inequivocamente diferenças quanto às duas populações: nas três tarefas os erros foram muito mais frequentes nos iletrados (Lecours, Mehler, Parente e colaboradores, 1987-a). Isto levou os autores a alertarem para a necessidade de atender às características socio-culturais dos sujeitos quando se avaliam os efeitos da afasia: poderá haver diferenças quanto ao desempenho de tarefas linguísticas, ou até outras, anteriores à ocorrência da lesão.

Vale a pena apreciar mais de perto estes resultados dos sujeitos controlo, em cada uma das tarefas. Na tarefa de repetição, o sujeito devia repetir as seguintes palavras: "pá, trem, pato, pratos, cavalo, cinzeiros, sabonete, embarcação", bem como as frases: "o céu está azul", "o cachorro preto grande da vizinha mordeu o menino" e "nós lhe daremos desde que ela reclame". É evidente que algumas palavras e frases são mais típicas do português do Brasil; igualmente flagrante para um "lusofono" (como dizem os autores) é a estranheza do "cachorro preto grande". De facto, quer esta frase, quer a seguinte ("nós lhe daremos...") suscitaram erros de repetição não só em iletrados como também em letrados não lesionados (cf. Lecours *et al.*, 1987-a). Os autores não referem quaisquer diferenças entre as respostas de sujeitos portugueses e brasileiros a esta tarefa, pelo que ou elas de facto não existiram ou passaram despercebidas.

Na tarefa de nomeação, apresentavam-se aos sujeitos oito desenhos de "objectos" comuns: pente, sino, banana, cachimbo, faca, violão

---

\* Trata-se da versão "M1-2" do "Protocole MT-86 de l'examen linguistique de l'Aphasie", de Lecours, Nespoulous, Joannette, Lumay, Puel, Lafonde, Cot e Rascol do Laboratório Théophile-Alajouanine em Montréal, referida nos artigos acima citados de Lecours e colaboradores.

(no Brasil)/televisão (em Portugal), uma orelha e um gato. O sujeito devia identificar estes desenhos e ainda as seguintes partes do gato, que o examinador ia apontando: "rabo; garras; bigode; olhos" (*ibidem*, p. 235, onde se podem também apreciar os estímulos). De uma maneira geral, esta tarefa era mais difícil para os iletrados. Isto não é surpreendente, atendendo a que certamente estes sujeitos estão menos familiarizados com material gráfico deste tipo.

Observando os estímulos, é evidente a dificuldade em identificar o desenho da orelha (item nº 3). De facto, os autores referem que "this particular item appeared to be somewhat difficult for literate controls as well" (não apresentam resultados por estímulo), e não deixam de salientar que para os iletrados este *item* era "nearly impossible to name": aqueles sujeitos ou diziam que não sabiam do que se tratava, ou sugeriam que talvez fosse um pato ou até genitais femininos. Outros estímulos eram de fácil identificação, não havendo quase nenhuma resposta errada em ambos os grupos de sujeitos: trata-se dos desenhos do pente, da faca e do cachimbo. Dos restantes, era mais difícil para os iletrados identificarem o desenho do gato (que de facto está numa pose algo ameaçadora e que alguns iletrados identificaram como uma onça); os iletrados portugueses também tiveram dificuldades em identificar o desenho da televisão. Apesar desta aparente interacção entre grupo e tipo de estímulo, os autores referem que mesmo eliminando as respostas ao desenho da orelha continuava a existir uma diferença significativa entre o desempenho de letrados e iletrados nesta tarefa de nomeação (*ibidem*, p. 239).

Finalmente na tarefa de "matching", os sujeitos deviam apontar, entre quatro alternativas possíveis, qual o desenho correspondente à frase emitida pelo examinador; as frases eram simples ("o homem come, a menina anda, o cachorro dorme") ou complexas ("o cavalo puxa o menino",



"o cachorro segue a mulher e o carro", "o menino pequeno empurra o grande na cadeira").\* Também nesta tarefa os erros foram mais frequentes nos iletrados que, por exemplo, apontavam para o desenho da menina a correr quando o examinador os solicitava a mostrarem "o desenho onde a menina anda" (*ibidem*, p. 240; cf. também Lecours *et al.*, 1987-b).

A comparação entre os efeitos de lesão cerebral esquerda ou direita conforme a alfabetização baseou-se no desempenho de sujeitos lesionados naquelas três tarefas, bem como no diagnóstico clínico e entrevista (Lecours *et al.*, 1988). Foram observados 95 iletrados, 48 com lesão à esquerda e 47 com lesão à direita, e 93 letrados, 61 com lesão à esquerda e 32 com lesão à direita. Os autores compararam a frequência de erros em cada caso com a frequência correspondente observada nos sujeitos controlo por meio de testes  $\chi^2$ . Por exemplo, na tarefa de repetição, a percentagem de respostas erradas nos iletrados com lesão à esquerda foi de cerca de 44%; nos iletrados saudáveis, tinha sido de 26%. O teste  $\chi^2$  revela que esta diferença é significativa.

Os resultados mostram que as tarefas de repetição e "matching" só são negativamente afectadas pela lesão no hemisfério esquerdo, tanto nos sujeitos letrados como nos iletrados. A lesão à esquerda tem também efeitos perniciosos na tarefa de nomeação, tanto para letrados como para iletrados. A única diferença intergrupar parece ocorrer quanto ao efeito da lesão no hemisfério direito nesta tarefa de nomeação (*ibidem*, pp. 575, 582). Verifica-se aqui que nos iletrados lesionados a frequência de erros é maior do que nos iletrados controlo (20.6 vs 14.1, respectivamente), sendo esta diferença

---

\* Mais uma vez se nota o sabor brasileiro das frases. A última por exemplo, e tendo em conta os desenhos apresentados, em português de Portugal seria mais "o menino pequeno empurra o grande para a cadeira". Tanto quanto o artigo esclarece, os sujeitos foram testados em Lisboa e em várias partes do Brasil da mesma maneira, o que ou não terá ocasionado diferenças, ou elas passaram despercebidas (cf. Lecours *et al.*, 1987-b; os desenhos podem observar-se na pág. 247).

significativa pelo teste  $\chi^2$ . Quanto aos letrados lesionados (à direita) a frequência de erros é 10.7, sendo a frequência de erros nos normais de 4.7. Curiosamente o acréscimo de erro nos sujeitos lesionados é, em média, praticamente idêntico, cerca de 6%, nos iletrados e nos letrados. Todavia, o teste  $\chi^2$  mostra que a diferença entre controlos e lesionados só é significativa nos sujeitos iletrados. Os autores referem também que na entrevista se evidenciavam mais perturbações nos pacientes com lesão à direita iletrados do que nos letrados. Essas perturbações eram as seguintes: redução da fala, dificuldade em encontrar palavras e produção de parafrásias fonémicas.

Com base naquelas duas diferenças entre letrados e iletrados, Lecours *et al.* (*ibidem*, p. 575) sugerem que "cerebral representation of language is more ambilateral in illiterates than it is in school educated subjects, although left cerebral dominance remains the rule in both". Perguntamo-nos até que ponto a primeira parte daquela conclusão é ditada pelos factos. Sobre os dados da entrevista, são os mesmos autores a prevenir a necessidade de ter em conta o "background" cultural dos indivíduos na avaliação da afasia - e logo, de perturbações de linguagem. Parece ser mais fácil dar o conselho do que segui-lo, pois na interpretação dos erros observados na entrevista fizeram-se apenas comparações quanto ao número total de erros nos letrados lesionados e os iletrados lesionados. Por exemplo, relativamente às dificuldades de encontrar palavras, é explicado que elas não se observavam em nenhum letrado com lesão à direita, mas verificavam-se em 7 dos 47 iletrados com lesão à direita. Factos como este são tomados como indicação de maior envolvimento do hemisfério direito em actividades linguísticas, sem os situar no contexto do desempenho de sujeitos não lesionados na mesma situação.

Sobre a diferença entre letrados e iletrados lesionados à direita, é de salientar que ela ocorre apenas numa das três tarefas estudadas. Como

se trata da tarefa de nomeação, onde, como se viu, há largas discrepâncias quanto ao grau de familiaridade dos *items* empregues, pena é que não se tenha apresentado uma análise por item, ou grupo de *items*.

Além disso, observando o Quadro 3 de Lecours *et al.*, 1987-a, não deixa de ser curioso que as diferenças entre os *scores* de letrados e iletrados sejam semelhantes nas três tarefas tanto para os sujeitos controlo como para os lesionados à direita. Na tarefa de nomeação, aliás, ela é idêntica: há 10% mais erros nos iletrados do que nos letrados, quer nos saudáveis, quer nos que sofreram lesão à direita\*.

Assim, duvidamos até que ponto é de atribuir algum significado especial ao facto do teste  $\chi^2$  ter sido significativo ao comparar os resultados de letrados e iletrados com lesão à direita naquela tarefa de nomeação.

---

\* Os valores são os seguintes: letrados controlo, 4.7; com lesão à direita 10.7; iletrados controlo, 14.1; com lesão à direita, 20.6.

## 2. Audição dicótica de palavras dissimilares e similares em adultos letrados, semiletrados e iletrados

### 2.1. Introdução

A possível influência da alfabetização na dominância hemisférica foi abordada através de estudos comportamentais com sujeitos normais em tarefas de audição dicótica.

O primeiro (Damásio, Damásio, Castro-Caldas e Hamsher, 1979) comparou o desempenho de 16 iletrados, 10 semiletrados e 21 letrados em três tarefas de audição dicótica. Os iletrados não tinham frequentado a escola, os semiletrados tinham-no feito por quatro anos e os letrados por no mínimo sete anos. Apresentavam-se-lhes i) pares de palavras dissimilares (*e.g.*, "colher-árvore", "gata-casa"), ii) pares de dígitos e iii) pares de palavras similares, diferindo apenas na consoante inicial (*e.g.*, "caneta-maneta", "mão-pão"); a tarefa era repetir as duas palavras ouvidas em cada ensaio. Como a *performance* dos iletrados e semiletrados era semelhante, estes sujeitos foram englobados num único grupo de "disletrados". Foi observada uma superioridade do ouvido direito em todos os casos, excepto na tarefa de palavras similares: nesta, o grupo "disletrado" exibiu uma vantagem do ouvido esquerdo. Enquanto com estes estímulos 13 dos 21 letrados apresentaram uma vantagem do ouvido direito e apenas 2 uma vantagem contrária, entre os 26 iletrados apenas num se observa a clássica vantagem do ouvido direito, tendo 21 exibido vantagem do ouvido esquerdo.

Este resultado é particularmente desconcertante tendo em conta que, nas tarefas com dígitos e palavras dissimilares, iletrados e semiletrados tiveram uma superioridade de ouvido direito. Convém notar que o nível de *performance* dos letrados era superior em todas as condições, principalmente nas palavras similares: os letrados identificaram

correctamente, em média, 1.8 dígitos, 1.6 palavras dissimilares e 1.6 similares por cada ensaio, enquanto os sujeitos "disletrados" deram 1.6, 1.2 e 1.0 respostas correctas respectivamente (H. Damásio\*). Como as palavras usadas na tarefa similar apenas diferiam na primeira consoante, torna-se provavelmente fácil identificar uma delas e tentar adivinhar a outra. O facto de nesta tarefa os "disletrados" terem em média apenas uma resposta correcta leva-nos a suspeitar que se não todos, a maioria deles terá entendido a tarefa como a de identificar apenas um *item*. Além disto, como as instruções permitiam a livre orientação da atenção, algum viés na situação experimental pode ter favorecido a atenção para o lado esquerdo.

O segundo estudo (Tzavaras, Kaprinis e Gatzoyas, 1981) consiste no teste clássico de audição dicótica, com três pares de dígitos em cada ensaio, devendo o sujeito imediatamente a seguir à sua apresentação relatar tudo quanto ouviu. Este teste foi passado a 60 iletrados (sem escolarização) e 51 letrados (um mínimo de 12 anos de estudo).

Numa primeira experiência foi usado o paradigma da evocação livre, podendo o sujeito orientar a atenção para o ouvido direito ou esquerdo. Verificou-se que os letrados tinham um nível de desempenho superior (79% de respostas correctas contra 54% dos iletrados). Quer uns quer outros deram mais respostas correctas no ouvido direito; todavia esta vantagem era muito mais pronunciada nos letrados. A diferença entre a média de respostas correctas à direita e à esquerda foi de 35 nos iletrados e apenas 5 nos letrados. Usando um índice de lateralidade em que esta diferença bruta é ponderada pelo nível geral de *performance*, mantinha-se a diferença significativa no grau de assimetria entre os dois grupos.

---

\* Comunicação pessoal a J. Morais

Prevenindo a possibilidade de que estes resultados fossem devidos a viés atencionais, os autores realizaram uma segunda experiência usando os mesmos estímulos, mas modificando a tarefa: agora pedia-se aos sujeitos que em cada ensaio relatassem apenas os dígitos apresentados a um ouvido, mudando a orientação da atenção em cada cinco ensaios. Nestas condições, a *performance* global de cada grupo subiu cerca de 15%, e de novo a vantagem do ouvido direito foi mais marcada nos sujeitos iletrados.

Também estes resultados suscitam perplexidade. Considerando os dados provenientes da neuropsicologia clínica, e pressupondo que as assimetrias auditivas são um índice da organização cerebral, previríamos só duas possibilidades: ou i) a vantagem de ouvido seria qualitativa e quantitativamente semelhante em iletrados e letrados, ou ii) os iletrados teriam uma *performance* menos assimétrica que os letrados. Os achados de Tzavaras *et al.* (*ibidem*) vão exactamente no outro sentido, e os autores especulam se o treino formal da leitura e escrita não contribuiria para uma maior flexibilidade na representação da função linguística. Porém, antes de nos aventurarmos por tais caminhos, convém atender a alguns aspectos metodológicos deste estudo.

Primeiro, os controlos (sujeitos letrados) eram muito mais jovens que os iletrados: a sua idade variava entre os 18 e os 48 anos (média de 30), enquanto os iletrados tinham entre 23 e 82 anos (média de 58). A maioria dos letrados (88%), e somente 14% dos iletrados, tinha menos de 40 anos. Quase metade dos iletrados (56%) tinham uma idade superior ao mais velho dos letrados. Assim, a idade pode ter sido um factor importante nas diferenças de ouvido observadas no estudo de Tzavaras e colaboradores. De facto, os *scores* de lateralidade eram muito variáveis, especialmente nos níveis etários mais avançados. Mesmo com a orientação da atenção controlada, alguns iletrados com mais de 60 anos tinham um desempenho do ouvido

esquerdo próximo da extinção; isto é pouco comum em indivíduos neurologicamente normais e está provavelmente mais ligado a perdas auditivas unilaterais do que à especialização hemisférica.

Um segundo problema diz respeito ao índice de lateralidade. Tzavaras e colaboradores usaram a diferença entre as respostas correctas à direita e à esquerda ponderadas pelo total de respostas correctas, um *score* apropriado quando o nível de *performance* é inferior a 50% (Bryden e Sprott, 1981; Marshall, Caplan e Holmes, 1975; Repp, 1977; Sprott e Bryden, 1983-a, -b). No entanto, o nível de desempenho de letrados e iletrados era cerca de 80% e 60% respectivamente; não podemos pois eliminar a possibilidade de que a magnitude dos *scores* de lateralidade esteja confundida com o nível de *performance*.

Os estudos que apresentamos a seguir visaram esclarecer a controvérsia sobre as assimetrias auditivas em sujeitos letrados e iletrados. Tendo em conta os resultados de Damásio *et al.* (1979), foi usada a técnica de apresentação dicótica com dois tipos de palavras, dissimilares e similares diferindo só na consoante inicial. Como na segunda experiência de Tzavaras *et al.* (1981), a tarefa consistia em prestar atenção a apenas um ouvido e indicar a palavra que nele tivesse sido apresentada. Trabalhámos com sujeitos iletrados, no sentido restrito do termo: pessoas que não sabem ler nem escrever por não terem frequentado a escola devido a razões socio-económicas. Os sujeitos letrados eram de dois tipos: pessoas com escolaridade reduzida, capazes de ler e escrever rudimentarmente; e sujeitos com no mínimo nove anos de escolarização, que lêem e escrevem fluente e correntemente. Chamaremos aos primeiros semiletrados, e aos segundos letrados. Como é difícil encontrar iletrados masculinos, todos os sujeitos foram femininos. A sua idade era aproximadamente a mesma nos vários grupos.

## 2.2. Experiência I

### 2.2.1. Método

#### Estímulos

Foram elaborados dois testes em que eram apresentadas dicoticamente palavras dissilábicas (CVCV) iniciadas por uma oclusiva /b, d, g, p, t, k/. Em cada teste havia 48 pares de palavras gravadas a uma velocidade de um par em cada cinco segundos. Cada par era precedido por um breve apito\*, que funcionava como um sinal de aviso. As palavras foram pronunciadas por um falante masculino português. Na construção do teste dicótico propriamente dito, elas foram sincronizadas manualmente pelo início da explosão com uma precisão de  $\pm 10$  milisegundos. Num dos testes (tarefa similar), as palavras apresentadas em cada ensaio diferiam apenas na consoante inicial. Os três tipos de contrastes fonéticos, a saber: de vozeamento ("cola-gola"), lugar de articulação ("capa-papa") e ambos ("carro-barro"), ocorriam igual número de vezes e estavam aleatoriamente distribuídos ao longo dos vários ensaios. No outro teste (tarefa dissimilar), foram usadas as mesmas palavras, mas emparelhadas de modo a serem diferentes em pelo menos os primeiros três fonemas ("pato-cura")\*. Os estímulos foram gravados num Revox A 77, e apresentados aos sujeitos através de auscultadores estereofônicos e o equipamento de alta fidelidade correspondente (leitor de cassetes portátil CP-105, Hatadi Pearce Simpson).

---

\* Um "beep", usando a onomatopeia inglesa.

\* Encontram-se os dois testes nos Anexos ED. 1 e ED. 2.

## Sujeitos

Participaram neste estudo 46 mulheres. Havia 18 iletradas, 14 semiletradas e 14 letradas. A idade das iletradas variava entre 43 e 61 anos (média de 51.8, desvio padrão 6.2). Na altura em que foram testadas, viviam numa zona rural do Norte do país\* e executavam trabalhos agrícolas e/ou domésticos. Por razões socio-económicas não tinham frequentado a escola enquanto crianças, nem recebido instrução depois de adultas.

Os grupos de semiletradas e letradas foram constituídos segundo um duplo critério: escolarização (número de anos) e fluência da leitura. Para avaliar esta última, elaborou-se um teste de leitura composto por 47 palavras (cf. Anexo L.1). Estas palavras foram dactilografadas e dispostas em três colunas de dificuldade crescente. Os sujeitos eram instruídos a ler do princípio da primeira coluna até ao fim da terceira, "tudo seguido", sendo registado o tempo de execução.

As semiletradas tinham entre 42 e 60 anos (média = 50.9, d.p. = 6.1). Oito delas habitavam em meio rural e trabalhavam em actividades agrícolas e domésticas. As seis restantes viviam numa pequena cidade\*\* onde trabalhavam como mulheres a dias. Tinham andado na escola no máximo quatro anos, o que lhes permitia um domínio básico da leitura e da escrita. Todavia, estas competências não eram necessárias, nem regularmente usadas no seu dia-a-dia. O tempo que demoravam a ler a lista de palavras foi, em média, 109 segundos (d.p.=43, mínimo de 65 e máximo de 174).

As letradas tinham entre 38 e 63 anos de idade (média de 50.9, d.p.=7.8). Dez sujeitos tinham 12 anos de escolarização, dois sujeitos 9 anos e

---

\* Vale de Cambra, S. João da Madeira; os dados foram recolhidos em 1984.

\*\*Espinho.

outros dois tinham cursos universitários. Todas liam e escreviam fluentemente, e usavam estas actividades no seu dia-a-dia. Demoravam em média 32 segundos a ler a mesma lista (d.p.=8, mínimo 19 e máximo 46).

Nenhum dos sujeitos tinha história conhecida de desordens neurológicas ou auditivas. Verificou-se a sua lateralidade manual perguntando-se-lhes qual a mão usada nas seguintes actividades: comer com colher, usar uma faca, agulha, tesoura (de costura e de poda), vassoura - que posição das mãos, acender fósforos e, conforme os sujeitos, escrever e/ou usar a enxada e a foicinha. Todos os sujeitos eram predominantemente dextros, com excepção de uma iletrada e uma semiletrada. Os resultados destas duas esquadras foram todavia integrados nas análises dos dados, pois eram semelhantes aos observados nos outros sujeitos.

### 2.3. Procedimento

Os sujeitos foram testados individualmente numa sala sossegada. Era-lhes explicado que ouviriam palavras faladas por uma voz masculina, através de auscultadores. De cada vez, haveria uma palavra num ouvido e outra, diferente, no outro ouvido, como se houvesse duas pessoas a falar, uma do lado direito e a outra do esquerdo. A tarefa consistia em prestar só atenção a um ouvido, conforme a experimentadora tivesse indicado, e em repetir o que tinham percebido nesse ouvido. Não deveriam "ligar nada" ao outro ouvido. Se não tivessem bem a certeza da resposta, deveriam mesmo assim dizer o que "lhes tinha parecido". No princípio, instruía-se metade dos sujeitos para prestarem atenção ao ouvido esquerdo (a outra metade ao ouvido direito). A cada seis ensaios, dizia-se aos sujeitos para mudarem o lado a que prestavam atenção; nessas ocasiões trocava-se também a posição

**QUADRO 1:** Média e desvio-padrão (em baixo) de respostas correctas (R.c.) e intrusões nas tarefas dissimilar e similar em cada grupo. O score máximo possível é 24. Entre parêntesis está o número de sujeitos com vantagem do ouvido esquerdo e do direito.

	Iletradas n = 18		Semiletradas n = 14		Letradas n = 14	
	Esq.	Dir.	Esq.	Dir.	Esq.	Dir.
<b>DISSIMILAR</b>						
R.c.	9.2 3.7	12.2 4.6	10.4 5.3	12.6 3.5	14.5 4.8	16.3 3.1
Intrusões	1.6 2.0 (5)	0.6 1.2 (12)	1.5 2.0 (4)	0.9 2.1 (10)	0.4 0.6 (3)	0.3 0.6 (8)
<b>SIMILAR</b>						
R.c.	10.0 2.9	11.7 2.7	9.9 2.7	11.5 3.2	12.0 3.3	15.2 2.5
Intrusões	10.0 0.9 (6)	8.6 2.2 (11)	10.2 2.9 (4)	9.8 2.6 (8)	9.4 3.0 (0)	6.9 1.8 (10)

dos auscultadores. Os estímulos eram apresentados nos dois ouvidos a uma intensidade de cerca de 70 dB SPL.

Todos os sujeitos fizeram primeiro a tarefa dissimilar, seguida pela similar, na mesma ordem de ouvido atendido. Numa curta pausa entre as duas tarefas, perguntava-se se notavam diferenças entre os dois ouvidos, principalmente de intensidade. Uma semiletrada e uma letrada queixaram-se de ouvir "mais baixinho", e de ser "mais difícil" num ouvido (respectivamente, o esquerdo e o direito); estes sujeitos foram substituídos.

Por razões que serão explicadas a seguir, os sujeitos foram retestados dois meses mais tarde na ordem contrária de ouvido atendido: os que na primeira sessão tinham começado a prestar atenção ao ouvido esquerdo, iniciavam agora o teste atendendo à direita, e vice-versa. Nesta segunda sessão, avaliava-se a acuidade auditiva dos sujeitos com um teste de audição monaural, composto por 18 dígitos (9 em cada ouvido) apresentados em três níveis de intensidade decrescentes (cerca de 75, 65 e 55 dB SPL). Não foi notada nenhuma perturbação auditiva uni- ou bilateral.

### 2.2.3. Resultados

Cada grupo deu, em média, mais respostas correctas ao atender ao ouvido direito do que ao esquerdo, quer na tarefa dissimilar, quer na similar (cf. Quadro 1). A *performance* global foi maior nas letradas (60%) do que nas semiletradas (46%) e iletradas (45%). O tipo de erros dependia fortemente da tarefa: enquanto na similar 75% do total de erros eram intrusões (relatar a palavra apresentada no ouvido não atendido), estas reduziam-se a 5% dos erros na tarefa dissimilar.

A inspecção dos resultados brutos levantou a suspeita de que a ordem de ouvido atendido tinha afectado o grau de assimetria auditiva; por isso, este factor - que, por conveniência de apresentação, denominaremos "ordem de atenção" - foi incluído nas análises estatísticas. Foi calculada uma análise de variância às respostas correctas com grupo e ordem de atenção como factores intersujeito, e ouvido e tarefa como factores intersujeito. Os efeitos principais do ouvido ( $F_{(1,40)}=23.41, p<.0001$ ) e grupo ( $F_{(2,40)}=11.54, p<.0001$ ) foram significativos. A interacção entre ouvido e ordem de atenção foi também significativa ( $F_{(1,40)}=14.95, p<.001$ ): a superioridade do ouvido direito era mais pronunciada nos sujeitos que prestavam atenção ao ouvido direito nos blocos ímpares, e ao ouvido esquerdo nos blocos pares. Nenhum outro efeito principal ou interacção atingiu significância: para o efeito da tarefa,  $F_{(1,40)}=2.16$ , para a interacção grupo x tarefa  $F_{(2,40)}=1.46$ , e para o grupo x ouvido  $F < 1^*$ .

A interacção ouvido x ordem foi um resultado inesperado, e deve ter contribuído para a grande variabilidade intersujeito. Numa tentativa de reduzir esta variabilidade, foi efectuada uma segunda sessão experimental, em que cada sujeito foi testado na ordem contrária de ouvido atendido. Infelizmente, nem todos estiveram disponíveis para esta segunda sessão: duas iletradas e 4 letradas não puderam ser retestadas.

Observámos novamente mais respostas correctas no ouvido direito para cada grupo, nas tarefas dissimilar e similar. As iletradas melhoraram o seu desempenho de 45% na primeira sessão para 52% na segunda, com aproximadamente o mesmo nível nas tarefas dissimilar e similar. Também nas semiletradas o nível de *performance* foi semelhante nas duas tarefas: 49%. As letradas, porém, obtiveram resultados superiores na tarefa dissimilar (68%) do que na similar (53%). Uma análise de variância às

\*Para pormenores, cf. Anexo ED.3/1.

respostas correctas (com grupo e ordem de atenção como factores intersujeito, e ouvido e tarefa como factores intrasujeito) confirma que os efeitos do ouvido ( $F_{(1,34)}=11.37, p=.001$ ) e grupo ( $F_{(2,34)}=5.08, p=.01$ ) foram significativos. O mesmo aconteceu com as interacções grupo x tarefa e ordem x ouvido,  $F_{(2,34)}=6.6, p=.003$  e  $F_{(1,34)}=15.1, p=.004$ , respectivamente. Nenhum outro efeito principal ou interacção foi significativo\*.

Como os grupos diferiam no nível de *performance* global e se pretendia comparar o grau de assimetrias auditivas nos vários grupos, calcularam-se índices de lateralidade supostamente não contaminados pelo nível de *performance*, para cada sujeito e condição. Usou-se a razão proposta por Halwes (1969) e Marshall, Caplan e Holmes (1975): a diferença bruta entre os *scores* à direita e à esquerda é ponderada pelo total de respostas correctas quando este é inferior a 50%, ou aquela diferença é ponderada pelo total de erros quando a *performance* global excede os 50%; este índice é conhecido por *f*. Temos pois  $f=(\text{Direita-Esquerda}) \text{ Correcto}/\text{Total correcto}$  ( $T_c$ ) para  $T_c \leq .5$  e  $f=(\text{Direita-Esquerda}) \text{ Correcto}/(1-T_c)$  para  $T_c \geq .5$ . Por clareza de exposição, apresentar-se-á este índice sob a forma  $f \times 100$ .

Como se evidencia no Quadro 2, onde se mostram os *scores f* médios para cada grupo por tarefa, sessão e ordem de atenção, o grau e por vezes a própria direcção da assimetria auditiva dependem fortemente da ordem de atenção. Os sujeitos que têm claras vantagens de ouvido direito quando atendem a este ouvido nos blocos ímpares são os mesmos que ao atender ao ouvido esquerdo nos mesmos blocos têm uma assimetria reduzida, nalguns casos até de sinal contrário. Para cada sujeito, calculou-se um valor único do índice de lateralidade, fazendo a média dos *scores f* da primeira e segunda sessões (vejam-se as médias e desvios padrões respectivos no Quadro 2). Uma análise de variância a estes valores, com

---

\*Para pormenores, cf. ED.3/2.

grupo e tarefa como factores respectivamente inter e inter-sujeito, não revelou quaisquer efeitos principais significativos ( $F < 1$ ). A interacção atingiu significância,  $F_{(2,37)} = 3.64$ ,  $p = .036$ ; isto deve-se provavelmente ao facto de nas iletradas não haver praticamente diferença entre as duas tarefas, enquanto nas letradas e semiletradas se obteve um padrão contrastante, as primeiras tendo scores  $f$  superiores na tarefa similar e as segundas na dissimilar.

De modo a confrontar mais directamente estes resultados com os de outros estudos, fizeram-se comparações emparelhadas *a posteriori* dos scores  $f$ , contrastando os grupos dois a dois separadamente em cada tarefa. Nenhum destes testes foi significativo, nem na tarefa dissimilar nem na similar.

Calcularam-se também coeficientes de correlação Pearson entre os scores  $f$  da primeira e segunda sessões: esta foi positiva e significativa na tarefa dissimilar ( $r = .31$ ,  $n = 40$ ,  $p = .26$ ), e negativa embora não significativa na similar ( $r = -.23$ ,  $n = 40$ ,  $p = .074$ ).

#### 2.2.4. Discussão

Observou-se em sujeitos letrados e iletrados uma vantagem do ouvido direito na audição dicótica de palavras; o grau desta vantagem não variava de modo significativo com a tarefa (palavras dissimilares *vs* similares), nem com o grupo (iletradas, semiletradas e letradas), mas era forte e significativamente afectado pela ordem de ouvido atendido.

Um resultado semelhante de interacção entre ouvido e ordem foi descrito por Hiscock e Kinsbourne (1980): ao relatar dígitos apresentados dicoticamente, crianças entre 3 a 12 anos tinham uma vantagem

significativa de ouvido direito só se tivessem começado a prestar atenção à direita (a vantagem desaparecia se inicialmente tivessem prestado atenção à esquerda). Para este achado foi proposta uma interpretação atencional: ele dever-se-ia à dificuldade de mudar a orientação da atenção do ouvido direito para o esquerdo (*ibidem*). Naquele estudo, todavia, a orientação da atenção mudava só ao fim de 36 ensaios (no meio da sessão), e o mesmo conjunto de estímulos era apresentado a cada ouvido. Esta situação distingue-se da nossa, em que a atenção era alterada de seis em seis ensaios, e havia diferentes subconjuntos de estímulos a serem apresentados a cada ouvido. Na nossa experiência temos pois ordem de ouvido atendido confundida com material.

Efectivamente, apesar de todos os estímulos serem palavras CVCV relativamente comuns\*, e de a ordem do seu aparecimento no teste ter sido aleatória, verificamos por inspecção dos resultados que o nível de exactidão variava ao longo dos blocos. Na tarefa dissimilar, o 1º e 7º blocos tinham os melhores níveis de exactidão (61% e 59%, respectivamente), enquanto o 4º e o 8º blocos eram os piores (47% e 39%, respectivamente); na tarefa similar, o 7º bloco era particularmente fácil (66% de exactidão), enquanto o 6º e o 8º eram os mais difíceis (38% e 34% de exactidão, respectivamente). Parece pois que o grau de dificuldade não se mantém homogéneo ao longo dos blocos: os blocos pares são mais difíceis do que os blocos ímpares, nas duas tarefas. Quando se atendia aos blocos mais fáceis no ouvido direito (e aos mais difíceis no esquerdo), obtinha-se uma vantagem de ouvido direito sobrestimada; pelo contrário, ao prestar atenção aos blocos mais difíceis no ouvido direito, e aos mais fáceis no esquerdo, obtinha-se uma fraca vantagem do ouvido direito ou até uma ligeira superioridade do esquerdo. Estas observações levam-nos a

---

\* Na falta de estatísticas sobre a frequência de palavras portuguesas na altura em que o teste foi elaborado (1984), recorreremos à nossa intuição de falantes para pronunciarmos tais juízos.

atribuir o efeito da ordem ao nosso material, e não a um *priming* atencional: o diferente grau de dificuldade das palavras nos blocos ímpares e pares enviesou os resultados no sentido de uma melhor *performance* no ouvido a que eram apresentados os blocos mais fáceis. Por isso, referir-nos-emos a este efeito indiferentemente como viés de material ou efeito de ordem.

Este efeito é uma ilustração flagrante da sensibilidade do teste de audição dicótica a variáveis que não a especialização hemisférica, e limita claramente inferências sobre a lateralização cerebral a partir das assimetrias auditivas observadas. Por exemplo, o facto de as correlações entre as primeira e segunda sessão serem particularmente baixas (na tarefa dissimilar,  $r=.31$ ), e mesmo negativas (na tarefa similar,  $r=-.27$ ), é facilmente compreensível à luz do efeito do viés de material. O mesmo raciocínio se pode aplicar à consistência da vantagem de ouvido: só 58% dos sujeitos mantiveram a sua assimetria inicial (à esquerda ou à direita) na tarefa dissimilar, e ainda menos (43%) na tarefa similar; trata-se de números relativamente baixos, comparando com os 70 a 80% registados em estudos de fidelidade das assimetrias laterais em situações de audição dicótica (veja-se Blumstein, Goodglass e Tartter, 1975; Bakker, van der Vlugt e Claushuis, 1978). Também a percentagem de sujeitos com superioridade do ouvido direito, 65% considerando todos os grupos, sessões e tarefas, fica aquém dos 70 a 75%, incidência tipicamente observada em experiências de audição dicótica (Geffen, 1978). Isto sugere que o viés de material pode atenuar não só a magnitude de assimetrias auditivas (Pipe, 1985), como também a própria incidência de vantagem do ouvido direito.

Finalmente, é possível que o efeito do viés de material tenha obscurecido o papel de outras variáveis, como a tarefa e o grupo, nomeadamente através da grande variabilidade intersujeito. Tendo em conta estas considerações passemos agora à comparação destes resultados

com os de estudos anteriores. Tomaremos primeiro os resultados observados na tarefa dissimilar e depois os da tarefa similar.

Os dados presentemente conhecidos partilham um achado comum: a clássica superioridade do ouvido direito (em letrados) encontra-se também em sujeitos iletrados desempenhando tarefas de audição dicótica com palavras foneticamente dissimilares (dígitos, em Tzavaras *et al.*, 1980; dígitos, di- e trissílabos, em Damásio *et al.*, 1979; palavras CVCV, nesta experiência). Todavia os três estudos obtiveram diferentes resultados quanto ao grau de assimetria dos vários grupos. No de Damásio *et al.*, as comparações intergrupo indicaram que os letrados tinham uma superioridade do ouvido direito mais pronunciada que os "disletrados" na tarefa de palavras dissimilares (22% e 11%\* respectivamente); o mesmo não acontecia com os dígitos, onde os dois grupos obtiveram assimetrias quantitativamente semelhantes. Pelo contrário, Tzavaras *et al.* descrevem uma vantagem de ouvido direito muito mais forte para iletrados do que para letrados (35% e 9%, respectivamente, considerando as duas experiências) na audição dicótica de dígitos. No nosso estudo, a vantagem de ouvido direito para as palavras dissimilares foi apenas ligeiramente, e de modo não significativo, superior nos iletrados (11% *vs* 9% nos semiletrados e 5% nos letrados, considerando as duas sessões).

Possíveis factores dando conta do achado de Tzavares *et al.* (*ibidem*), nomeadamente a idade dos sujeitos, já foram apresentados na Introdução. O facto de este resultado não ter sido replicado por Damásio *et al.* (1979), nem corroborado (pelo menos claramente) neste estudo, leva-nos a concluir para já que não há nenhum factor especificamente ligado ao facto

---

\* Optamos aqui por usar como índice de lateralidade a diferença bruta expressa em percentagem porque, apesar de não ter em conta o nível de performance, é a única maneira possível de fazer comparações entre os vários estudos.

de ser iletrado que promova maior assimetria auditiva na situação de audição dicótica de palavras.

Na tarefa similar, todos os grupos tiveram, em média, uma superioridade do ouvido direito. Este resultado replica parcialmente o de Damásio *et al.* observado nos sujeitos letrados, mas é claramente contraditório com o seu achado de uma inversão no tipo de assimetria auditiva nos sujeitos iletrados e semiletrados. Como explicar esta discrepância? Na nossa experiência, iletradas e semiletradas deram quase tantas intrusões como respostas correctas, enquanto as letradas tinham uma incidência relativamente menor de intrusões. Isto sugere que os sujeitos letrados conseguiam controlar e orientar mais eficazmente a sua atenção e, reciprocamente, que iletrados e semiletrados eram mais afectados pelos estímulos que iam aparecendo no ouvido não atendido. Podemos especular que, ao permitir que a atenção se orientasse livremente (como na situação de Damásio *et al.*) iletrados e semiletrados tendem a ser mais sensíveis do que letrados a variáveis ambientais, como, por exemplo, a posição do experimentador. O facto de não termos replicado a vantagem de ouvido esquerdo para palavras similares numa tarefa semelhante, mas em que a orientação da atenção é controlada, apoia a interpretação daquele resultado algo insólito em termos de viés atencional.

Comparações mais finas das diferenças de ouvido nos vários grupos foram sem dúvida dificultadas devido ao efeito/viés de material. Apesar de ser claro que a semelhança fonética das palavras não provoca *per se* uma alteração no sentido das assimetrias auditivas nos iletrados, o certo é que resta por esclarecer qual a sua influência na *performance* em cada um dos grupos considerados. É também conveniente apreciar a magnitude e incidência da vantagem de ouvido direito para palavras dissimilares não contaminadas pelo viés de material.

Para dispor de dados mais fiáveis através da eliminação dos efeitos de material e de um aumento no número de sujeitos observado, foi realizada uma segunda experiência. Pensou-se que seria preferível misturar ensaios com pares de palavras dissimilares e similares para contrariar a eventual adopção de uma estratégia especial para lidar com a semelhança fonética, nomeadamente tentando adivinhar as palavras.

## 2.3. Experiência II

### 2.3.1. Método

#### Estímulos

As gravações usadas na Experiência I foram modificadas do modo que a seguir se explica. Selecionaram-se 72 pares de palavras com base na percentagem de erros observada na Experiência I, 1ª sessão, eliminando os pares ora muito fáceis (percentagem de erros igual ou inferior a 12%), ora muito difíceis (percentagem de erros igual ou excedendo 83%). Os pares seleccionados foram reordenados em 12 blocos com 3 pares de palavras dissimilares e 3 pares de palavras similares cada um. Os blocos foram compostos de modo a manter aproximadamente o mesmo nível de exactidão ao longo do teste, nomeadamente para o conjunto dos blocos pares e o dos ímpares; conseguiu-se de facto um nível de exactidão equilibrado nestes blocos para as palavras dissimilares (44.1% vs 44.5%, em média, 1ª sessão da Experiência I nos blocos ímpares e pares, respectivamente) e para as palavras similares (45.3% vs 44.9%, *idem*).

Antes dos ensaios experimentais incluíram-se dois ensaios de treino; estes foram escolhidos entre os pares com uma taxa de erros inferior

a 12%, sendo um dissimilar e outro similar\*. Como na primeira experiência, cada par era precedido por um "beep" e gravado a um ritmo de um par cada cinco segundos. Usou-se o mesmo equipamento.

### Sujeitos

Participaram nesta experiência 74 mulheres dextras, 22 iletradas, 20 semiletradas e 32 letradas. As iletradas tinham entre 37 e 66 anos (média=51.2, d.p.=9.4); pertenciam a meios rural e semi-rural do Norte do país\*, onde trabalhavam em actividades agrícolas ou domésticas. Nunca tinham frequentado a escola por razões socio-económicas. As semiletradas tinham entre 33 e 64 anos de idade (média=46.2, d.p.=6.8). O seu meio e ocupação era semelhante ao das iletradas. Tinham no máximo quatro anos de escolarização, e não usavam regularmente as suas competências básicas de leitura e escrita. O tempo que levaram, em média, a ler a lista previamente descrita era de 80.3 segundos (entre um mínimo de 40 e um máximo de 130, d.p.=27.3). A idade média das letradas era de 45.8 anos (entre 36 e 66 anos, d.p.=7.3). Tinham frequentado a escola no mínimo doze anos. A maioria (19 das 32) tinha completado um curso universitário, e trabalhava como docente liceal; das restantes, 6 eram enfermeiras e 7 professoras primárias. Para ler a lista de palavras demoraram em média 26.7 segundos (entre 21 e 35, d.p.=3.7).

De nenhum participante era conhecida uma história de perturbação neurológica ou auditiva. Não se observou nenhuma perda auditiva evidente num teste monaural de acuidade auditiva, composto por 18 palavras CVCV, 9 em cada ouvido, apresentadas em três níveis de

---

\* O teste encontra-se no Anexo ED.4.

\*Vale de Cambra, Vila da Feira.

**QUADRO 3:** Média e desvio-padrão (entre parêntesis) das respostas correctas nos ouvidos esquerdo (Esq.) e direito (Dir.), para os pares dissimilares e similares, em cada grupo. O score máximo possível é 18. Incluem-se scores  $f$  (desvio-padrão) e a probabilidade de testes  $t$  respectivos contra zero.

GRUPO	ILETRADAS N = 22		SEMILETRADAS N = 20		LETRADAS N = 32	
	Esq.	Dir.	Esq.	Dir.	Esq.	Dir.
DISSIMILAR	10.3 (2.7)	12.5 (2.2)	12.6 (2.6)	14.1 (2.1)	14.8 (2.0)	15.7 (1.8)
$f$ MEDIO	16.2 (20.2) $p = .0007$		15.9 (31.7) $p = .0214$		19.6 (38.1) $p = .0037$	
SIMILAR	7.7 (1.8)	9.8 (2.3)	8.5 (1.9)	11.0 (2.4)	9.5 (1.7)	13.4 (1.7)
$f$ MEDIO	13.1 (20.1) $p = .0035$		17.1 (19.8) $p = .0006$		30.9 (20.1) $p < .0001$	

intensidade decrescente (75, 55 e 45 dB). As palavras usadas neste teste de acuidade auditiva foram as que se tinham eliminado do primeiro teste de audição dicótica. A lateralidade manual foi verificada como na Experiência I. Recolheram-se também informações sobre sinistrismo familiar. Quatro letradas, 4 semiletradas e 2 iletradas tinham filhos esquerdinos; uma letrada tinha o pai, e três semiletradas um(a) irmã(o) esquerdinos.

### Procedimento

O procedimento foi semelhante ao da primeira experiência, excepto no seguinte. Os sujeitos fizeram primeiro o teste de acuidade auditiva monaural, e depois a tarefa experimental. A seguir a terem ouvido os ensaios de treino, explicava-se-lhes que as palavras que apareceriam em cada ouvido seriam por vezes muito diferentes, como "cara" de um lado, e "dente" do outro; por vezes, seriam muito parecidas, como "duro" de um lado, e "puro" do outro. Mesmo se tivessem a impressão de estarem a ouvir palavras idênticas, ou quase, em ambos os ouvidos, deveriam esforçar-se sempre por prestar atenção só ao ouvido previamente indicado. Em cada grupo, metade dos sujeitos começou a prestar atenção ao ouvido esquerdo e metade ao direito; a orientação da atenção mudava a cada seis ensaios, havendo pois seis blocos para cada ouvido.

### 2.3.2. Resultados

Como se vê no Quadro 3, observou-se uma vantagem do ouvido direito em todos os grupos, quer para as palavras similares, quer para as dissimilares.

Esta vantagem estava presente em 64% das iletradas, 60% das semiletradas e 53% das letradas para os pares dissimilares; nos pares similares, os valores correspondentes foram 77%, 80% e 94%, respectivamente. A vantagem oposta do ouvido esquerdo observou-se, para os pares dissimilares, em 5% das iletradas, 10% das semiletradas e 13% das letradas; para os pares similares, temos respectivamente, 23%, 14% e 3%. Calcularam-se testes  $\chi^2$  separadamente para os pares dissimilares e similares, comparando a incidência da vantagem do ouvido direito nas letradas e iletradas; os resultados,  $\chi^2=0.237$  e  $\chi^2=1.846$ , revelaram que aquelas diferenças não são significativas.

O nível de *performance* foi mais elevado para os pares dissimilares, 74%, do que para os similares, 56%, e para as letradas, 75%, do que para as iletradas, 56%, ficando as semiletradas aproximadamente a meio (64%). Calculou-se uma análise de variância das respostas correctas com grupo e ordem de ouvido atendido como factores intersujeito, e ouvido e tipo de par (dissimilar *vs* similar) como factores intrasujeito. Dos efeitos principais, foram significativos os de grupo ( $F_{(1,68)}=37.2$ ,  $p<.0001$ ) e tipo de par ( $F_{(1,68)}=276.52$ ,  $p<.0001$ ). A interacção entre ouvido e tipo de par foi também significativa,  $F_{(1,68)}=16.42$ ,  $p=.0001$ , bem como a interacção tripla grupo x ouvido x tipo de par,  $F_{(2,68)}=6.23$ ,  $p=.0003$ . Isto deve-se provavelmente ao facto de iletradas e semiletradas terem mais ou menos a mesma assimetria em ambos os tipos de palavras, enquanto as letradas evidenciam uma assimetria mais marcada nos pares similares. Para esclarecer melhor este ponto, analisamos a interacção entre grupo e ouvido separadamente para os pares de palavras dissimilares e similares: no primeiro caso, ela revelou-se não significativa ( $F_{(2,68)}=2.59$ ), e aproximou-se da significância no segundo,  $F_{(2,68)}=3.08$ ,  $p=.0524$ .

A interacção entre grupo e tipo de par esteve também próxima da significância,  $F_{(2,68)}=3.02$ ,  $p = .0554$ , reflectindo o facto de o efeito pernicioso da semelhança fonética ser ligeiramente maior para as letradas e semiletradas (-21% e -20%, respectivamente), do que para as iletradas (-14%). Todavia, a interacção tripla grupo x tipo de par x ordem foi significativa,  $F_{(2,68)}=3.46$ ,  $p = .0371$ : o decréscimo relativamente menor das iletradas entre pares dissimilares e similares verificava-se apenas se elas tinham prestado atenção ao ouvido direito nos blocos ímpares. Apesar de não termos conseguido eliminar alguns efeitos locais da ordem das palavras no novo teste, vale a pena notar que nem o efeito principal da ordem, nem a interacção ouvido x ordem foram significativos ( $F < 1$  em ambos os casos). Todas as interacções restantes, nomeadamente a do grupo x ouvido ( $F < 1$ ) não foram significativas.\*

Como o nosso interesse principal é comparar as diferenças de ouvido entre letradas e iletradas, e a diferença bruta está confundida com o nível de desempenho, fez-se uma análise de variância dos *scores f* com ordem, grupo e tipo de par como variáveis, as duas primeiras intersujeito e a última intrasujeito. Esta análise mostrou que nenhum dos efeitos principais nem as interacções são significativos, nomeadamente o efeito do grupo  $F_{(2,68)}=2.15$ , e a interacção grupo x tipo de par,  $F_{(2,68)}=1.17$ .\* Para possibilitar comparações mais claras com outros estudos, efectuaram-se análises separadas por tipo de par. Estas mostraram que o efeito de grupo era significativo nos pares similares ( $F_{(2,18)}= 6.0$ ,  $p=.004$ ), mas não nos dissimilares ( $F < 1$ ). Este efeito para pares similares deve-se ao facto de o *score f* das letradas ser muito superior ao do dos outros dois grupos (cf. Quadro 3). Veremos na Discussão interpretações possíveis deste resultado, sendo na altura apresentadas outras análises relevantes.

\*Para pormenores, cf. Anexo ED.3/3.

\*\*Para pormenores, cf. Anexo ED.3/4.

Os erros observados dependiam do tipo de par: as intrusões foram muito mais frequentes nos pares similares (cerca de 35% do total de respostas) do que nos dissimilares (cerca de 2%). Outros erros foram mais frequentes nos pares dissimilares (cerca de 25%) do que nos similares (cerca de 10%). Os *scores f* para pares dissimilares e similares estavam positivamente correlacionados só no caso das letradas ( $r = .29$ ,  $n=32$ ,  $p = .049$ ). As correlações correspondentes para as iletradas e semiletradas foram  $r = .20$  e  $r = .02$ , respectivamente.

Como a idade dos sujeitos variava num leque relativamente vasto, e a idade pode ser um factor importante na determinação quer do nível de *performance* quer dos índices de lateralidade, calcularam-se coeficientes de correlação entre a idade e estas variáveis. O nível de *performance* nos pares dissimilares apresentava uma correlação negativa com a idade,  $r = -.24$ ,  $n=74$ ,  $p = .01$ ; não havia porém correlações significativas entre a idade e o desempenho nos pares similares, nem entre idade e *scores f*.

Poderia supor-se que estes resultados estariam enviesados pelo sinistrismo familiar (SF+), principalmente no grupo das semiletradas em que 7 dos 20 sujeitos a relataram. Porém, a inspecção dos resultados dos 14 sujeitos SF+ não revelou nenhum padrão particular nem na magnitude de assimetria auditiva, nem na incidência da superioridade do ouvido esquerdo.

#### 2.4. Discussão

De modo consistente com os resultados da primeira experiência, observou-se uma superioridade do ouvido direito na identificação de palavras dissimilares e similares apresentadas dicoticamente, não só em sujeitos de escolaridade elevada (grupo das letradas), como também em sujeitos de

escolaridade reduzida ou nula (grupos das semiletradas e iletradas). A reordenação dos pares de palavras ao longo do teste, para que a dificuldade dos estímulos apresentados a cada ouvido fosse aproximadamente a mesma, eliminou a interacção ouvido x ordem previamente observada, confirmando assim que se tratava de um efeito de viés de material. Como no caso anterior, examinar-se-ão em separado os resultados obtidos com os pares dissimilares e com os pares similares.

Relativamente aos pares dissimilares, nem nas análises das respostas correctas nem nas dos *scores f* emergiram de modo significativo diferenças inter-grupo quanto à vantagem de ouvido direito. Estes resultados replicam os obtidos na Experiência I e corroboram a conclusão de que o facto de ser iletrado não provoca uma vantagem de ouvido direito mais marcada na audição dicótica de palavras dissimilares.

Quanto aos pares similares, as letradas tiveram em média uma vantagem de ouvido direito superior à das semiletradas e iletradas. Dever-se-ia então concluir que, não a alfabetização *per se*, mas o grau de escolarização, afecta o grau de dominância do hemisfério esquerdo para discriminações fonéticas mais subtis? O problema de uma tal interpretação é que havia diferenças importantes nas *performances* globais, quer intra, quer inter-grupo, e as diferenças na magnitude de assimetria auditiva podem estar mais relacionadas com a variabilidade inter-sujeito no nível de *performance* do que com características grupais de escolarização.

Como mostra a Figura 1, em que estão representados os *scores f* individuais para os pares similares consoante o nível de *performance* global, os *scores f* maiores estão associados a um melhor desempenho, enquanto *scores f* pequenos ou negativos tendem a ocorrer em níveis baixos de *performance*. Efectivamente, os *scores* de lateralidade estão positivamente correlacionados com o nível de *performance*,  $r=.45$ ,  $n=74$ ,

$p=.001$ . Para os pares dissimilares, o  $r$  correspondente é de .06. A distribuição dos grupos quanto aos níveis de *performance* é claramente desigual: não se encontra nenhuma iletrada nos níveis superiores e só uma letrada (3%) vs 64% das iletradas e 40% das semiletradas se situam abaixo ou ao nível de 50% de exactidão.

Um dos problemas ao inferir sobre diferenças grupais na assimetria hemisférica diz respeito à homogeneidade das características do grupo para além das especificamente em questão (cf. Schwartz e Kirsner, 1984, para uma discussão detalhada). Tendo em conta esta advertência, tentámos encontrar entre as iletradas, semiletradas e letradas um subgrupo que tivesse níveis de *performance* semelhantes. Como se pode ver na Figura 1, 36% das iletradas, 40% das semiletradas e 34% das letradas situam-se num espaço comum de exactidão, entre 63% e 53%, limites que correspondem aproximadamente à melhor *performance* das iletradas e à pior das letradas\*. Estes sujeitos caracterizam-se por níveis de *performance* intermédios, 56% nas iletradas, 57% nas semiletradas e 58% nas letradas. Os *scores f* destes subgrupos são, em média, 20.9 (d.p.=15.1), 20.0 (d.p.=22.3) e 25.6 (d.p.=15.2), respectivamente. Uma análise de variância a estes *scores* mostra que o efeito de grupo não é significativo ( $F<1$ ); mais ainda, os *scores f* das letradas não diferem significativamente dos das iletradas, como revelou uma comparação *a posteriori* ( $F<1$ ).

Os nossos dados oferecem pois o seguinte quadro: sujeitos com baixo nível de desempenho, constituídos principalmente por iletradas e semiletradas, apresentam em média assimetrias pouco marcadas (as 14 iletradas e 8 semiletradas com *performances* inferiores a 50% tiveram um *f* médio de 8.4, d.p.=21.8, e de 9.8, d.p.=16.9, respectivamente); sujeitos com

---

\* De facto, um sujeito letrado teve uma *performance* global de 47%. Não se tomou este valor como critério para definir o limite inferior porque ele se desvia mais de 2 desvios-padrões da média das letradas (63.5% d.p. = 6.6) e é provavelmente atípico para estes sujeitos.

um alto nível de desempenho, indivíduos escolarizados em quem se observam maiores assimetrias (as 4 semiletradas e as 20 letradas cuja performance era superior a 62.5% tiveram um  $f$  médio de 31.7, d.p.=21.5%); e sujeitos de nível médio, com aproximadamente a mesma proporção de iletradas, semiletradas e letradas, em quem se observam graus de assimetria intermédios e semelhantes para os vários grupos.

Poder-se-ia especular que diferentes níveis de exactidão numa determinada tarefa serão conseguidos pela acção de processos ou mecanismos parcialmente distintos, mais ou menos lateralizados, que teriam uma relação probabilista com o grau de escolarização. Infelizmente, não podemos para já e com base nos dados apresentados, decidir se as diferenças de *performance* que observámos se devem à menor ou maior capacidade de atenção selectiva ou à eficácia e/ou disponibilidade de mecanismos específicos de percepção da fala; o papel da alfabetização e escolarização naquelas variáveis está ainda por esclarecer. Todavia, os nossos resultados mostram claramente que, se se toma o cuidado de comparar grupos homogéneos quanto ao nível de *performance*, o grau de vantagem de ouvido direito para a linguagem não depende da alfabetização.

### 3. Conclusão geral

O presente estudo, usando a técnica de audição dicótica com pares de palavras dissimilares e similares, fornece provas em favor da dominância do hemisfério esquerdo, inferida através da vantagem do ouvido direito, para a compreensão da linguagem em sujeitos semiletrados e iletrados, tal como nos letrados. Além disso, supondo que a variação intergrupo na magnitude da vantagem de ouvido reflete diferenças grupais no grau de especialização hemisférica - uma suposição implícita em parte

da literatura revista - , não obtivemos indicação de diferenças no grau de lateralização da linguagem entre as três populações. Apesar de só podermos especular sobre as razões pelas quais Tzavaras *et al.* (1981) encontraram uma maior vantagem de ouvido direito para palavras dissimilares nos iletrados do que letrados, e porque Damásio *et al.* observou uma vantagem de ouvido esquerdo para palavras similares nos iletrados, vale a pena salientar que este estudo evitou duas fontes de artefacto possíveis: a não-dissociação das variáveis alfabetização e idade, como aconteceu no estudo grego, e a falta de controlo da orientação da atenção, como no primeiro trabalho português.

Finalmente, este estudo alerta para duas importantes questões metodológicas. Primeiro, a precaução de controlar os estímulos em termos da sua estrutura fonética não garante a adequação do teste para avaliar diferenças de ouvido (cf. Experiência I). Se, por alguma razão, for preferível não apresentar o mesmo estímulo igual número de vezes a cada ouvido, torna-se imprescindível determinar a dificuldade do reconhecimento de cada par de estímulos antes de passar à investigação sobre lateralidade. Segundo, ficou claramente estabelecido que, para os pares similares, o índice de lateralidade usado estava correlacionado com a exactidão. Assim, comparações de assimetrias auditivas entre diferentes grupos só podem ser validamente interpretadas tomando explicitamente em conta o nível de *performance* característico dos grupos em questão.

## CAPÍTULO V

## IDENTIFICAÇÃO DE PALAVRAS FALADAS EM CONDIÇÕES DIFÍCEIS

Ao longo deste capítulo debruçar-nos-emos sobre a influência do conhecimento da linguagem escrita na identificação de palavras faladas em condições difíceis. Começaremos por apresentar dados relevantes provenientes de variados paradigmas experimentais. Descreveremos alguns exemplos não triviais do efeito da escrita em tarefas que envolvem a linguagem falada, observados em sujeitos letrados. Passaremos então uma breve revista à investigação com adultos iletrados em tarefas de análise explícita da fala.

Os resultados descritos na literatura e algumas das nossas observações quanto ao desempenho de adultos iletrados na tarefa de audição dicótica tornam plausível a hipótese segundo a qual a alfabetização afectaria o modo como se percebe a fala, em condições onde estejam ausentes ajudas contextuais e pragmáticas. Para testarmos aquela hipótese foi efectuado um conjunto de três experiências de identificação de palavras em condições difíceis, em que se comparou o desempenho de iletrados adultos e crianças pré-letradas com o de sujeitos letrados de várias idades e graus de escolarização. Delas se dará conta detalhadamente no restante deste capítulo.

### 1. Introdução geral

Os estudos sobre percepção de fala em adultos têm sido feitos quase exclusivamente com estudantes universitários. Pressupõe-se que

compreender a fala é uma capacidade tão básica que é possível generalizar os resultados do jovem universitário para o ser humano adulto em geral. Todavia, será que podemos aceitar incondicionalmente este pressuposto?

Os resultados das experiências I e II mostraram uma diferença inequívoca quanto ao nível de exactidão no reconhecimento de palavras faladas conforme a escolarização. Este achado é tão robusto (cf. também Lecours *et al.*, 1987-a) que se torna fácil sucumbir à tentação de *a posteriori* o considerarmos trivial. Ele coloca-nos, todavia, perante uma questão que é importante resolver: a que se deverá esta aparente dificuldade dos iletrados relativamente aos letrados no reconhecimento de palavras faladas? Traduzirá ela alguma diferença subjacente quanto aos processos de compreensão da mensagem falada? Ou tratar-se-á de uma diferença meramente quantitativa, até de um qualquer artefacto metodológico? Estas questões justificam por si só a necessidade de examinar se e até que ponto o reconhecimento da fala é afectado do nível de escolarização. Elas não são porém o único argumento em prol de tal iniciativa. Há outras duas razões importantes. Uma provém de estudos com adultos iletrados em tarefas de manipulação da fala; outra tem a ver com efeitos do conhecimento da escrita observados na *performance* de letrados em certas tarefas com linguagem falada. Começemos por apreciar estes últimos.

### 1.1. A influência do conhecimento da escrita em certas tarefas com linguagem falada (sujeitos letrados)

Talvez a demonstração mais flagrante da influência do conhecimento da escrita em tarefas que apenas envolvem a linguagem falada se encontre no achado de Seidenberg e Tannenhaus (1979) sobre o efeito da ortografia na detecção de rima. É mais rápido detectar a rima entre duas palavras faladas quando elas se escrevem de maneira semelhante

(e.g., "pie" - "tie") do que quando são ortograficamente dissemelhantes ("pie" - "guy"). Acontece um fenómeno análogo para detectar que duas palavras não rimam: demora mais tempo responder "não" a "howl-bowl" ([haul] - [boul]) do que a "howl-roll" (*ibidem*, Experiência 2; p. 551).

Outra demonstração dos efeitos da escrita na representação da linguagem falada é o fenómeno descoberto por Bertelson (1972) de que a localização aparente de um ruído numa frase é afectada pela direcção da escrita. Suponhamos uma situação em que se ouve uma frase com um "click" que acontece num determinado ponto daquela frase; a tarefa é localizar a ocorrência desse "click" no contexto da frase (entre as palavras x e y, por exemplo). O "click" é apresentado de modo a que pareça provir de fontes espaciais diversas relativamente à frase, ou seja, ou mais à esquerda ou à direita. Quando a fonte espacial aparente do ruído se situa à esquerda da frase, os sujeitos julgam a sua ocorrência mais cedo do que realmente aconteceu (por exemplo, se cada letra representar uma palavra, uma situação "a b c /"click"/ d" é percebida como "a b /"click"/ c d". Se a fonte aparente do ruído relativamente à frase se situa à direita, então os sujeitos localizam o ruído ainda mais à direita. Este padrão de resultados é observado com frases em inglês e francês, línguas cuja direcção de leitura é da esquerda para a direita.

Sujeitos bilingues em francês e hebreu evidenciam uma interessante dissociação: se testados em francês, os seus resultados são como descrevemos acima, se testados em hebreu (cuja direcção é da direita para a esquerda), observa-se um padrão invertido. Estes resultados são consistentes com a ideia de que o viés quanto à localização do ruído está associado à direcção da escrita (*ibidem*). Dados recentes ainda não publicados mostram que aquele fenómeno só se observa em crianças com pelo menos 7 anos de escolarização (referido por Bertelson e De Gelder, no

prelo). Este facto vem apoiar a suposição de que o fenómeno em questão é mediado por uma representação da fala determinada em parte pelo conhecimento da escrita.

Finalmente, tem-se observado a influência da ortografia nouro tipo de tarefas, aquelas que envolvem análise explícita da fala. Ehri e Wilce (1979) relataram que ao contar o número de fonemas de palavras faladas as crianças são por vezes induzidas em erro pelo número de letras das respectivas representações ortográficas: por exemplo, contam menos um fonema em "rich" [ritʃ] do que em "pitch" [pitʃ]. Resultados análogos são mencionados por Mann (1986) a propósito da *performance* de crianças japonesas na contagem de fonemas e *moræ*\*.

É evidente que os vários factos experimentais que acabámos de ver se referem a tarefas bem distintas e que provavelmente os processos implicados em cada uma delas são diferentes (isto é particularmente óbvio quanto às tarefas de análise explícita da fala, como se discutirá mais à frente). Nessa medida, eles não permitem elucidar inequivocamente a que nível ocorrem os presumíveis efeitos da escrita. Mostram porém de forma clara que o tratamento da linguagem falada pode ser afectado de modo não trivial pelo conhecimento da linguagem escrita.

## 1.2. Análise da fala em sujeitos iletrados

Investigação recente com adultos iletrados, iniciada por J. Morais e colaboradores, tem vindo a desvendar um importante efeito da alfabetização em certas tarefas que envolvem a linguagem falada. Trata-se

---

\**Mora* é uma unidade rítmica da língua japonesa, que corresponde aproximadamente à sílaba.

da análise explícita das unidades da fala (por exemplo, segmentar uma palavra nas suas sílabas e fones constituintes).

Claro que o problema da análise consciente dos constituintes da fala não é equivalente à questão que colocamos sobre o eventual efeito da alfabetização na percepção de fala; enquanto esta é, em larga medida, dependente de processos inconscientes, a primeira, por definição, refere-se aos processos conscientes de manipulação do percepto. Isto porém não impede que possamos ganhar alguns *insights* relevantes a partir daquela investigação. Como ela é apresentada em pormenor num trabalho recente em língua portuguesa (Cary, 1988), limitar-nos-emos aqui a revê-la muito sucintamente.

O interesse pelas capacidades de análise da fala em adultos iletrados surgiu no contexto da investigação sobre leitura. Façamos pois uma breve incursão nesse domínio.

#### 1.2.1. A relação entre a aprendizagem da leitura e a segmentação da fala

Falar e escutar ("listening") são, nas palavras de Mattingly (1972), actividades linguísticas primárias; elas desenvolvem-se em praticamente todas as crianças (com excepção de casos de surdez, ou outras deficiências graves) por mera exposição ao meio linguístico adulto, dispensando um treino específico. A leitura, porém, é uma actividade linguística secundária (*ibidem*), que não aparece espontaneamente, requer antes instrução formal.

Conseguir entender uma ortografia implica manipular as unidades linguísticas representadas através de sinais gráficos e estabelecer a sua correspondência com a fala. Isto acontece provavelmente de modos diferentes consoante o sistema de escrita (cf., e.g., Holender, 1985; Rozin e Gleitman, 1977; Taylor, 1981). Além da capacidade em fazer discriminações

fonéticas mínimas, a leitura hábil requer um conhecimento explícito da estrutura fonética da fala. Não basta perceber que [patu] é diferente de [gatu], é importante reconhecer que a diferença se encontra no primeiro segmento. O leitor tem de operar uma análise da fala, e atingir a chamada "linguistic awareness" (Mattingly, 1972), ou seja, a capacidade de representar a fala como uma sucessão das unidades linguísticas empregues na escrita; no caso da escrita alfabética, como uma sucessão de fonemas - daí a expressão de "phonological awareness", (Liberman, 1971, Mattingly, 1972), que referiremos em português como a consciência da estrutura fonética da fala (Cary e Morais, 1979) ou consciência fonética\*.

A noção de "linguistic awareness" é operacionalizada através de tarefas de segmentação da fala de vários tipos: subtrair ou adicionar um fonema ou sílaba - e.g., dizer "belt" sem /t/ ou "carpet" sem /car/ (Rosner e Simon, 1971), inverter a ordem de sílabas ou fonemas (Alegria e Morais, 1979), dizer uma pequena parte de uma frase, palavra ou sílaba (Fox e Routh, 1980), ou até usar uma regra indutiva baseada na extracção de fonemas (e.g., Zhurova, 1973: por exemplo, "o boneco com a camisola verde é Van, com a camisola preta é Pan", ...). Outro paradigma comum recorre à tarefa de batimento: por exemplo, bate-se três vezes para "butterfly" - segmentação silábica - ou também três vezes para "but" - segmentação fonémica (Liberman, Shankweiler, Fischer e Carter, 1974).

Nas duas últimas décadas tem-se vindo a demonstrar uma firme associação entre o sucesso na aprendizagem da leitura e a consciência da estrutura fonética da fala: por um lado, o desempenho em tarefas de segmentação fonémica melhora sensivelmente pelos seis, sete anos - quando se aprende a ler e escrever; por outro lado, a capacidade de, em idade pré-escolar, categorizar palavras com base na semelhança fonética permite

\* Usaremos, neste contexto, os termos fonético e fonémico um tanto livremente.

predizer em certa medida o sucesso posterior na aprendizagem da leitura e escrita (e.g., Bradley e Bryant, 1983; Lundberg, Olofsson e Wall, 1980; para revisões recentes, veja-se Content, 1984; 1985).

### 1.2.2. Comparação entre adultos iletrados e letrados em tarefas de manipulação da fala

Uma questão crucial, não só em termos teóricos, mas também pelas suas implicações práticas, é a da natureza da correlação entre aprendizagem da leitura e consciência da estrutura fonética da fala\*. É justamente aqui que entram os trabalhos com adultos iletrados: eles permitem clarificar o que é produto do treino específico e da instrução da leitura, e o que depende do desenvolvimento geral, "espontâneo", do sistema cognitivo.

Morais, Cary, Alegria e Bertelson (1979) mostraram que adultos iletrados tinham severas dificuldades, e nalguns casos eram mesmo incapazes, de segmentar explicitamente a fala em unidades subsilábicas (por exemplo, reproduzir sem a primeira consoante os vocábulos pronunciados pelo experimentador); o mesmo não acontecia com indivíduos de idade e condições de vida semelhantes, mas que tinham aprendido a ler já quando adultos. Poder-se-ia atribuir estes resultados a alguma incapacidade global de manipular segmentos da fala, ou até de compreender uma instrução por indução? Não, pois os adultos iletrados foram capazes de desempenhar com sucesso, por exemplo, uma tarefa de inversão de sílabas na sequência de treino indutivo (e.g., Cary e Moraes, 1979; cf. também Bertelson, Moraes, Cary e Alegria, 1986-b). A conclusão de que "the ability to

---

\* O leitor interessado em aprofundar esta questão encontrará contributos importantes num número recente dos "Cahiers de Psychologie Cognitive/European Bulletin of Cognitive Psychology", vol 7(5), Outubro 1987.

deal explicitly with the phonetic units of speech is not acquired spontaneously" (Morais *et al.*, 1979; p. 330) foi corroborada recentemente por novos dados experimentais (Morais, Bertelson, Cary e Alegria, 1986 ; Read Zhang, Nie e Ding, 1986; Scholes, 1987).

Em geral, o pior desempenho dos iletrados é mais evidente em tarefas que envolvam a manipulação de segmentos fonémicos e atenua-se em tarefas que envolvam segmentos silábicos ou detecção de rima. Estes resultados são consistentes com a ideia de que algumas formas de manipulação da fala se desenvolvem até certo ponto espontaneamente (cf. Bradley e Bryant, 1983; Mann, 1986), o que não impede que possam ser melhoradas através de treino específico. Porém, a manipulação de certas classes fonéticas, nomeadamente as oclusivas, requer tal treino (Bertelson, 1986; Bertelson e De Gelder, no prelo; Morais, 1987; Morais *et al.*, 1986-a; Morais, Alegria e Content, 1987-a; b)\*.

A questão da especificidade da competência metafonológica (*vs* uma incapacidade global de análise, por exemplo) torna-se mais clara ao ter presente a comparação entre letrados num sistema logográfico e num sistema alfabético. Read *et al.* (1986) verificaram que letrados chineses que conheciam apenas a escrita logográfica desempenham as tarefas de segmentação da fala como os adultos iletrados observados por Morais *et al.* (1979) ou como crianças em idade pré-escolar (cf. Alegria e Morais, 1979); porém, letrados chineses que conheciam a escrita alfabética "pinyin" têm resultados semelhantes aos dos ex-iletrados portugueses (Morais *et al.*, 1979). Como notou Bertelson (1986), trata-se de um resultado importante, pois permite eliminar a possibilidade de que a diferença entre iletrados e ex-

---

\* O leitor interessado na relação entre alfabetização ("literacy") e os processos cognitivos envolvidos na aquisição da leitura poderá consultar a colectânea editada por Bertelson (1987).

letrados portugueses fosse devida a factores motivacionais, de nível intelectual, ou outros.

Enfim, os resultados revistos mostram claramente que a aquisição da leitura influencia o modo como se executam tarefas de segmentação da fala; esta influência é particularmente marcada se aquela segmentação exige uma referência explícita às unidades linguísticas usadas na escrita.

### 1.3. Consciência da estrutura fonética da fala e processos perceptivos

Vimos o papel desempenhado pela alfabetização na análise explícita das unidades da fala; esta análise deve basear-se nos julgamentos que o auditor faz a propósito das propriedades fonéticas da fala. Trata-se pois de uma competência meta-fonológica. Uma questão importante é se a "phonetic awareness", essa competência meta-fonológica, poderia afectar a própria percepção e compreensão da fala, nomeadamente os processos envolvidos na identificação de fonemas, sílabas ou palavras.

Claro que o estabelecimento de um efeito de alfabetização a nível de análise explícita da fala, claramente pós-perceptivo, não permite inferir esse mesmo efeito a nível perceptivo. Assumimos que este nível perceptivo envolve em larga medida processos inconscientes de elaboração do percepto, enquanto a análise explícita da fala recorre em grande parte a processos conscientes de manipulação do percepto. A crença num isomorfismo entre aqueles dois níveis não é justificável - ou a Psicologia Cognitiva não se distinguiria da introspecção (cf., *e.g.*, Fodor, 1983; Marcel, 1983; 1986, para discussões gerais; e Morais, 1985; Norris e Cutter, 1988, quanto à questão da sílaba ou fonema enquanto unidades de processamento ou alvos de detecção consciente). Todavia, o estabelecimento da diferença entre letrados e iletrados quanto às competências meta-fonológicas, por um lado, e quanto ao

desempenho em tarefas de identificação de palavras, por outro, justifica que seja levantada a hipótese de que a aquisição daquelas competências possa favorecer o desenvolvimento de estratégias de atenção à estrutura segmental dos vocábulos. Mehler, Morton e Jusczyk (1984; p. 101) fizeram uma sugestão na mesma linha, ao considerar que as várias unidades linguísticas poderão ter diferentes papéis na percepção da fala, consoante os períodos etários. Pode parecer pouco plausível que processos conscientes de manipulação do percepto possam de algum modo afectar estratégias mais ou menos automáticas de elaboração do percepto. Convém lembrar a este propósito que os processos conscientes desempenham um papel importante na aquisição e desenvolvimento de muitas actividades automáticas. Além disso, e como foi já sugerido por Bertelson e De Gelder (no prelo, p. 23), a questão das inibições e constrangimentos mútuos entre o processamento perceptivo e as representações conscientes deve ser resolvida empiricamente.

A hipótese segundo a qual o conhecimento da linguagem escrita, através da aquisição de competências meta-fonológicas, promove o desenvolvimento de uma estratégia de atenção à estrutura segmental dos vocábulos será examinada nas três experiências seguintes: uma compara adultos letrados, semiletrados e iletrados numa tarefa de reconhecimento de palavras em audição dicótica; outra compara o desempenho de crianças pré-letradas, crianças e jovens de várias idades e grupos de escolarização na mesma situação; finalmente a terceira aborda o desempenho de adultos letrados, semiletrados e iletrados numa tarefa de identificação de palavras em presença de ruído.

## 2. Identificação de palavras dicóticas - a: Análise dos erros de adultos letrados, semiletrados e iletrados (Experiência III)

### 2.1. A técnica de audição dicótica como meio de abordar os processos perceptivos

Propomo-nos analisar o desempenho na situação de audição dicótica numa perspectiva diferente da neuropsicológica. Nesta, a análise centra-se nos efeitos de lateralidade auditiva (cf. Experiências I e II), que pelo menos em certas condições dão indicação sobre a organização cerebral. A técnica de apresentação dicótica presta-se porém a utilizações bem diversas. Ela foi aliás introduzida por Broadbent (1954) como teste de atenção e memória.

Na medida em que um ouvido recebe uma estimulação enquanto o outro ouvido recebe, simultaneamente, uma informação diferente, a técnica de apresentação dicótica instaura uma situação de competição entre os estímulos. Ora perante competição entre estímulos, podem observar-se erros de identificação que consistem em combinar partes ou características desses estímulos. Uma análise desses erros poderá fornecer indicações valiosas sobre o(s) tipo(s) de informação relevante(s) ao longo do processo perceptivo. O alcance desta análise tornar-se-á mais claro após vermos alguns exemplos.

Tomemos primeiro o trabalho de A. Treisman e colaboradores no domínio da percepção visual. Quando se apresenta ao sujeito, por exemplo, um "X" verde e "O" vermelho, ele pode perceber um "X" vermelho ou um "O" verde (Treisman e Schmidt, 1982). Este fenómeno é conhecido por conjunção ilusória; "ilusória" porque não se deve a uma combinação explícita ou voluntária por parte do sujeito. Fenómenos deste tipo sugerem que o processo perceptivo comporta uma análise precoce das partes ou propriedades componentes do estímulo (para uma revisão recente da evidência empírica e discussão aprofundada, cf. Treisman, no prelo).

No domínio auditivo é conhecido pelo menos um fenómeno análogo às conjunções ilusórias visuais: a fusão de traços fonéticos ou *blending*\* (cf. Treisman e Patterson, 1984; p. 14). Quando os dois ouvidos recebem simultaneamente fonemas diferentes quanto ao lugar de articulação e ao vozeamento, há mais respostas com os traços correctos, mas mal combinados, do que com novos traços: *e.g.*, perante /p/ e /d/, ouve-se mais frequentemente /b/ e /t/ do que /g/ e /k/ (Blumstein, 1974; Cutting, 1976; Studdert-Kennedy e Shankweiler, 1970; Tartter e Blumstein, 1981). Os *blendings* serão discutidos mais à frente, tratemos agora a questão mais geral da interpretação de fenómenos deste tipo.

Segundo Treisman e Patterson (1984, p. 14), "if features can be wrongly recombined, they must at some level have been separately registered as separate entities"; encontra-se a mesma posição em Studdert-Kennedy e Shankweiler (1970) e Cutting (1976). Este pressuposto enforma efectivamente o nome com que se baptizaram estes fenómenos: conjunção ilusória, fusão de traços. É claro que uma análise lógica em termos de "traços", propriedades ou características quer da estimulação, quer da resposta, não implica necessariamente a sua "realidade psicológica", ou seja, que o sistema percepto-cognitivo tenha seguido a lógica do investigador. Estas dificuldades têm vindo a ser explicitadas pela própria Treisman e colaboradores (*e.g.*, Treisman e Souther, 1986; p. 286). Porém, conscientes da armadilha em que consistiria hipostasiar precipitadamente a análise lógica na cabeça do sujeito, é inegável que a comparação das características das respostas em termos de fusão ou separação de propriedades dos estímulos traz indicações importantes sobre o papel dessas propriedades no processo perceptivo propriamente dito. A vantagem crucial desta estratégia é que ela

\* Apesar do termo *blending* ser, em rigor, menos bom porque não específico (pode "misturar-se" ou "fundir-se" muita coisa), usá-lo-emos frequentemente em vez de "fusão de traços fonéticos" por ser mais sucinto.



não se encontra dependente da experiência fenomenológica; teremos assim uma janela sobre os processos automáticos, obrigatórios e inconscientes empregues na elaboração do percepto. Para inferir sobre a relevância dos traços fonéticos na percepção de fala, não se pede ao sujeito que, por exemplo, detecte contrastes duplos; analisam-se antes as suas respostas nessas situações. A primeira estratégia, de detecção, dar-nos-ia possivelmente informações sobre a acessibilidade à consciência duma determinada propriedade; a segunda, de análise das respostas, informa-nos sobre a relevância dessa propriedade na elaboração do percepto.

Em suma, utilizaremos a técnica de audição dicótica como meio de abordar os processos não necessariamente conscientes de percepção da fala. A situação de competição entre os estímulos permitir-nos-á observar dois fenómenos de fusão descritos na literatura, os *blendings* e as fusões fonológicas.

Uma segunda vantagem deste modo de apresentação é que o reconhecimento é mais difícil, pois as palavras aparecem isoladas e sem ajudas do contexto. Isto permitir-nos-á observar vários tipos de erro, conforme se dêem em um ou mais fonemas e sílabas dos estímulos. Possivelmente o facto de as palavras aparecerem desinseridas dum contexto significativo realçará a utilidade da atenção às propriedades do som e à estrutura fonética das palavras.

## 2.2. Fusão de traços fonéticos

Como as fusões fonológicas foram já objecto de tratamento detalhado (cf. Capítulo III), limitar-nos-emos agora a abordar a fusão de traços fonéticos ou *blendings*. Se se constituir um teste de audição dicótica composto por estímulos iniciados por consoantes oclusivas, verifica-se que os sujeitos

(estudantes universitários) dão uma proporção relativamente mais elevada de erros que consistem na combinação errada de traços fonéticos de um e outro estímulo (em vez de combinações com um novo traço). Por exemplo, para o par [ba] - [ka], há mais respostas erradas [pa] e [ga] do que [da] e [ta]; havendo competição entre bilabial sonora, [b], e velar surda, [g], há mais respostas bilabiais surdas [p] e velares sonoras [g], do que respostas com traços fonéticos inexistentes na estimulação, neste caso, as dentais [t] e [d].

A primeira interpretação proposta para este efeito foi que se tratava de um indício de que a fala era analisada em termos de traços fonéticos (cf. Studdert-Kennedy e Shankweiler, 1970; Cutting, 1976; Tartter e Blumstein, 1981), sendo por isso baptizado de *blending* ou combinação de traços fonéticos. Assim, a existência deste erro não só comprovaria a realidade psicológica dos traços fonéticos como, por definição, ocorreria numa etapa específica do processo perceptivo: aquela em que se teria já dado a extracção daquelas entidades, *i.e.*, no modo fonético. Consistente com esta interpretação é o facto de a incidência de *blendings* não ser afectada por ligeiras assincronias entre os dois estímulos, nem pela diferença das respectivas frequências fundamentais (Cutting, 1976). Blumstein e Tartter (1981) relataram também que a ocorrência de *blendings* não depende de diferenças simultâneas quanto à altura tonal e às propriedades do tracto vocal (em estímulos sintéticos).

No entanto, Repp (1977-a) mostrou que interpretar os *blendings* em termos estritamente fonéticos não é correcto; certas características acústicas são também relevantes. A frequência de *blendings* era maior para pares de estímulos que incluíssem [ba] em vez de [pa] e dependia de manipulações intra-categóricas do "Voice Onset Time" e das transições de formantes. Assim, interpretações deste efeito exclusivamente em termos de modo auditivo ou modo fonético não são satisfatórias; parece antes que os *blendings* ocorrem numa altura em que as características acústicas são

ainda relevantes, mas foram já de algum modo categorizadas ou agrupadas segundo critérios fonéticos.

### 2.3. Predições quanto ao tipo de erros nos vários grupos

Optamos por analisar com pormenor as respostas dadas aos pares dissimilares, com pelo menos três fones diferentes. Estes permitem a comparação das intrusões segmentais da primeira consoante ou da primeira vogal com as intrusões de primeira sílaba; e, devido à diferença de maneira de articulação na segunda consoante, há a possibilidade de ocorrência de fusões fonológicas. Em alguns ensaios, as consoantes iniciais diferem em vozeamento e lugar de articulação, pelo que se podem analisar fusões de traços fonéticos. As nossas predições são as seguintes: (i) se o domínio da escrita alfabética torna o sujeito mais atento à estrutura segmental das palavras, então sujeitos letrados terão relativamente mais erros limitados a um só fonema e menos erros globais (silábicos ou em toda a palavra) do que sujeitos iletrados; (ii) se o modo fonético é um processo de percepção da fala universal e automático, então os erros de fusão fonética distribuir-se-ão uniformemente pelos vários grupos.

Relativamente aos sujeitos semiletrados supomos que se situarão algures entre os letrados e iletrados; a sua posição relativamente àqueles dois grupos permitirá justamente verificar até que ponto a atenção aos segmentos da fala em situações de competição de estímulos e não-redundância é mais uma consequência da alfabetização *per se* ou requer antes prática frequente em ler e escrever.

Como o processo subjacente às fusões fonológicas é desconhecido, limitar-nos-emos a observar a sua ocorrência e tipo. Note-se que a mera existência de fusões fonológicas tem implicações importantes para a questão

de se as unidades de percepção da fala são sílabas ou fonemas: dois estímulos CV não poderiam dar azo a uma sílaba CCV se as mais pequenas unidades extraídas no processo perceptivo fossem sílabas.

#### 2.4. Método

Sujeitos, estímulos e procedimento são os mesmos da experiência I, já descrita. Trata-se pois de 18 iletradas, 14 semiletradas e 14 letradas de idades semelhantes. A sua tarefa era de prestar atenção a um só ouvido e repetir uma em cada par de palavras apresentado dicoticamente. O teste era constituído por 48 pares de palavras CVCV, diferindo em pelo menos as três consoantes iniciais. Estas eram sempre oclusivas, e diferiam ou só em vozeamento, ou só em lugar, ou nos dois simultaneamente. A segunda consoante podia ser oclusiva, fricativa ou nasal, tendo-se usado vários tipos de contraste.

#### 2.5. Resultados

Como se viu (Experiência I), a *performance* foi maior nas letradas (64%) do que nas semiletradas (49%) e iletradas (44%), e em todos os grupos houve vantagem do ouvido direito.

##### 2.5.1. Erros locais vs globais

Podem observar-se no Quadro 1 as proporções das várias categorias de erro. Usam-se as proporções de cada tipo de erro relativamente ao total de erros (em vez dos valores absolutos correspondentes) porque o

**QUADRO 1:** Proporções médias dos vários tipos de erro: Intrusões (Int., relato da palavra apresentada no canal não atendido); mudança exclusivamente na: 1ª consoante (C<sub>1</sub>), 1ª vogal (V<sub>1</sub>), 2ª consoante (C<sub>2</sub>), 1ª sílaba (C<sub>1</sub>V<sub>1</sub>), duas consoantes (C<sub>1</sub>C<sub>2</sub>); mudança de pelo menos os três segmentos iniciais, excluindo intrusões (C<sub>1</sub>V<sub>1</sub>C<sub>2</sub>(V<sub>2</sub>)); erros que não respeitam a estrutura CVCV dos estímulos (Estruturais). As outras categorias (erros exclusivamente na 2ª vogal, ou em dois ou três segmentos que incluam a 2ª vogal) foram agrupados na categoria Outros. As proporções foram calculadas para cada sujeito relativamente ao total de erros por ele exibido.

	ILETRADAS	SEMILETRADAS	LETRADAS
Intrusão	7.1	9.3	3.8
C <sub>1</sub>	23.8	29.1	38.7
V <sub>1</sub>	5.7	6.3	5.4
C <sub>2</sub>	7.6	7.5	8.8
C <sub>1</sub> V <sub>1</sub>	10.9	10.4	7.5
C <sub>1</sub> C <sub>2</sub>	6.4	8.7	6.6
C <sub>1</sub> V <sub>1</sub> C <sub>2</sub> (V <sub>2</sub> )	12.0	8.1	6.7
Estrutural	6.9	10.1	10.0
Outros	19.6	10.5	12.5

objectivo é avaliar a importância relativa das várias fontes de erro; esta é também a maneira que permite fazer comparações entre grupos com diferentes níveis de *performance*. Subjacente a esta opção está a ideia de que as categorias de erro provavelmente reflectirão diferentes estratégias ou mecanismos de reconhecimento; a análise da incidência relativa de diferentes erros elucidam-nos sobre até que ponto diferentes processos ou mecanismos determinaram a resposta do sujeito.

A inspecção do Quadro 1 sugere que as letradas dão relativamente mais erros restringidos à primeira consoante ( $C_1$ ) do que as iletradas; estas por sua vez dão mais erros na primeira sílaba ( $C_1V_1$ ) e em toda a palavra ( $C_1V_1C_2(V_2)$ ).

Uma análise de variância às proporções de erros  $C_1$ , depois de submetidos à transformação arcoseno, revelou um efeito significativo do grupo ( $F(2,43)=3,87, p<.05$ ). Testes Tukey indicaram que só a diferença entre letradas *vs* iletradas era significativa. Não havia efeito de ouvido ( $F < 1$ ), apesar da ligeira tendência em estes erros serem mais frequentes no ouvido direito do que no esquerdo (considerando os três grupos em conjunto, .321 *vs* .287, respectivamente). Isto acontecia, *grosso modo*, igualmente para os três grupos; não se observou pois nenhuma interacção entre grupo e ouvido ( $F < 1$ ). Entre os erros de tipo  $C_1$ , um era particularmente frequente: responder com a consoante inicial da palavra apresentada no canal não-atendido. Chamemos-lhes intrusões de consoante inicial; elas foram responsáveis por 9%, 13% e 16% de todos os erros nas iletradas, semiletradas e letradas, respectivamente. Porém, entre os erros de tipo  $C_1$ , estas intrusões eram igualmente frequentes nos três grupos: .39, .46 e .40, respectivamente.

A correlação entre a proporção de erros  $C_1$  e o número de respostas correctas foi positiva em todos os grupos: não significativa nas letradas (.28), perto da significância a  $p = .05$  nas iletradas (.44) e significativa ( $p < .01$ ) nas

semiletradas (.75). Como todas as correlações foram positivas, parece haver uma associação entre o nível de *performance* e a proporção de erros  $C_1$ . Por esta razão é importante esclarecer se o efeito da escolarização na proporção de erros  $C_1$  se mantem com subgrupos de níveis de *performance* semelhantes.

Foi possível constituir dois subgrupos de iletradas e letradas com a mesma percentagem média de respostas correctas: foram as dez melhores letradas e as sete piores iletradas, cujo desempenho foi de 53%. Um teste *t* unicaudal às proporções de erros  $C_1$  nestes subgrupos mostrou que se mantem a sua incidência significativamente maior nas letradas ( $t_{(15)}=1.75$ ,  $p=.05$ ). Assim, mesmo se o nível de *performance* é semelhante, há um efeito da alfabetização na proporção de erros  $C_1$ . Por outro lado, verifica-se que o nível de *performance* é ele próprio afectado pelo facto de se saber ou não ler e escrever, independentemente da proporção de erros  $C_1$ . Comparamos subgrupos de iletradas e letradas com aproximadamente o mesmo nível de proporção de erros  $C_1$  (.286 e .287, respectivamente); eles eram compostos pelas 13 iletradas com a maior, e as 7 letradas com a menor, incidência de erros  $C_1$ . A *performance* das letradas era significativamente superior à das iletradas ( $t_{(18)}=2.35$ ,  $p < .025$ ).

Efectuou-se também uma análise de variância às proporções de erros globais, ou seja, o conjunto de erros na primeira sílaba,\*  $C_1V_1$ , e em pelo menos três fonemas,  $C_1V_1C_2(CV_2)$ , excluindo as intrusões da palavra apresentada no canal não atendido (provavelmente estas são devidas a uma

\* Chamaremos erros silábicos aos erros  $C_1V_1$ ; não são considerados os erros de segunda sílaba,  $C_2V_2$ , porque (i) a construção do teste, permitindo por vezes que a segunda vogal seja idêntica nas duas palavras, impede de interpretar claramente os erros  $C_2V_2$  como propriamente silábicos; (ii) os erros  $C_2V_2$  foram muito raros, apenas 1.2% nos universitários, liceais, primários e pré-primários.

insuficiência de atenção selectiva). Esta análise revelou um efeito do grupo,  $F_{(2, 43)}=3.71$ ,  $p < .05$ . Testes Tukey mostraram que só a comparação entre letradas e iletradas era significativa. Houve também um efeito de ouvido,  $F_{(1, 43)}=4.41$ ,  $p < .05$ , reflectindo o facto de a proporção de erros globais ser maior no ouvido esquerdo (.197) do que no direito (.140). Todos os grupos exibiam esta tendência; assim, não se observou nenhuma interacção entre ouvido e grupo ( $F < 1$ ). Ao contrário do que acontecia com os erros  $C_1$ , a correlação entre a percentagem de respostas correctas e a proporção de erros globais foi negativa em todos os grupos: -.45 (não significativa), -.68 e -.67 ( $p < .01$ ) nas letradas, semiletradas e iletradas, respectivamente.

Poderia ser que os resultados até agora revistos estivessem de algum modo associados à tendência ou relutância por parte dos sujeitos em dar respostas de tipo não-palavra. Por exemplo, se as iletradas tivessem um viés contra repetir "coisas esquisitas", e os erros de tipo  $C_1$  fossem frequentemente não-palavras, então as diferenças intergrupo observadas dever-se-iam a um mero viés de resposta. Contrariamente àquela suposição, a percentagem de respostas não-palavra relativamente ao total de erros era maior nas iletradas (26%) do que nas semiletradas (23%) e letradas (16%). Uma análise de variância àquelas percentagens, previamente submetidas à transformação arcoseno, revelou um efeito significativo do grupo,  $F_{(2,43)}=3.45$ ,  $p < .05$ . Testes Tukey indicaram que só as letradas e iletradas tinham resultados significativamente diferentes. Este efeito de grupo é provavelmente devido a um viés de resposta trivial: como todos os estímulos eram palavras, as letradas, que conseguiram identificar correctamente muito mais *items*, tenderiam também a produzir palavras perante situações dúbias.

## 2.5.2. Tipos de contraste

### Contrastes de vozeamento

Relativamente aos erros de primeira consoante\* nos pares que diferiam apenas quanto ao vozeamento, obteve-se um efeito certamente atribuível às características fonéticas da língua portuguesa: havia cerca de quatro vezes mais erros de vozeamento quando o estímulo atendido era surdo (/p, t, k/) do que quando era sonoro (/b, d, g/). No português, tal como no espanhol, as consoantes sonoras são pré-vozeadas, *i. e.*, a vibração das cordas vocais antecede cerca de 0.1 segundos a libertação da oclusão. Assim, aqueles erros poderão ser devidos a um erro de localização da fonte sonora onde ocorria o pré-vozeamento\*\*. Tratar-se-á, talvez, de um efeito análogo ao descoberto por Repp (1975): quando se apresenta a um ouvido alvos CV (ingleses) e ao outro vogais isoladas como máscaras, observa-se uma tendência em dar respostas sonoras se o início da vogal precede o início da sílaba. A relação temporal inter-aural parece pois competir com os índices temporais de vozeamento do alvo. Como as sonoras portuguesas são pré-vozeadas, elas teriam um efeito semelhante ao das vogais usada por Repp (*ibidem*).

### Contraste duplo

Analisaram-se os erros limitados à primeira consoante nos pares de contraste duplo (vozeamento e lugar de articulação) com vista à detecção de fusões de traços fonéticos. Tendo por exemplo o par [b] - [k], a possibilidade de respostas erradas tipo *blending*, [p] ou [g], é a mesma que de

\* Consideramos nesta análise apenas os erros C<sub>1</sub>, para evitar a confusão com outros factores, como dependência de contexto, regras fono-tácticas, etc.

\*\* Usaremos o termo "pré-vozeamento" para nos referirmos a "Voicing lead".

outras oclusivas anómalas, [d] ou [t]. O que se verifica é que são mais frequentes os *blendings* do que as respostas oclusivas anómalas, aproximadamente 4 vezes mais nas iletradas e letradas, e 2.5 vezes mais, nas semiletradas (cf. Quadro 2). A razão entre o número de *blendings* e a soma de *blendings* e anómalas é semelhante nos três grupos: .80 nas letradas, .72 nas semiletradas e .81 nas iletradas. Como os *blendings* sonoros poderão resultar de uma tendência a aceitar o vozeamento independentemente de ele surgir no canal atendido ou não-atendido, vale a pena observar separadamente os *blendings* surdos. Como se vê no Quadro 2, mantem-se claramente a sua preponderância face às respostas anómalas, nos três grupos.

A correlação entre o número de respostas correctas, por um lado, e a diferença entre *blendings* e anómalas, por outro, é baixa e não significativa em qualquer dos grupos: .18, .14 e -.01 nas letradas, semiletradas e iletradas, respectivamente.

### 2.5.3. Fusões fonológicas

Considerou-se fusão fonológica qualquer resposta que englobasse os cinco fonemas que tivessem estado presentes na estimulação dicótica; aceitaram-se também transformações para segmentos muito parecidos, nomeadamente as requeridas por regras fono-tácticas do português (como a transformação do [ʀ] em [r]). Houve um total de 39 fusões fonológicas, 11 nas iletradas, 18 nas semiletradas e 10 nas letradas. Eram todas de tipo CVCCV, exceptuando três. Destas, em dois casos a consoante adicional foi transformada na semivogal [w]; no outro, a estrutura foi CCVCV. Cerca de 20% eram não-palavras. Um resumo destas observações encontra-se no Quadro 3.

## 2.6. Discussão

Numa tarefa de reconhecimento de palavras faladas apresentadas dicoticamente, sujeitos letrados dão mais respostas correctas do que semiletrados e iletrados. A análise das respostas erradas sugere que, apesar de em circunstâncias normais tanto letrados como iletrados perceberem a fala, pelo menos em condições difíceis eles não o fazem exactamente da mesma maneira. Os resultados desta experiência trazem alguma luz sobre dois tipos de processos envolvidos no reconhecimento da fala: processos pré-atencionais e automáticos de descodificação fonética e processos atencionais opcionais.

Os processos pré-atencionais de percepção da fala não parecem ter contribuído decisivamente para as variações de *performance* observadas. Não só não houve correlação entre exactidão e a estimativa de *blendings*, como a proporção de *blendings* era semelhante em letradas e iletradas. Este último achado é particularmente importante. Trata-se da replicação de um fenómeno clássico na literatura da percepção da fala, estendido a sujeitos menos habituais quer pela sua idade (mais avançada que a dos estudantes universitários), quer pela sua instrução. Centrando-nos agora na questão da hipotética influência da alfabetização no processamento da fala, o facto da incidência de *blendings* ser praticamente independente do grupo confirma a plausível suposição de que processos básicos de extracção de informação fonética do *input* acústico não dependem de factores experienciais como a escolarização. Tal como os letrados frequentadores típicos dos laboratórios psicológicos, também iletrados recorrem a processos de descodificação fonética na percepção da fala.

Vejamos agora os processos atencionais. Uma das interpretações possíveis sobre o pior desempenho das iletradas e semiletradas é a de uma incapacidade ou dificuldade de controlo espacial da atenção: estes sujeitos

teriam mais dificuldade do que os letrados em orientar e manter a sua atenção focada neste ou naquele ouvido. Os resultados obtidos não apoiam tal interpretação: apesar de os sujeitos com pior *performance* darem também mais intrusões, este acréscimo bruto de intrusões é pequeno relativamente ao decréscimo de respostas correctas.

Detenhamo-nos nas diferenças intergrupo quanto às proporções de erros na primeira consoante e os globais. O nível de *performance* está positivamente correlacionado com os erros  $C_1$  e negativamente com os erros globais. Além disso, para níveis de desempenho semelhante, as letradas têm mais erros  $C_1$  e menos erros globais do que as iletradas. Há pois uma associação entre estes tipos de erros e o nível de desempenho, por um lado, e a escolarização, por outro.

O que propomos, com J. Morais, R. Kolinsky e A. Content, é que a elevada proporção de erros num segmento único reflectirá um processo de atenção à estrutura segmental da fala. Este modo de atenção seria particularmente útil em situações de pouca clareza e falta de redundância do sinal. O seu desenvolvimento poderia ser favorecido, ou mesmo desencadeado, pela consciência da estrutura fonética da fala. Sem conseguir elaborar representações conscientes dos segmentos fonémicos (iletrados) ou apenas com falta de prática na análise segmental da fala (semiletrados), iletrados e semiletrados não atenderiam à estrutura segmental do estímulo.

Que observações nos permitem inferir sobre um tal processo de atenção à estrutura segmental das palavras? A ocorrência de erros limitados a um só segmento não indica por si só que se exerceu atenção segmental. Mesmo prestando atenção ao estímulo integral poderá haver erros num só fonema, nomeadamente mais erros em consoantes do que vogais (se as vogais forem mais facilmente discrimináveis do que as consoantes). Um aumento de esforço atencional pode ser acompanhado por um acréscimo de

erros num segmento único. Todavia, para um nível de exactidão constante, a existência de mais ou menos erros segmentais indica provavelmente a utilização de uma estratégia especial. Nesta experiência verificámos que as letradas tinham uma proporção de erros na primeira consoante superior à das iletradas, mesmo quando o seu nível de desempenho era idêntico. Isto indica que a alfabetização exerce um efeito específico a nível da compreensão da palavra falada em situação de competição dicótica. Tendo em conta o que se sabe sobre a relação entre alfabetização e consciência da fala como sequência de fones, é plausível que a estratégia específica indutora de mais erros segmentais nas letradas esteja associada a uma atenção particular à estrutura fonética das palavras. Designá-la-emos por atenção segmental.

Verificou-se também que, mesmo com idêntica proporção de erros na primeira consoante, as letradas mantinham um nível de desempenho superior ao das iletradas. Este resultado sugere que a atenção à estrutura segmental da fala, apesar de certamente contribuir para melhorar a *performance*, não é o único factor a determiná-la. Independentemente do alvo da atenção ser um fonema, uma sílaba, ou uma palavra, o facto de a atenção ser mais ou menos intensa, recorrendo a mais ou menos "recursos" (no sentido de "resources", cf. Kahneman, 1973; Norman e Brobow, 1975; Wickens, 1984) afecta o nível de *performance*. O melhor desempenho das letradas deve-se provavelmente não só à sua capacidade de atender à estrutura segmental da fala, como a outros factores, que convirá esclarecer em futuras investigações.

As semiletradas enquanto grupo tiveram um comportamento mais próximo do das iletradas do que das letradas. Isto pode atribuir-se a um, ou ambos, dos factores seguintes. Primeiro, uma mestria rudimentar da leitura e escrita pode não assegurar o desenvolvimento completo de mecanismos atencionais centrados nos segmentos da fala. Para esclarecer

este aspecto, é importante realizar estudos semelhantes em crianças em vários níveis de escolarização. Segundo, a manutenção destes mecanismos pode depender de uma prática regular da leitura e/ou escrita. Para elucidar esta questão, valeria a pena avaliar mais detalhadamente as competências dos sujeitos semiletrados, quer na própria mestria da leitura e escrita, quer nas competências meta-fonológicas.

Outro aspecto a salientar tem que ver com os efeitos de lateralidade. Já tínhamos observado e discutido este assunto, considerando globalmente a quantidade de erros: não havia interacção entre grupo e ouvido. É importante acentuar que também para as diferenças de ouvido na proporção de erros, segmentais ou globais, não se verificou qualquer interacção com o grupo. Isto permite-nos afirmar que as diferenças grupais discutidas não estão contaminadas por factores de lateralidade.

Uma observação pertinente é que a proporção de erros globais foi maior no ouvido esquerdo do que no direito. À primeira vista, trata-se de um achado interessante, consistente com a hipótese de que o processamento típico do hemisfério direito é menos analítico e mais holístico que o do hemisfério esquerdo (cf. Bradshaw e Nettleton, 1981; Morais, 1982). Segundo esta óptica, os erros globais resultariam provavelmente de tentativas frustradas de emparelhamento de propriedades acústicas globais do *input* com representações lexicais do mesmo tipo, uma estratégia "típica" do hemisfério direito. Porém, outra interpretação possível é que os erros globais foram relativamente mais frequentes no ouvido esquerdo, não porque ele reflecta o processamento do hemisfério direito, mas porque os sujeitos teriam mais dificuldade em prestar-lhe atenção e a *performance* é inferior. Isto é, a nível de observáveis, a maior proporção de erros globais no ouvido esquerdo está associada ao pior desempenho nesse ouvido; e, na falta de outras provas, parece mais parcimonioso não atribuir este padrão específico de

erros, associado a um pior desempenho, a um tipo de processamento hemisférico especial. Note-se que não houve diferenças de ouvido na proporção de erros segmentais, o que seria predito por um modelo hemisférico perfilhando a dicotomia analítico/holístico.

Detenhamo-nos agora nas fusões fonológicas. Elas foram observadas nos três grupos. Assim, não parecem depender da alfabetização, nem da consciência fonológica. Verificou-se todavia um maior número de fusões fonológicas nas semiletradas. Isto pode ser meramente ocasional; não tendo este trabalho sido elaborado para examinar esta questão, valerá a pena retomá-la em investigações posteriores. Por ora, ficamo-nos por mais alguns comentários descritivos. Tal como tinha já sido observado noutros estudos, é mais fácil haver fusão (i) entre oclusivas e líquidas do que entre fricativas e líquidas; (ii) entre oclusivas e /l/, sem transformações fonémicas, do que entre oclusivas e /r/; (iii) há substituições de /r/ por /l/, mas não o contrário (Cutting, 1975). Diferentemente de investigações expressamente desenhadas para examinar fusões fonológicas, que usam em regra monossílabos ou polissílabos contrastando posições iniciais ou finais, neste trabalho observou-se que as fusões fonológicas não respeitam as fronteiras silábicas: a consoante da sílaba final de um estímulo pode ser fusionada com a sílaba inicial do outro estímulo, produzindo uma sílaba CVC (*e.g.*, "bode+pala" → "balde"). Para os partidários da sílaba como unidade de processamento da fala, trata-se de um fenómeno dificilmente explicável.

Em suma, este trabalho sugere que processos pré-atentivos de análise fonética não são influenciados pela alfabetização; esta contribui porém para desenvolver processos de atenção à estrutura segmental da fala. Estes processos terão certamente um papel insignificante em situações onde a mensagem acústica seja clara, redundante e integrada num contexto pragmático, mas serão decisivos em situações de mais difícil inteligibilidade.

### 3. Identificação de palavras dicóticas - b: Análise dos erros de crianças pré-letradas, crianças e jovens letrados (Experiência IV)

#### 3.1. Objectivos

A experiência que se segue visa testar a hipótese sugerida pelo trabalho anterior: que os processos de análise fonética não dependem da alfabetização e que a atenção à estrutura segmental da fala depende do conhecimento da escrita. A mesma situação de audição dicótica será utilizada com crianças pré-letradas e letradas, e jovens liceais e universitários. Poderemos assim observar um eventual efeito do nível da escolarização para além do efeito da alfabetização propriamente dita.

Se a atenção à estrutura segmental da fala for desencadeada pela aprendizagem da leitura e da escrita, deveríamos encontrar uma incidência menor de erros segmentais nas crianças pré-letradas do que nas crianças e jovens letrados. Além disso, se o conhecimento com a escrita for a variável determinante, a incidência de erros segmentais deveria ser semelhante nas crianças pré-letradas e nas adultas iletradas. Caso a idade ou outras experiências, nomeadamente a experiência com a fala, desempenhem um papel importante, dever-se-ia encontrar uma dissociação entre crianças pré-letradas e adultos iletrados.

Quanto aos processos de análise fonética, esperaríamos o seguinte: na medida em que são pré-atencionais e automáticos, a preponderância de *blendings* face às oclusivas anómalas deveria observar-se independentemente da alfabetização e do nível de escolarização; se observarmos um padrão semelhante nos sujeitos infantis e nos jovens, teremos encontrado evidência convergente em favor da extração precoce e obrigatória de traços fonéticos da mensagem falada.

**QUADRO 1:** Características etárias dos grupos observados: média e desvio-padrão.\*

GRUPO	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
IDADE MÉDIA	22.2	15.3	7.8	5.9
(d. p.)	(1.8)	(.8)	(.5)	(.5)

\* No caso dos grupos infantis, as idades médias foram calculadas com a aproximação de mês; no caso dos dois grupos de jovens, usou-se apenas o ano.

### 3.2. Método

#### 3.2.1. Sujeitos

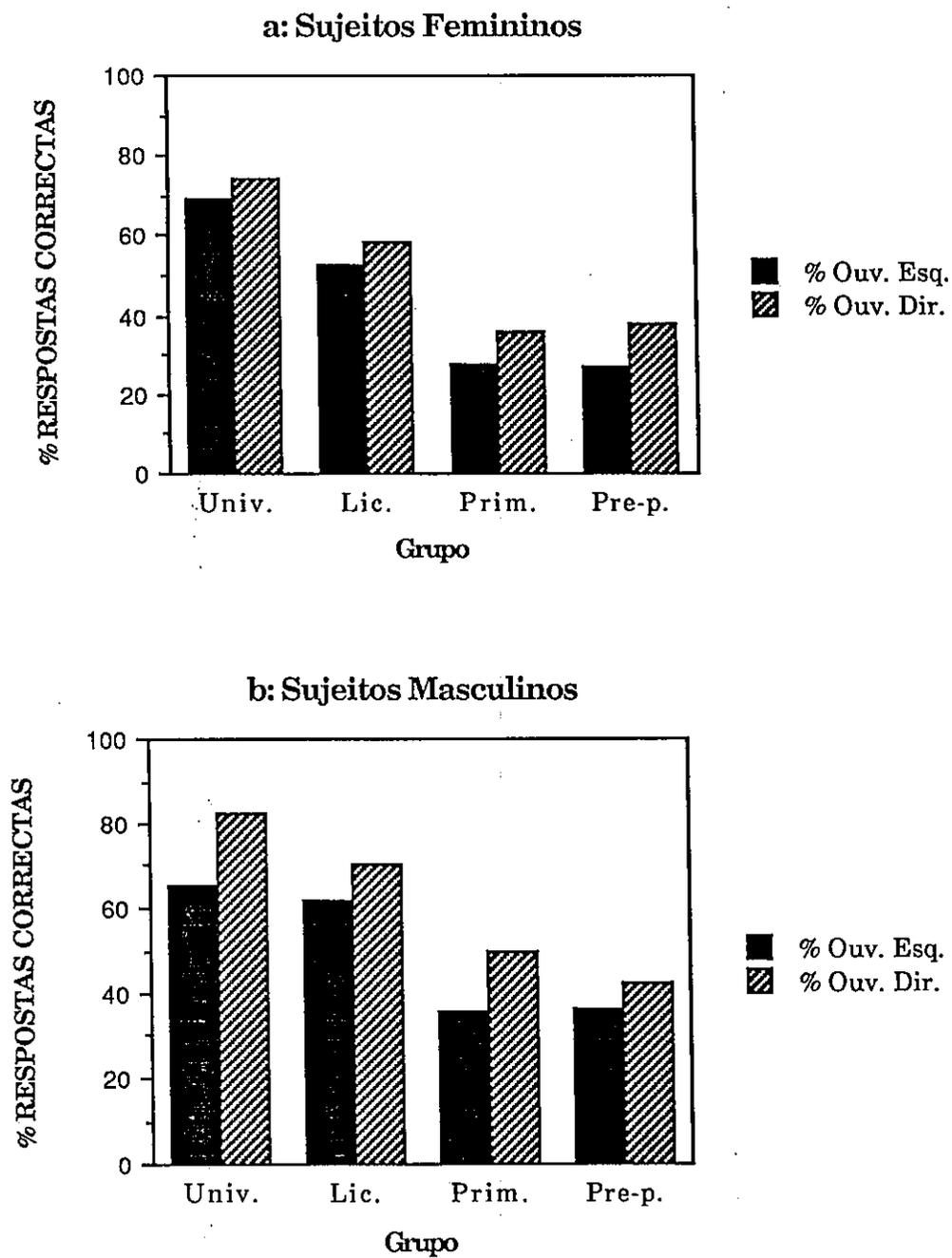
Participaram nesta experiência 96 sujeitos dextros, metade do sexo masculino e metade do sexo feminino, equitativamente distribuídos em quatro grupos etários e de grau de escolarização: jovens universitários (Univ.), adolescentes liceais (Lic.), crianças da escola primária (Prim.) e crianças pré-letradas frequentando o jardim infantil (Pré-prim.). Cada grupo é composto por 12 sujeitos do sexo masculino e 12 do sexo feminino de idades equivalentes (cf. Quadro 1).

Todos os sujeitos letrados completaram o teste de fluência de leitura (cf. Anexo L.1). Este teste foi adaptado para as crianças da primária, tendo-se usado apenas as palavras mais simples da primeira coluna (maioritariamente dissílabos) em caligrafia manual. Todos os alunos da primária que participaram nesta experiência tinham sido previamente qualificados pela professora como bons leitores.

Nenhum dos sujeitos tinha uma história conhecida de problemas auditivos ou neurológicos. Na sessão experimental perguntava-se-lhes se sentiam diferenças de volume nos dois ouvidos; se houvesse queixas de maior dificuldade em perceber num ou noutro ouvido, o sujeito era substituído.

#### 3.2.2. Estímulos e procedimento

Estímulos e procedimento foram idênticos aos da Experiência III. Trata-se pois de 48 pares de palavras CVCV diferindo em pelo menos os três segmentos iniciais, apresentados aos sujeitos através de audição dicótica. O



**FIGURA 1:** Percentagem média de respostas correctas por ouvido e grupo

procedimento foi ajustado para as crianças pré-primárias: fazia-se uma pausa a meio do teste e, se fosse necessário para manter a atenção da criança, faziam-se ao longo do teste breves interrupções acompanhadas por algum incentivo (por exemplo, "vamos descobrir o que vem deste lado", etc).

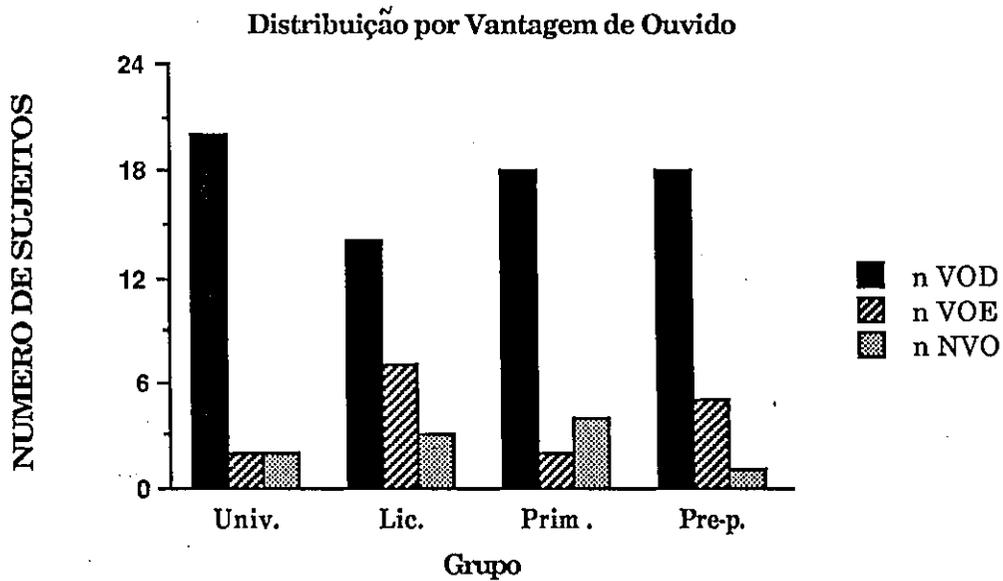
### 3.3. Resultados e discussão

#### 3.3.1. Análise geral do desempenho

O desempenho - palavras correctamente identificadas no ouvido a que se presta atenção - foi melhor nos universitários (média de 73%, entre um mínimo de 50% e um máximo de 88%, desvio-padrão de 10%) e adolescentes (média de 61%, entre 23 e 96%, d.p.=18), e pior nos primários (média de 37%, entre 13 e 67%, d.p.=14) e pré-primários (média de 36%, entre 13 e 81%, d.p.=15).

A figura 1 documenta os resultados com maior detalhe, mostrando a exactidão consoante grupo e ouvido, por sexo. Está bem patente a superioridade do ouvido direito em ambos os sexos. Uma análise de variância às respostas correctas confirma o efeito de ouvido,  $F_{(1,88)}=38.11$ ,  $p<.001$  e de grupo,  $F_{(3,88)}=37.89$ ,  $p<.0001$ . Também o sexo exerce um efeito significativo,  $F_{(1,88)}=6.71$ ,  $p=.01$ : os sujeitos masculinos têm um nível de desempenho superior ao dos femininos, respectivamente 56% e 48% de respostas correctas; os desvios-padrão correspondentes são 21 e 22. Nenhuma interacção atinge a significância\*. Quanto ao efeito do grupo, comparações dois a dois *a posteriori* através do teste Tukey ( $\alpha=.05$ ) mostraram que não existem diferenças significativas entre universitários e liceais, nem entre pré-primários e primários. Os restantes contrastes atingem a significância, nomeadamente a diferença entre primários e liceais.

\* Para pormenores, cf. Anexo EO.1.



**FIGURA 2:** Número de sujeitos exibindo uma superioridade de ouvido direito (VOD), ou do esquerdo (VOE), ou nenhuma vantagem de ouvido (NVO), por grupo. O número máximo de sujeitos em cada grupo é 24.

**QUADRO 2:** Proporções dos vários tipos de erro (percentagens médias): Intrusões (Int., relato da palavra apresentada no canal não atendido); mudança exclusivamente na: 1ª consoante ( $C_1$ ), 1ª vogal ( $V_1$ ), 2ª consoante ( $C_2$ ), 1ª sílaba ( $C_1V_1$ ), duas consoantes ( $C_1C_2$ ); mudança de pelo menos os três segmentos iniciais excluindo intrusões,  $C_1V_1C_2(V_2)$ ; erros que não respeitam a estrutura CVCV dos estímulos (Estruturais). Os outros erros (exclusivamente na 2ª vogal), ou em dois ou três segmentos que incluem a 2ª vogal) foram agrupados na categoria "Outros". As proporções foram calculadas para cada sujeito relativamente ao total de erros por ele exibido.

	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
Intrusão	2.3	4.9	10.2	24.9
$C_1$	48.5	41.1	26.8	14.8
$V_1$	7.8	5.2	4.5	4.7
$C_2$	5.9	4.4	6.9	4.6
$C_1V_1$	5.3	11.7	10.7	7.1
$C_1C_2$	7.9	7.4	9.7	6.8
$C_1V_1C_2(V_2)$	2.0	5.2	10.1	17.3
Estrutural	7.1	8.1	6.5	6.0
Outros	13.1	12.0	14.5	13.8

Na figura 2 pode observar-se a distribuição dos sujeitos conforme o tipo de assimetria auditiva: cerca de 73% tiveram uma vantagem do ouvido direito, e apenas 17% a vantagem oposta do ouvido esquerdo.

### 3.3.2. Tipo de erros

As proporções das várias categorias de erro aparecem no Quadro 2. Usam-se as proporções de cada tipo de erro relativamente ao total (em vez dos valores absolutos correspondentes) porque o objectivo é avaliar a importância relativa das várias fontes de erro; esta é também a maneira que permite fazer comparações entre grupos com diferentes níveis de *performance*.

O exame daquele quadro sugere que a incidência dos vários erros varia segundo o grupo. As intrusões são claramente mais frequentes nos sujeitos mais jovens, indicando assim que os mecanismos de controlo da atenção vão sofrendo uma melhoria com o aumento da idade. Contrariamente às intrusões, os erros na primeira consoante ( $C_1$ ) são mais comuns nos sujeitos menos jovens. Uma análise de variância às proporções de erro  $C_1$  (depois de submetidas à transformação arcoseno) confirma o efeito do grupo, altamente significativo:  $F_{(3,88)}=15.85$ ,  $p < .0001$ . Nenhum dos outros factores, sexo ou ouvido, foi significativo, nem nenhuma interacção\*. Testes Tukey ( $\alpha=.05$ ) revelaram que a única diferença não-significativa é entre liceais e universitários; assim, os pré-primários têm significativamente menos erros  $C_1$  do que os primários, e estes menos que os liceais e universitários.

Para avaliar eventuais diferenças intergrupais noutros tipos de erro, consideremos primeiro os erros globais. Como no estudo anterior, esta

\* Para pormenores, cf. Anexo EO.2.

categoria agrupa os erros de primeira sílaba,  $C_1V_1$ , e os erros em pelo menos os três primeiros segmentos. Os dados mostram uma diminuição mais ou menos gradual deste tipo de erros com o aumento da idade e escolarização. Uma análise de variância a estes valores (depois de transformados) revelou um efeito significativo, o de grupo,  $F_{(3,88)}=23.28$ ,  $p<.0001$ , e um efeito quase significativo do sexo,  $F_{(1,88)}=3.88$ ,  $p=.051$ . Nem o ouvido, nem nenhuma interação atingiram significância\*. Quanto ao efeito do sexo, verifica-se que os sujeitos femininos têm uma proporção de erros globais ligeiramente superior à dos sujeitos masculinos (19.0% vs 15.7%, respectivamente). Quanto ao efeito do grupo, foram feitas comparações *a posteriori*, através de testes Tukey para  $\alpha=.05$ . Elas mostraram que os universitários têm uma proporção de erros globais significativamente inferior a todos os outros grupos (cerca de 7% vs 17% nos liceais, 21% nos primários e 24% nos pré-primários); das restantes comparações, só a diferença entre pré-primários e liceais é significativa.

As diferenças que observámos quanto à distribuição dos erros globais nos vários grupos dever-se-ão essencialmente à progressão dos erros  $C_1V_1C_2(V_2)$  dos pré-primários até aos universitários; para clarificar este ponto, analisámos separadamente este tipo de erros. Uma análise de variância análoga às anteriores replicou a inexistência de interações e o efeito do grupo e sexo: respectivamente,  $F_{(3,88)}=38.47$ ,  $p=.0001$  e  $F_{(1,88)}=6.08$ ,  $p<.0155^{**}$ . Os sujeitos femininos dão em média 9.8% de erros deste tipo e os masculinos 7.5%. O efeito do ouvido esteve próximo da significância,  $F_{(1,88)}=3.87$ ,  $p=.0521$ : os erros  $C_1V_1C_2(V_2)$  eram mais frequentes quando se atendia ao ouvido esquerdo (9.4%, d.p. = 10.2) do que ao direito (7.5%, d.p. = 9.4). Quanto às diferenças grupais, testes Tukey ( $\alpha=.05$ ) revelaram que nem as diferenças entre universitários e liceais (respectivamente, 2% e 5%), nem entre

---

\* Cf. Anexo EO.3.

\*\* Cf. Anexo EO.4.

**QUADRO 3:** Correlação ( $r$  de Pearson) entre exactidão e proporção de erros de tipo global (% GI) e  $C_1$  (%  $C_1$ ) por grupo. Em baixo do valor de  $r$ , a respectiva probabilidade por hipótese nula, com  $n = 24$ .

GRUPO	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
% GI	-.49	-.53	-.35	-.59
$p$	.01	.007	.09	.002
% $C_1$	.36	.38	.57	.08
$p$	.08	.06	.003	.68

primários e pré-primários (respectivamente, 10% e 17%) atingem significância; todas as outras comparações são significativas, nomeadamente entre primários e liceais.

### 3.3.3. Comparação de subgrupos com o mesmo nível de desempenho

Verificámos que a incidência de erros de tipo segmental ( $C_1$ ) e global (e.g.,  $C_1V_1C_2(V_2)$ ) varia consoante a idade, escolarização e até, nalguns casos, com o sexo. Tínhamos também visto que os grupos de diferentes idades ou grau de escolarização, bem como em geral sujeitos masculinos e femininos, desempenham a tarefa a diferentes níveis de eficácia. A questão que agora se coloca é de esclarecer até que ponto as diferenças encontradas quanto à preponderância deste ou daquele tipo de erros não estarão associadas ao melhor ou pior desempenho dos sujeitos. Formulando a questão mais radicalmente, poder-se-á atribuir os efeitos observados a diferenças no nível geral de *performance*? O primeiro passo no exame desta questão é calcular as correlações entre a percentagem de respostas correctas e a proporção de erros  $C_1$  e globais\*.

Como se vê no Quadro 3, existe uma relação inversa entre a proporção de erros globais e o nível de desempenho: nos quatro grupos há uma correlação negativa, que só não atinge a significância nos sujeitos primários. Assim, é possível que os efeitos significativos do grupo e do sexo na proporção de erros globais estejam associados, até confundidos, com o melhor ou pior desempenho dos sujeitos.

A relação entre *performance* e proporção de erros  $C_1$  revela um quadro diferente. Quanto melhor é a exactidão, mais os erros tendem a con-

\* Considerando, em vez dos erros globais ( $C_1V_1 + C_1V_1C_2(V_2)$ ), apenas os erros  $C_1V_1C_2(V_2)$ , os resultados foram semelhantes.

**QUADRO 4:** Características de desempenho (exactidão, % R.C., proporção de erros C<sub>1</sub>, % C<sub>1</sub>, e globais, % GI) de subgrupos de sujeitos pré-primários, primários e liceais com nível de exactidão semelhante\*. Apresenta-se o desvio-padrão entre parêntesis.

	%R.C.	%C <sub>1</sub>	%GI
Pré-prim. n=14	41.8 (5.6)	14.7 (7.6)	21.7 (6.7)
Prim. n=14	41.7 (7.8)	30.2 (7.6)	17.4 (7.6)
Lic. n=10	42.9 (11.4)	32.3 (14.4)	22.0 (9.0)

\* Foi possível emparelhar 14 sujeitos pré-primários quanto à exactidão; só 10 sujeitos do grupo liceal tinham um nível de performance equivalente.

centrar-se na primeira consoante, em todos os grupos com excepção dos pré-primários. Nestes, não se observa praticamente nenhuma relação entre as duas variáveis; nos universitários e liceais as correlações são positivas e aproximam-se da significância; nos primários, a correlação, positiva, é mais forte e significativa. Assim, é possível que algumas das diferenças intergrupais quanto à incidência de erros de tipo  $C_1$  sejam atribuíveis ao nível de exactidão. Para elucidar este ponto, comparemos subgrupos de diferentes idades e escolarização, mas emparelhados por nível de desempenho.

Como o nível de exactidão é bastante díspare nos universitários e pré-primários, não é possível fazer comparações directas entre sujeitos destes grupos. Felizmente, é possível encontrar entre os sujeitos liceais um subgrupo com uma exactidão semelhante à dos melhores sujeitos infantis, de cerca de 42% (cf. Quadro 4). A inspecção deste quadro revela que, mantendo constante a exactidão, desaparece a diferença anteriormente observada quanto aos erros globais. Quanto aos erros  $C_1$ , uma análise de variância confirmou que o efeito de grupo é significativo,  $F_{(2,35)}=12.36, p=.0001^*$ . As diferenças significativas ocorrem entre os pré-primários e os outros dois grupos, como a inspecção dos dados sugere e testes Tukey confirmaram.

A diferença entre primários e pré-primários quanto à incidência de erros  $C_1$  é pois suficientemente robusta para ser replicada num subconjunto de cerca de metade dos sujeitos.

O subgrupo dos sujeitos liceais com relativamente pior desempenho tem uma incidência de erros  $C_1$  praticamente idêntica à dos sujeitos mais jovens e menos escolarizados, de cerca de 30%. A diferença anteriormente observada entre liceais e primários (estes com menos 14% de erros  $C_1$ ) está pois associada ao facto de os grupos totais terem em média

---

\* Cf. Anexo EO.5.

**QUADRO 5:** Características do desempenho (exactidão, % R.C., de subgrupos de sujeitos liceais e universitários com nível de exactidão semelhante.

	% R.C.	%C <sub>1</sub>	%G1
Lic. (n=12)	70.3 (8.3)	52.9 (16.7)	10.7 (8.7)
Univ. (n=12)	70.2 (8.0)	48.8 (15.5)	7.6 (7.9)

**QUADRO 6:** Exactidão conforme o tipo de contraste (de lugar, de vozeamento ou duplo). O número total de estímulos por tipo de contraste está indicado. As percentagens de exactidão foram calculadas relativamente a este total.\*

CONTRASTE			
	LUGAR (n = 18)	DUPLO (n = 14)	VOZEAMENTO (n = 16)
% R.C.	56	53	51
(d.p.)	(25)	(22)	(23)

\*Apresentam-se os dados globais para todos os sujeitos, N = 96, por não se ter verificado nenhuma interacção do tipo de contraste com o grupo.

diferentes níveis de *performance*, os liceais com cerca de 61% de respostas correctas e os primários apenas com 37%.

A associação entre nível de desempenho e incidência de erros  $C_1$  e globais aparece claramente ao comparar grupos de idade e escolarização mais avançadas: se emparelharmos sujeitos universitários e liceais com a mesma percentagem de respostas correctas, verificamos que as respectivas proporções de erros  $C_1$  e globais são semelhantes (cf. Quadro 5).

Em suma, para níveis de exactidão semelhantes, a única diferença intergrupar que subsiste é a menor incidência de erros  $C_1$  nos pré-primários. Este resultado indica que o facto de saber ler e escrever se reflecte na proporção de erros segmentais.

Uma questão relacionada é se haverá efeito da escolarização no desempenho, mesmo para subgrupos com a mesma incidência de erros  $C_1$ . Doze pré-primários e 13 primários deram a mesma proporção de erros  $C_1$ , 21% (desvio-padrão de 6 e 7, respectivamente); a exactidão foi semelhante nos dois subgrupos, 38% (d.p.=18) e 34% (d.p.=15), respectivamente. Assim, o efeito da escolarização em crianças da faixa etária considerada, de 6 a 8 anos, e com níveis de desempenho semelhante, aparece essencialmente a nível da incidência de erros  $C_1$ .

#### 3.3.4. Tipo de contraste

O teste é constituído por três tipos de contraste: de vozeamento, de lugar, ou duplo. Separando as respostas correctas conforme o tipo de contraste, verificamos que a percentagem de exactidão é ligeiramente menor para contrastes de vozeamento, e relativamente maior nos de lugar (cf. Quadro 6).

QUADRO 7: Incidência de erros de tipo  $C_1$  e  $C_1--V_2$  (percentagem) para contraste de vozeamento, conforme a atenção esteja dirigida para consoante surda ou sonora. As percentagens foram calculadas relativamente ao total de erros por grupo, em 12 contrastes de vozeamento labiais, 6 com atenção a surdo, 6 com atenção a sonoro.

	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
At. Surdo	63	86	94	82
At. Sonoro	37	15	6	18

Uma análise de variância às respostas correctas, com grupo como factor intersujeito, e contraste como factor intrasujeito, mostrou que o tipo de contraste é um factor significativo na determinação da exactidão, mas não interage com o grupo\*. Considerando separadamente cada tipo de contraste, foram analisados os erros restringidos à primeira consoante,  $C_1$ , ou aqueles acompanhados por erros na segunda vogal,  $C_1$ -- $V_2$ .

### Contraste de vozeamento

Quanto aos contrastes de vozeamento, o fenómeno mais saliente é a maior incidência de erros quando a palavra atendida começa por consoante surda. Isto acontece em todos os grupos, dando-se aproximadamente três vezes mais erros quando a palavra atendida começa por surda do que quando ela começa por sonora (82% vs 28%, respectivamente). A maior parte desses erros consiste na intrusão da consoante sonora do canal secundário; aproximadamente 77% dos erros  $C_1$  e  $C_1$ -- $V_2$  quando se atende à consoante surda são intrusões da sonora homorgânica. Pelo contrário, a incidência de intrusões de  $C_1$  quando a palavra a que se presta atenção começa por sonora é mais reduzida, aproximadamente 54%. Este padrão de resultados só não é flagrante nos universitários\*\*. A relativa dificuldade quando se atende a surdo (em vez da "facilidade" quando se atende a sonoro) é também mais flagrante nos grupos de sujeitos não universitários, como se pode observar no Quadro 7. Isto talvez seja devido a uma maior eficácia no controlo da atenção no caso dos universitários.

---

\* Cf. Anexo EO.6.

\*\* Para os universitários, num total de 16 erros quando atendiam à sonora, 13 consistem em intrusões da surda correspondente (81%); num total de 27 erros quando atendiam à surda, 23 foram intrusões da sonora correspondente (85%).

**QUADRO 8:** Incidência de erros  $C_1$  e  $C_1-V_2$  para contrastes duplos, conforme a atenção esteja dirigida para consoante surda ( $n=7$ ) ou sonora ( $n=7$ ). As percentagens foram calculadas relativamente ao total de erros por grupo. Entre parêntesis, a percentagem de respostas sonoras relativamente ao total das surdas + sonoras\*.

	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
At. Surdo	67 (67)	75 (57)	56 (73)	53 (38)
At. Sonoro	33 (38)	25 (71)	44 (58)	47 (68)

**QUADRO 9:** Número de *blendings* e respostas anómalas para os 14 pares de contraste duplo, por grupo. Apresenta-se primeiro o total para respostas surdas e sonoras e depois apenas os casos em que resposta foi surda (R. Surda).

	Total		R. Surda	
	<i>Blendings</i>	Anómalas	<i>Blendings</i>	Anómalas
UNIV.	35	7	17	2
LIC.	43	13	25	2
PRIM.	38	15	20	3
PRÉ-PRIM.	21	12	15	3

\* Por exemplo, nos 7 contrastes em que a atenção incide na surda, os universitários deram 11 erros  $C_1(V_2)$  surdos e 22 sonoros; nos 7 contrastes com atenção a sonoro, deram 10 erros surdos e 6 sonoros. A percentagem de erros com atenção a sonoro é de  $(10 + 6)/(10 + 6 + 11 + 22) = 33\%$ ; nesta condição, a percentagem de respostas sonoras é de  $6/(10 + 6) = 38\%$ .

O facto de haver mais erros quando se atende a surdo, e esses erros consistirem maioritariamente em intrusões, certamente tem a ver com as características do português, como já referimos no estudo anterior.

### Contraste de lugar

Quanto aos contrastes de lugar, separámos os pares de estímulos e respectivas respostas erradas conforme fossem sonoras ou surdas. É possível verificar se se mantem a nível de resposta o carácter surdo ou sonoro dos estímulos. Como seria de esperar, observa-se em geral essa manutenção da sonoridade. Se ambos os estímulos são sonoros, há mais respostas erradas sonoras (60%) do que surdas (44%). Se ambos os estímulos são surdos, é ainda mais flagrante como a resposta mantém o valor de vozeamento dos estímulos: 81% das respostas erradas são também surdas, e apenas 19% sonoras.

### Contrastes duplos: combinações de traços fonéticos

Tal como nos contrastes de vozeamento, também nos contrastes duplos tende a haver mais erros se a atenção está orientada para a palavra começada por surda, nomeadamente nos liceais e universitários (cf. Quadro 8).

Nesta situação de contraste duplo, é particularmente relevante uma análise dos erros em termos de *blendings* e respostas anómalas. No Quadro 9 pode ver-se que em todos os grupos houve uma preponderância de *blendings* relativamente às anómalas (coluna "TOTAL"). Esta preponderância foi significativa em todos os grupos (pelo *sign test*,  $p < .02$ , nos

universitários, liceais e primários), excepto no pré-primário\*. Separámos as respostas conforme fossem surdas ou sonoras pois, como as análises prévias demonstraram, parece haver uma tendência em aceitar o vozeamento independentemente do ouvido a que tenha sido apresentado; assim, os *blendings* sonoros poderiam resultar mais de um erro de lugar do que de uma combinação de traços, correcta e separadamente extraídos, mas incorrectamente recombinaados. Considerando apenas as respostas surdas, mantem-se claramente a preponderância dos *blendings*: a razão entre o número de *blendings*, por um lado, e o total dos *blendings* mais as respostas anómalas por outro, é de .89 nos universitários, .93 nos liceais, .87 nos primários e .83 nos pré-primários. Todavia, os resultados do *sign test* revelaram de novo que não é significativa a preponderância de *blendings* no caso das crianças pré-primárias ( $p=.37$ ).

A correlação entre, por um lado, a diferença entre o número de *blendings* e anómalas (*blendings* - anómalas) e, por outro, o total das respostas correctas, foi relativamente baixa e não significativa para os quatro grupos considerados ( $n = 24$ ): nos pré-primários,  $r = -.16$ ,  $p = .226$ , nos primários,  $r = -.01$ ,  $p = .472$ , nos liceais,  $r = -.07$ ,  $p = .371$  e nos universitários,  $r = -.33$ ,  $p = .056$ . Estes resultados são semelhantes aos obtidos na experiência anterior.

### 3.3.5. Fusões fonológicas

Os erros que não respeitaram a estrutura CVCV dos estímulos, ou erros estruturais, foram examinados com vista a detectar a hipotética ocorrência de fusões fonológicas. Verificaram-se respostas de tipo fusão

---

\* Os valores foram os seguintes: para os universitários,  $p < .001$ ; para os liceais,  $p = .006$ ; para os primários,  $p = .01$ ; para os pré-primários  $p = .50$ .

**QUADRO 10:** Resumo das observações de fusões fonológicas, por grupo e par de palavras. Entre aspas, a resposta dos sujeitos.

	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
BODE-pala "Balde"	3	3	12	11
TODA-calor "caldo"	0	0	0	3
BARRO-pouco "Barco"	1	0	2	1
POSSE-gula "Bolsa"	0	0	2	1
CACHO-goto "Gosto"	0	0	2	1
COCO-tara "Porco"	0	1	0	0
BURRA-ponto "Burto"	0	0	0	1
Total (T)	4	4	18	18
(T/Erros)%	1.3	0.9	2.5	2.4

**QUADRO 11:** Média (d.p.) de erros de tipo não-palavra e respectiva proporção média relativamente ao total dos erros, por grupo. Todos os cálculos foram efectuados individualmente.

GRUPO	UNIV.	LIC.	PRIM.	PRÉ-PRIM.
Média	2.2 (1.7)	3.0 (2.0)	6.9 (4.0)	6.6 (4.0)
Proporção	.16 (.11)	.18 (.14)	.23 (.10)	.22 (.13)

fonológica em reduzido número. A mais frequente dava-se perante o par BODE-pala, que repetidamente foi ouvido como "balde". Este erro ocorreu em 3 do total das 4 fusões observadas nos universitários, e em 12 do total das 18 fusões dos primários (para mais detalhes, cf. Quadro 10).

O número bruto de fusões fonológicas parece estar inversamente associado à exactidão, tal como acontecia com os erros globais. A discrepância intergrupo quanto à incidência de fusões atenua-se se as considerarmos em relação ao total de erros; todavia, a quantidade reduzida de observações e o próprio planeamento da experiência não permitem extrair qualquer conclusão.

### 3.3.6. Respostas de tipo não-palavra; fluência de leitura

Poderão os resultados anteriormente apresentados estar confundidos com a tendência ou relutância em produzir não-palavras? Claro que o número bruto de não-palavras está em todos os casos correlacionado com a exactidão; todavia, a sua proporção relativamente ao total dos erros foi semelhante para os quatro grupos (cf. Quadro 11). Uma análise de variância aos *scores* transformados não revelou qualquer efeito significativo\*.

Uma severa limitação à análise dos *blendings* e anómalas seria a existência de factores lexicais que sistematicamente favoreceriam a ocorrência dos primeiros (como por exemplo as respostas anómalas darem não-palavras e os *blendings* palavras). Um controle *a posteriori* mostra que esta possível interpretação em termos de viés lexical não dá conta dos resultados (cf. Anexo EO. Blend). Considerando o total dos 14 contrastes duplos, entre as respostas anómalas possíveis há mais não-palavras do que

---

\* Cf. Anexo EO.7.

QUADRO 12: Tempo médio (em segundos) requerido para a leitura de três colunas (Univ. e Lic.) ou uma (Prim. I), por sexo. Entre parêntesis, o desvio-padrão.

GRUPO	UNIV.	LIC.	PRIM.
MASC.	24 (2.2)	23 (3.2)	16 (3.6)
FEM.	27 (4.4)	25 (4.7)	16 (4.4)

QUADRO 13: Comparação da exactidão (%R.C.) e incidência de erros C<sub>1</sub> (% C<sub>1</sub>) nos grupos infantis, jovens e adultos (Let., letradas, Slet., semiletradas, Ilet., iletradas).

	% R.C.	% C <sub>1</sub>
Pré-prim. n=24	32	15
Prim. n=24	32	27
Lic. n=24	67	41
Univ. n=24	72	48
Let. n=14	64	39
Slet. n=14	49	29
Ilet. n=18	44	24

palavras; todavia, para os *blendings* possíveis, palavras e não-palavras são equiprováveis. Além disso, considerando apenas os casos de resposta surda e atenção a surdo, a probabilidade de respostas de tipo palavra é idêntica para *blendings* e anómalas, .85, ou seja 6 palavras e uma não-palavra. Ora nestas condições a razão de *blendings*, bl., para anómalas, an., (bl.: an.) foi de 10:1 nos universitários, 20:2 nos liceais, 11:0 nos primários e 10:3 nos pré-primários. Não se pode pois atribuir a preponderância de *blendings* a um viés lexical.

Os resultados obtidos no teste de fluência de leitura podem observar-se no Quadro 12. Em nenhum caso se observou correlação significativa entre fluência de leitura e exactidão.

### 3.3.7. Comparação dos grupos infantis e jovens com os adultos

Particularmente importante para o esclarecimento do efeito da alfabetização na percepção de fala é dissociar os efeitos da alfabetização propriamente dita, dos da idade e desenvolvimento em geral. Neste contexto, assume particular relevância a comparação dos resultados obtidos pelos adultos (Experiência III) com os dos sujeitos desta experiência, nomeadamente as três comparações seguintes: entre adultos iletrados e crianças pré-letradas, entre adultos semiletrados e crianças primárias, e entre letradas menos jovens e universitários. Evidentemente que não é só a idade a diferenciar estes grupos; todavia, mesmo aceitando que estas comparações estão contaminadas por muitos outros factores, elas ajudam à compreensão dos fenómenos em foco.

Como os grupos em questão tinham diferentes níveis de *performance* (para apreciação dos resultados globais, cf. Quadro 13), as comparações foram efectuadas para subgrupos emparelhados por

percentagem de respostas correctas. As catorze letradas e 13 dos universitários tinham em média 64% de exactidão; a proporção de erros  $C_1$  nos dois grupos foi praticamente idêntica: 38.7 e 39.3, respectivamente. Catorze dos primários\* e 11 das semiletradas tiveram em média 42% de respostas correctas; as proporções de erros  $C_1$  foram 32% e 25%, respectivamente; esta diferença não é significativa,  $t_{(25)}=1.31$ ,  $p=.204^{**}$ . Catorze pré-primários\*\*\* e 15 iletradas tiveram um nível de *performance* de, respectivamente, 42 e 41%; as proporções correspondentes de erros  $C_1$  são 15% e 22%. Esta diferença não atinge a significância,  $t_{(27)}=1.89$ ,  $p=.069$ .

Assim, entre sujeitos com o mesmo nível de desempenho, as diferenças quanto à incidência de erros de primeira consoante podem atribuir-se ao conhecimento e experiência com a escrita. A experiência com a fala e a idade em geral parecem desempenhar um papel negligenciável.

No contexto da comparação entre os grupos de adultos e os grupos de jovens e crianças, vale a pena apreciarmos a incidência de intrusões. Adultos menos jovens (entre 38 e 63 anos), universitários (cerca de 20 anos) e estudantes liceais (cerca de 15 anos) têm uma percentagem de intrusões notavelmente semelhante, entre 2 a 5%, em média\*\*\*\*. As intrusões parecem ser ligeiramente mais frequentes nas semiletradas e iletradas adultas, onde se observam em média cerca de 10% de intrusões. Este nível é semelhante ao das crianças primárias. Finalmente, os pré-primários constituem um grupo à parte, com 25% de intrusões, em média. Isto indica claramente que a proficiência de controlo da atenção sofre um

\* O subgrupo emparelhado com os pré-primários, cf. Quadro 4.

\*\* Cálculos com valores transformados. Os valores da probabilidade de  $t$  referem-se a testes bicaudais.

\*\*\* O subgrupo emparelhado com os primários, cf. Quadro 4.

\*\*\*\* Os valores, apresentados no Quadro 2 da presente experiência e no Quadro 1 da Experiência III (p.182'), são: para os universitários, 2.3% para as letradas, 3.8%, e para os liceais, 4.9%.

desenvolvimento associado quer à idade (veja-se a diferença entre iletradas, com 7% de intrusões, e pré-primários, com 25%), quer ao nível de escolarização (veja-se o decréscimo de intrusões dos primários, 10%, aos universitários, 2%).

### 3.4. Conclusão

Numa situação de audição dicótica de palavras CVCV, crianças de 6 e 9 anos têm um desempenho marcadamente inferior ao de jovens de 15 e 22 anos. Melhores resultados estão associados a um padrão de respostas erradas em há menos erros globais (que envolvam as sílabas completas ou os quatro segmentos de palavra) e mais erros segmentais, restritos à primeira consoante.

Considerando níveis de exactidão semelhantes, observou-se uma clara dissociação entre pré-primários e restantes sujeitos quanto aos erros de primeira consoante.

Tínhamos sugerido que, em níveis de exactidão semelhante, a preponderância de erros segmentais reflecte um efeito da alfabetização: sujeitos que adquiriram uma consciência da estrutura fonética da fala através da aprendizagem da leitura e da escrita num sistema alfabético terão desenvolvido processos de atenção à estrutura segmental da fala. Estes processos seriam particularmente úteis em situações como a de audição dicótica, em que a identificação é dificultada pela ausência de informação contextual ou pragmática.

Se assim fosse, crianças pré-letradas deveriam dar menos erros segmentais do que crianças e jovens letrados. Ora foi exactamente isso que observámos na presente experiência. Replicámos pois o efeito da

alfabetização descoberto na experiência anterior, em que adultos iletrados deram menos erros segmentais do que adultos letrados com o mesmo nível de desempenho. Crianças pré-letradas e adultos iletrados constituem um grupo homogêneo quanto à proporção de erros segmentais, distinguindo-se quer de crianças letradas e adultos semiletrados, quer de universitários e adultos letrados, em quem aquele tipo de erros surge mais frequentemente.

Encontrámos assim evidência convergente em favor da hipótese de que a alfabetização afecta o modo como se identificam palavras em condições difíceis, através da utilização, ou não, de uma estratégia de atenção aos segmentos da fala.

Quanto aos processos automáticos de análise fonética, esta experiência replicou os achados de que (i) são em geral mais frequentes os erros que envolvem a combinação errónea de traços fonéticos existentes na estimulação do que erros com um novo traço; (ii) não aparece nenhuma relação entre o nível de desempenho e o reflexo provável da acção desses processos (uma medida da preponderância dos *blendings* face às respostas de oclusiva anómala). De notar ainda que a preponderância de *blendings* face às respostas anómalas parece ser menos marcada nas crianças de 6 anos do que em crianças mais velhas e adultos. O facto de não termos encontrado diferença significativa através do *sign test* entre *blendings* e anómalas neste grupo de sujeitos, pode ser devido a uma falta de poder estatístico (temos relativamente poucas observações) ou pode traduzir uma diferença real subjacente. É possível que se trate de um efeito relacionado com a idade, e que a ocorrência de fusão de traços fonéticos dependa da maturação das vias de comunicação hemisféricas, ou outro factor ontogenético.

Finalmente, abordemos a questão da validade ecológica das experiências III e IV. Com efeito, uma das críticas que lhes poderiam ser

tecidas é que os resultados observados não teriam quaisquer implicações fora da situação "artificial" de audição dicótica. Além disso, poder-se-ia objectar que uma situação artificial de audição dicótica induziria um sentimento de estranheza nos sujeitos, principalmente nos de menor nível educacional. A pouca familiaridade com a situação poderia de alguma maneira determinar um comportamento "bizarro" por parte destes sujeitos (cf. Koopmans, 1986, a propósito das tarefas de segmentação em iletradas). Paradoxalmente, este ponto de vista tem mais a ver com teorizações de gabinete do que com observações em situação. Nem as crianças, nem as semiletradas e iletradas pareciam ficar constrangidas perante a situação de audição dicótica. No caso das crianças, tentava-se à partida dar à tarefa um certo carácter lúdico. Quanto aos sujeitos adultos, eles próprios comentavam espontaneamente "agora vou ligar para este lado", ou "agora é daqui", com a maior das naturalidades. Uma iletrada expressava a situação com uma ponta de animismo: "agora vou ouvir a porta, há bocado era a janela" (a porta estava à sua esquerda, a janela à sua direita). Assim, infirmando a especulação de que os resultados estariam enviesados pela artificialidade da situação, está o comportamento *de facto* dos sujeitos: nesta como noutras situações do dia-a-dia era preciso focalizar a atenção num acontecimento relevante entre outros a que "não se liga".

O facto de uma situação de observação ser "artificial" (leia-se especial) não implica necessariamente que as variáveis em jogo na determinação do comportamento enfermem de uma limitadora especificidade situacional. Pelo contrário, a tal situação especial pode mesmo ser indispensável para desvendar processos e variáveis relevantes em circunstâncias naturais (leia-se habituais). Ninguém se atreveria a pensar que os letrados adultos e as crianças da primária tivessem desenvolvido uma estratégia especial na tarefa de audição dicótica de modo a obterem melhores resultados e mais erros segmentais. Em contrapartida, se

é verdade que no decurso da aquisição e prática da leitura e/ou escrita os sujeitos se tornaram capazes de prestar atenção aos segmentos da fala, então reflexos desta competência deverão ser observados noutras situações; por exemplo, ao ouvir palavras em ruído sem informação pragmática esclarecedora\*. Valerá a pena, sem dúvida, testar situações daquele tipo.

---

\* No sentido de "disambiguating".

#### 4. Inteligibilidade de palavras em ruído: Comparação entre letrados, semiletrados e iletrados (Experiência V)

##### 4.1. Introdução

A experiência que se segue explora mais uma situação de reconhecimento de palavras faladas em condições difíceis, através da apresentação simultânea de ruído.

É conhecido que o ruído deteriora a inteligibilidade da fala, nomeadamente das consoantes relativamente às vogais (e.g., Sorin e Thouin-Daniels, 1983). O estudo clássico de Miller e Nicely (1955) mostrou que os erros ao perceber fala em ruído obedecem a um padrão interpretável à luz de factores acústico-linguísticos. A análise da matriz de confusões de 16 consoantes inglesas (as oclusivas /p, t, k, b, d, g, m, n/, e as fricativas /f, θ, s, ʃ, v, z, ʒ/) revelou que o ruído exerce um efeito particularmente pernicioso quanto ao lugar de articulação: são mais frequentes os erros em que se troca o lugar de articulação, do que a maneira ou o vozeamento. Por exemplo, confunde-se mais /p/ com /t/, do que com /v/ ou /b/. A nasalidade e o vozeamento são as características mais preservadas. Há relativamente menos confusões deste tipo mesmo a uma razão sinal/ruído de -12dB, enquanto já se torna difícil distinguir o lugar de articulação a razões inferiores a 6 dB (*ibidem*).

O objectivo desta experiência não é analisar o efeito do ruído *per se*; é antes avaliar se a identificação de palavras em ruído é afectada, quer em termos quantitativos, quer qualitativos, pelo facto de se saber ler e escrever. Comparar-se-á pois a percentagem de respostas correctas e o padrão de erros de sujeitos letrados, semiletrados e iletrados nas seguintes condições: palavras apresentadas a um só ouvido (condição monaural ou controlo); palavras num ouvido e simultaneamente ruído no outro ouvido (condição de

ruído contralateral); palavras e ruído misturados, nos dois ouvidos (condição de ruído binaural).

Optamos por usar as mesmas palavras nas três condições. Evitamos assim a possível confusão entre o efeito do ruído e a maior ou menor facilidade das palavras por questões lexicais ou outras (cf. Experiência I, onde apesar do cuidado em escolher palavras foneticamente comparáveis e relativamente comuns se verificaram grandes diferenças quanto à sua inteligibilidade); com este planeamento, cada palavra é controlo de si própria - ou seja, vê-se o efeito do ruído na identificação da palavra "x" relativamente à sua identificação sem ruído. Isto obriga-nos, claro, a dar especial atenção à ordem das condições.

A introdução de duas condições de ruído visa, por um lado, diversificar as situações de observação numa fase relativamente exploratória da investigação; por outro, permite colher observações relevantes sobre a eficácia do controlo da atenção nos sujeitos considerados.

De acordo com a hipótese de que a aquisição da leitura e da escrita promove o desenvolvimento de uma estratégia de atenção à estrutura fonética das palavras, e supondo que essa estratégia é útil em situações de reconhecimento da fala em que são inexistentes ajudas contextuais ou pragmáticas, o efeito do ruído não deveria ser idêntico para os grupos de letrados e iletrados. O decréscimo de *performance* provocado pelo ruído deveria ser maior em sujeitos iletrados do que em sujeitos letrados.

Além deste efeito quantitativo, deveria também observar-se um padrão de erros diferente conforme a alfabetização: sujeitos letrados deveriam dar relativamente menos erros globais e mais erros segmentais do que sujeitos iletrados.

A posição das semiletradas face aos outros dois grupos permitirá esclarecer até que ponto os eventuais efeitos observados se poderão atribuir à alfabetização propriamente dita ou a outros factores, como o nível de escolarização e a experiência com a fala e a escrita.

Como é difícil encontrar sujeitos das populações com que trabalhamos, e deveras moroso realizar as sessões experimentais\*, resolvemos o problema da ordem das condições da seguinte maneira: em cada grupo, cerca de metade dos sujeitos ouvia primeiro as palavras sem ruído, seguindo-se o ruído contralateral e finalmente o ruído binaural (ordem I); a outra metade começava por ouvir as palavras em ruído binaural, passando-se a seguir para a condição de ruído contralateral e finalmente a condição controlo (ordem II). Este tipo de planeamento permitir-nos-á comparar o desempenho na condição de ruído binaural após os sujeitos terem já experiência com as palavras (na ordem I) ou não (ordem II). Certamente a última situação é mais difícil. Poderemos assim verificar até que ponto a hipotética estratégia de atenção segmental é útil em situações de diferentes graus de dificuldade. Uma suposição plausível é que esta estratégia seja particularmente relevante quando há já uma certa familiaridade com as palavras (cf. Samuel e Ressler, 1986), ou quando o nível de dificuldade da tarefa seja intermédio. Assim, deveria haver um maior efeito de grupo na ordem I do que na ordem II.

Com o objectivo de dispor de dados adicionais sobre o funcionamento cognitivo dos sujeitos, particularmente dos que não são frequentadores dos laboratórios de Psicologia, todos os participantes fizeram um breve teste de amplitude de memória imediata de dígitos, no fim da

---

\* Além das deslocações até aos sujeitos, por vezes através de caminhos onde não chega o automóvel, ocasiões havia em que a sessão experimental se prolongava em longas conversas. Num fim-de-semana de trabalho (pois em geral só ao fim-de-semana é que letradas e semiletradas estavam disponíveis), era vulgar recolher dados de apenas 4 sujeitos.

sessão experimental. Recorremos à adaptação portuguesa da Escala de Inteligência de Wechsler para Crianças (Marques, 1969), de onde extraímos o teste de memória de dígitos\*. Escolhemos o teste de memória de dígitos porque, além de ser rápido e de fácil aplicação, foi apontada na literatura uma relação entre a escolarização e a amplitude de memória de dígitos (Read e Ruyter, 1985), que seria importante replicar.

## 4.2. Método

### 4.2.1. Sujeitos

Participaram nesta experiência 96 sujeitos, 36 iletradas, 32 letradas e 28 semiletradas, todos do sexo feminino. As iletradas tinham entre 40 e 65 anos (idade média de 52.3, desvio padrão de 6.3); eram habitantes de zonas rurais ou semi-urbanas do Norte do país\*\*, onde trabalhavam em serviços agrícolas e/ou domésticos. As semiletradas tinham entre 36 e 65 anos de idade (média de 48.0, d.p. = 7.0). Vinte e cinco habitavam em meio rural\*\*, onde trabalhavam em serviços agrícolas e/ou domésticos; as restantes viviam em meio urbano ou semi-urbano,\*\*\* e trabalhavam em serviços domésticos. Todas tinham no máximo 4 anos de escolaridade. No grupo das letradas, as idades variavam entre os 35 e os 66 anos (média de 48.8, d.p. = 7.0). Todas foram recrutadas entre docentes e funcionárias de secretaria do ensino primário e secundário em estabelecimentos da zona do Porto e Grande Porto\*\*\*\*. Catorze possuíam

---

\* Optamos por usar o subteste da WISC por existirem normas portuguesas que poderão ser úteis em futuros trabalhos onde se faça uma comparação destes dados com os de crianças de vários níveis etários.

\*\*Vale de Cambra, lugar de Lordelo e Vila da Feira, lugar de Lourosa.

\*\*\* Porto, Espinho e arredores.

\*\*\*\* Águas Santas e Gondomar.

grau universitário (16 anos de formação); as restantes tinham no mínimo 12 anos de formação.

Letradas e semiletradas fizeram um teste de leitura constituído pelas duas primeiras colunas do teste previamente usado (cf. Anexo L.1). Os sujeitos eram instruídos a ler cada coluna a um ritmo confortável e sem pausas, sendo anotado o tempo levado a ler cada uma delas. Infelizmente, seis letradas e quatro semiletradas não puderam realizar este teste por motivo de falta de tempo (o teste era feito no fim da sessão experimental). Se as semiletradas evidenciassem grande dificuldade na leitura da coluna mais simples, soletrando a custo e lentamente ( $\leq 30$  segundos), não lhes era solicitado que lessem a segunda coluna. Isto aconteceu com seis sujeitos.

Em nenhum dos participantes era conhecida a existência de problemas neurológicos, intelectuais ou auditivos. Nenhum evidenciou perdas auditivas aparentes, uni- ou bilaterais, quando submetido a um teste sumário de acuidade constituído por uma série monaural de 12 palavras CVCV (6 em cada ouvido), apresentadas em dois níveis de intensidade decrescente (70 e 50 dB SPL). Tolerou-se no máximo dois erros, espontaneamente corrigidos pelo sujeito numa segunda audição.

#### 4.2.2. Estímulos

Foram usadas 40 palavras dissilábicas CVCV, todas começando por uma oclusiva (/b, d, g, p, t, k/), com acentuação na primeira sílaba e pronunciadas por um falante masculino português (cf. Anexo IP. 1/1). Este conjunto foi seleccionado do material usado na experiência I, com base nos resultados do pré-teste que a seguir se descreve.

Gravaram-se 48 palavras sob três condições: só num canal, ou as palavras num canal e ruído-máscara no outro canal, ou palavras e ruído misturados nos dois canais\*. Estas condições correspondem respectivamente, à situação controlo ou de apresentação monaural, à situação de ruído contralateral e à situação de ruído binaural. O ruído consistia numa sobreposição de vários ruídos de rádio; fenomenologicamente, assemelhava-se a ruído branco, com um timbre ligeiramente mais metálico. Na condição binaural, foi gravado com a mesma intensidade das palavras. Na condição de ruído contralateral, usaram-se dois níveis sonoros: para metade dos estímulos, a intensidade do ruído era idêntica à das palavras; na outra metade, estas tinham uma atenuação de -20 dB. Cada palavra era precedida por um "beep", que actuava como sinal de aviso.

A sequência dos 48 estímulos nas condições controlo, ruído contralateral e ruído binaural, foi apresentada nesta ordem a 12 estudantes universitários (1º ano do Curso de Psicologia, 1985/86), 8 semiletradas e 4 iletradas. Os resultados mostraram que: (i) a manipulação do nível de intensidade do ruído contralateral não provocava efeitos no desempenho; (ii) havia palavras sistematicamente mal percebidas. Assim, foram refeitas as gravações usando apenas 40 palavras cujo nível de desempenho em ruído binaural se situava entre cerca de 10 e 90%, e eliminando a manipulação do nível sonoro em ruído contralateral; foi tomada a condição em que a intensidade da palavra é idêntica à do ruído.

A gravação definitiva foi apresentada aos sujeitos a um nível de aproximadamente 75 dB SPL através de auscultadores Philips EM 6421 e de um gravador Philips D 6920 MK 2.

---

\*O material experimental foi preparado no Laboratório de Psicologia Experimental da Universidade Livre de Bruxelas.

#### 4.2.3. Procedimento

Os sujeitos foram testados individualmente. Experimentadora e sujeito estavam sentados frente a frente numa sala sossegada\*. Explicava-se que se tratava de "um estudo sobre a maneira como as pessoas ouvem palavras, e como o ruído ou barulho pode dificultar a audição". A tarefa consistia em "repetir o que vai ouvindo através dos auscultadores. Quando as palavras aparecem sózinhas, isso é muito fácil, como vai ver. Quando as palavras aparecem misturadas com ruído, vai ser mais difícil e por vezes certamente vai achar que não se percebe nada. Mesmo assim, por favor tente perceber e diga o que ouviu. Mesmo que lhe pareça disparate, ou que não tenha bem a certeza, é muito importante que repita exactamente/igualzinho ao que lhe pareceu. Antes de cada palavra, há um apito que serve para avisar que logo a seguir vem a palavra. Por isso, quando ouvir o apito por favor preste atenção". Entre cada condição, fazia-se uma breve pausa e explicava-se de novo a situação seguinte; antes da condição de ruído contralateral, dizia-se "agora vai ouvir as palavras num lado e o ruído no outro; por favor preste atenção só ao lado onde surgem as palavras", e indicava-se a que lado se deveria estar atento. Por ocasião da audição em monaural perguntava-se ao sujeito se sentia algumas dificuldades em ouvir (como som turvo, diferenças de ouvido ou outras); se isso se passasse, ou se notasse um padrão anormal de respostas nesta condição, o sujeito era substituído.

Tínhamos observado no pré-teste que a condição de ruído binaural era muito difícil; por vezes os sujeitos relatavam que não conseguiam "perceber nada". Isto por vezes acontecia com vários estímulos consecutivamente, pelo que a motivação e atenção do sujeito podiam sofrer

---

\* Para a maioria dos sujeitos iletrados e semiletrados o local de realização da experiência foi o seu próprio domicílio.

QUADRO 1: Número médio de respostas correctas (desvio-padrão entre parêntesis) por grupo (iletradas, semiletradas e letradas) e condição (controlo, ruído contralateral e binaural), nesta ordem (ordem 1, 0<sub>1</sub>) ou na ordem contrária (ordem 2, 0<sub>2</sub>). O máximo possível de respostas correctas é 40. Indica-se o número de sujeitos testados em cada grupo e ordem.

		Iletradas		Semiletradas		Letradas	
O <sub>1</sub>	Controlo	35.3	(2.2)	36.7	(2.1)	38.0	(1.5)
	R. Contral.	34.4	(2.8)	35.9	(2.8)	36.1	(2.0)
	R. Binaural	14.0	(4.8)	17.9	(5.2)	19.8	(5.1)
		(n=16)		(n=14)		(n=14)	
O <sub>2</sub>	Controlo	37.5	(1.2)	37.8	(1.7)	38.6	(1.5)
	R. Contral.	34.6	(3.0)	33.5	(3.6)	37.2	(2.0)
	R. Binaural	6.6	(5.1)	8.8	(4.5)	12.3	(6.0)
		(n=20)		(n=14)		(n=18)	

um decréscimo. Por estas razões, frisava-se que era importante "não desistir de tentar perceber"; o sujeito era solicitado a dizer-nos se porventura estivesse distraído, para que pudéssemos parar o gravador e tornar a apresentar a palavra.

Dezasseis iletradas, 14 semiletradas e 14 letradas fizeram a experiência na ordem controlo, ruído contralateral, ruído binaural; 20 iletradas, 14 semiletradas e 18 letradas fizeram-na na ordem contrária. Nas condições controlo e ruído contralateral, as palavras eram apresentadas em dois blocos, um ao ouvido esquerdo e outro ao direito. A ordem de ouvido estava contrabalançada dentro de cada grupo.

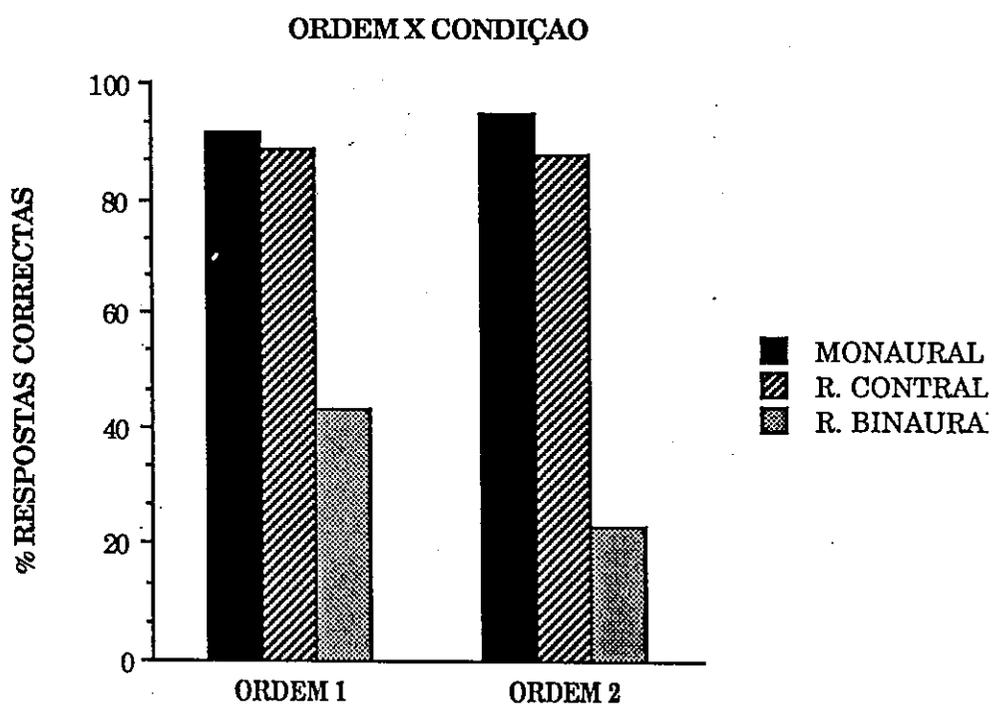
Antes da experiência, os sujeitos completavam o teste rápido de acuidade auditiva. No fim da sessão experimental era realizado o teste de Memória de Dígitos (subteste da Escala e Inteligência de Wechsler para Crianças, WISC) e de fluência de leitura (sujeitos semiletrados e letrados).

Nesta mesma sessão experimental, alguns sujeitos tinham previamente ouvido a gravação com os *continua* (cf. Experiência VI). Esta tinha uma duração reduzida - cerca de cinco minutos - e os sujeitos não se queixavam que fosse difícil ou exigisse muita concentração. Como não se notaram efeitos de fadiga ou outros, após um intervalo passava-se à apresentação dos estímulos aqui descritos.

#### 4.3. Resultados e discussão

##### 4.3.1. Nível de desempenho nas três condições

Como mostra o Quadro 1, a presença de ruído binaural afecta marcadamente a exactidão de resposta, que desce de cerca de 88 a 100% na condição controlo para cerca de 18 a 50% perante ruído binaural. O nível de



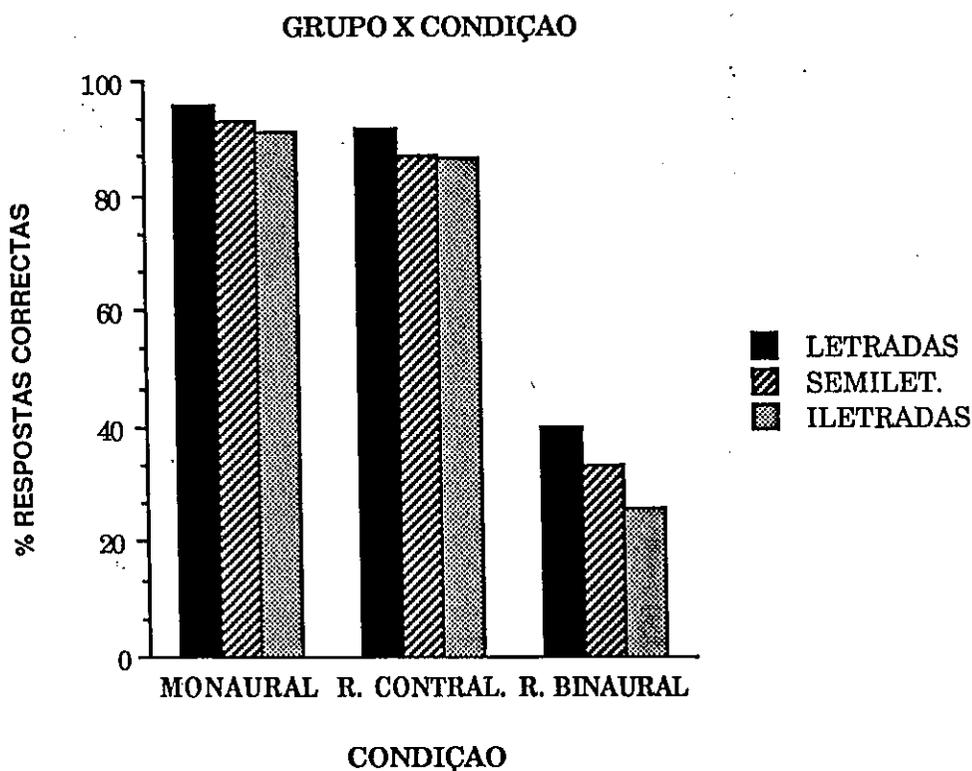
**Figura 1:** Efeito da ordem das condições no desempenho: comparação da exactidão (% de respostas correctas) na audição de palavras, monaural (MON.), com ruído contralateral (R. CONTR.) e com ruído binaural (R. BIN.), efectuadas nesta ordem (ORD. I) ou na ordem contrária (ORD. II). Apresentam-se os dados relativos aos três grupos por não se ter verificado interacção tripla ( $F < 1$ ).

desempenho na condição de ruído contralateral varia entre cerca de 84 e 93%, sendo muito próximo do alcançado na condição controlo.

Em todos os casos, as letradas apresentam resultados superiores aos das semiletradas, sendo estes por sua vez superiores aos das iletradas. A diferença entre estes grupos é aparentemente mais marcada na condição mais difícil, a de ruído binaural (percentagem de exactidão de 31%, 22% e 17%, respectivamente), do que na condição mais fácil de ausência de ruído, em que os valores correspondentes são 96%, 95% e 94%. A influência da ordem das condições evidencia-se claramente: o treino prévio beneficia a *performance* na condição que é realizada em último lugar, quer esta seja a mais difícil (ruído binaural), ou a mais fácil (sem ruído, apresentação monaural ou controlo).

Para avaliar mais precisamente o papel das várias fontes de variação, foi feita uma análise de variância ao número de respostas correctas com duas variáveis intersujeito, grupo e ordem, e uma variável intrasujeito, ruído. São altamente significativas as interacções da ordem com a condição, e do grupo com a condição, respectivamente  $F_{(4,180)}=4.58$ ,  $p=.0015$  e  $F_{(2,180)}=61.38$ ,  $p=.001$ . As restantes interacções são praticamente inexistentes ( $F \approx 1$  e  $F < 1$ ), e todos os efeitos principais são significativos.\* A interacção entre ordem e condição pode ser apreciada na Figura 1. Testes Tukey de comparação de médias a  $\alpha=.05$  revelaram que a diferença entre condição controlo e ruído contralateral só é significativa na ordem II. Se os sujeitos ouvem primeiro as palavras na ausência de ruído, o decréscimo de *performance* provocado pelo ruído contralateral não chega a atingir significância. Apesar deste decréscimo ser mais saliente no caso das iletradas e semiletradas, o facto de a interacção tripla entre ruído, ordem e grupo não ser significativa impede-nos de atribuir especial significado

\* Para pormenores, cf. Anexo IP.2/1



**Figura 2:** Efeito do grupo no nível de desempenho, por condição. Apresenta-se a percentagem de respostas correctas das iletradas, semiletradas e letradas em audição monaural, com ruído contralateral e com ruído binaural. Mostram-se os resultados independentemente da ordem por não ter havido interação tripla ( $F < 1$ ).

àquelas diferenças grupais. Todas as outras comparações são significativas: nas duas ordens, o decréscimo de monaural para binaural, e de ruído contralateral para binaural.

Particularmente relevante neste estudo é a interação entre grupo e condição, mostrada na Figura 2. Testes Tukey ( $\alpha=.05$ ) revelaram que a diferença entre letradas e iletradas é significativa em todos os casos, ou seja, em ambas as ordens e nas três condições. A posição das semiletradas relativamente aos outros dois grupos varia conforme a condição: em monaural e em ruído contralateral, o nível de *performance* das semiletradas é significativamente inferior ao das letradas, e semelhante ao das iletradas; com ruído binaural, porém, a diferença significativa é entre semiletradas e iletradas. Assim, nas tarefas mais fáceis - audição em monaural ou com ruído contralateral - as letradas constituem um grupo à parte, distinguindo-se dos outros de grau de escolarização reduzido ou inexistente; na tarefa difícil, são porém as iletradas que constituem o grupo à parte, com um desempenho significativamente inferior ao dos grupos escolarizados, quer em grau reduzido (semiletradas), quer em grau elevado (letradas).

Em suma, os resultados mostram claramente que a condição de ruído binaural é a mais difícil, particularmente se é a primeira a ser executada. As diferenças grupais são mais marcadas nesta situação, em que as iletradas têm um nível inferior ao das semiletradas e letradas.

Apesar de se ter tentado constituir grupos etários homogêneos e semelhantes, não foi possível eliminar uma razoável variabilidade intra e intergrupo. Uma possível explicação dos piores resultados das iletradas seria a sua idade relativamente mais avançada. Todavia, a correlação entre a sua idade e desempenho foi praticamente inexistente (e não significativa, cf. Anexo IP.2/2), tanto na condição monaural, como nas de ruído contralateral e binaural.



Foram calculadas as correlações entre o desempenho nas duas condições com ruído. As *performances* na situação de ruído contralateral e ruído binaural estão positivamente correlacionadas no caso dos sujeitos iletrados e semiletrados: respectivamente, na ordem 1,  $r = .77$  ( $n=16$ ,  $p = .0004$ ) e  $r = .68$  ( $n=14$ ,  $p = .0074$ ), e na ordem 2,  $r = .65$  ( $n = 20$ ,  $p = .0016$ ) e  $r = .64$  ( $n=14$ ,  $p = .01$ ). Esta relação não aparece no grupo das letradas: as correlações correspondentes são de  $-.28$  e  $.20$ , ambas com elevada probabilidade ( $p = .32$ ). Isto poderá ter acontecido porque na condição de ruído contralateral as letradas têm resultados muito elevados, havendo assim um efeito de tecto.

#### 4.3.2. A diferença intergrupo na condição controlo

O resultado do pior desempenho dos sujeitos semiletrados e iletrados na condição controlo foi inesperado e coloca vários problemas: como interpretar esta observação *per se*, e que significado atribuir às diferentes magnitudes de decréscimo entre condição controlo ou linha de base e condições experimentais propriamente ditas? Uma interpretação aparentemente mais parcimoniosa coloca-se em termos de acuidade auditiva: apesar dos cuidados em evitar que sujeitos com problemas auditivos mais ou menos marcados fizessem parte da amostra definitiva, teria havido alguns participantes, mormente iletrados, com deficiências auditivas mais subtis que foram incorrectamente englobados na amostra. Estes sujeitos distorceriam os valores médios do seu grupo, quer na condição controlo, quer nas condições experimentais. De facto, não podemos, em rigor, eliminar *a priori* esta possibilidade; tal exigiria num estudo audiométrico aprofundado, de difícil exequibilidade, que não corresponde aos objectivos do presente estudo.

Coloquemos de lado, para já, a questão da pior *performance* das letradas e semiletradas na condição controlo, e concentremo-nos no problema

do decréscimo de *performance* provocado pela introdução de ruído. Aqui é possível descortinar maneira de resolver se os relativamente piores resultados das iletradas e semiletradas proviriam de factores gerais, tipo deficiência auditiva de base, ou pelo contrário, se deverão a factores especificamente grupais relacionados com a escolarização. Trata-se de tentar eliminar *a posteriori* a diferença intergrupar na linha de base; foram exploradas duas vias, uma por controlo dos *items* e outra por emparelhamento dos sujeitos em função do desempenho na condição controlo\*. Apreciemos uma e outra.

#### Reanálise do efeito das condições por controlo dos *items*

Trata-se de analisar o efeito das manipulações experimentais (ruído contralateral e binaural) apenas nos *items* que cada sujeito identificou correctamente em monaural. O *rationale* para este procedimento

---

\* A utilização de medidas de decréscimo ponderadas pelo nível global de desempenho, à imagem do que é feito em estudos de lateralidade, não nos parece adequada neste caso, por razões de ordem conceptual e empírica. Estas medidas, como o *f* de Marshall *et al.* (1979) ou o  $\lambda$  de Bryden e Sprott (1983), visam fornecer um valor de diferença que não esteja artefactualmente limitado pela *performance* global (como é óbvio, se um indivíduo desempenha certas tarefas ou muito bem ou muito mal, a diferença de *performance* entre elas será necessariamente reduzida; em termos gerais, a magnitude da diferença está dependente das magnitudes originais). Para os nossos dados, o cálculo do  $\lambda$  só é possível para alguns sujeitos, visto outros terem alcançado numa das condições a *performance* máxima. O score *f* transforma os dados brutos de modo tão drástico que utilizá-lo só nos pareceria válido se dispuséssemos de uma teoria convincente sobre a sua adequação ao fenómeno em foco. Para apreciar a transformação induzida pelo score *f*, tomemos o caso dos sujeitos iletrada 3, I<sub>3</sub>, e letrada L<sub>3</sub>. A primeira deu 38 respostas correctas em monaural e apenas cinco em binaural; os valores correspondentes para L<sub>3</sub> são 39 e 23. O decréscimo bruto de I<sub>3</sub> é aproximadamente o dobro do de L<sub>3</sub>. Ora os respectivos scores *f* são idênticos, .89, visto aqueles valores brutos serem ponderados pelo número de erros. Este resultado é contra-intuitivo e a sua utilização parece-nos injustificada. De facto, na presente experiência o decréscimo de *performance* entre as condições não é, cremos, psicologicamente análogo aos efeitos de lateralidade para os quais o score *f* foi desenvolvido. Nestes, pretende avaliar-se até que ponto o desempenho em certa tarefa é afectado por uma variável (ouvido); na nossa experiência pretende averiguar-se não o efeito do valor 1 ou 2 na tarefa A, mas sim o efeito relativo das tarefas A e B. O significado psicológico do nível geral de *performance* não é equivalente nas duas situações e usá-lo como critério do tipo de ponderação a usar (erros ou respostas correctas) é, no mínimo, discutível.

**QUADRO 2:** Percentagem média de decréscimo relativamente aos *items* correctamente identificados em monaural, por sujeito, nas condições de ruído contralateral e de ruído binaural, por grupo e ordem\* .

	Iletradas	Semiletradas	Letradas
Ordem 1			
R. Contral.	90.7	92.7	92.2
R. Binaural	37.8	46.9	50.1
Ordem 2			
R. Contral.	90.6	87.4	94.5
R. Binaural	17.2	22.7	31.4

\* Por exemplo, se o sujeito tivesse identificado correctamente 38 palavras em monaural, e dessas 38 apenas 15 em ruído binaural a sua percentagem de decréscimo foi de  $15/38 = 39.5\%$ . Estes cálculos foram efectuados individualmente, após o que se calculou a média por grupo e condição.

é muito simples: apesar de os estímulos terem sido escolhidos com base no pré-teste e constituírem parte do vocabulário comum da língua portuguesa, é possível que idiossincrasias individuais e de grupo levem a que certos estímulos sejam incorrectamente percebidos. De facto, uma análise dos erros por *item* revelou um nítido conjunto de estímulos, "pomba", "bela", "garra", "papa" e "cacho" que as iletradas sistematicamente erraram, mesmo em monaural: para o total das 5 palavras, 47.5% das respostas são erradas ("papa" era ouvido como "pata", e "cacho" como "tacho", por exemplo); o valor correspondente para as letradas é 11.4% e para as semiletradas 22.9%. Aquelas idiossincrasias não são controláveis e dificultam uma interpretação inequívoca dos resultados em termos do efeito pernicioso da manipulação experimental.

Reanalizando os dados nesta perspectiva, é possível encarar as percentagens de respostas correctas em ruído contralateral e binaural relativamente ao número total de palavras identificadas em monaural como uma medida de resistência ao efeito do ruído (cf. Quadro 2). Estes valores, a que chamaremos "percentagem de decréscimo", são muito semelhantes às percentagens de respostas correctas relativamente ao total dos estímulos, e têm a vantagem de permitir fazer análises estatísticas eliminando o factor trivial condição. Uma análise de variância às percentagens de decréscimo em ruído binaural, previamente submetidas à transformação arcoseno, com grupo e ordem como factores intersujeito, revelou serem significativos os dois efeitos principais e não haver entre eles interacção. Para o efeito de grupo,  $F_{(2,90)}=8.27$ ,  $p=.0005$ ; para a ordem,  $F_{(1,90)}=56.59$ ,  $p=.0001$ .\* Só as comparações entre letradas e iletradas, por um lado e semiletradas e iletradas, por outro, eram significativas (testes Tukey a  $\alpha=.05$ ), replicando assim os resultados obtidos na primeira análise.

---

\*Para resultados pormenorizados, cf. Anexo IP.2/3.

**QUADRO 3:** Resumo das características etárias e de desempenho dos subgrupos emparelhados. Apresenta-se o número médio (desvio-padrão entre parêntesis) de respostas correctas por condição, monoural (MON.), ruído contralateral (CONT.) e binaural (BIN.), nesta ordem (ORDEM I) ou na ordem contrária (ORDEM II).

	ORDEM I			ORDEM II		
	ILET. n = 8	SLET. n = 8	LET. n = 8	ILET. n = 11	SLET. n = 11	LET. n = 11
IDADE	53.5 (4.6)	46.1 (6.2)	53.5 (2.8)	48.4 (3.5)	48.9 (8.1)	48.2 (7.8)
MON.	37.1 (1.5)	37.3 (1.0)	37.3 (1.5)	37.8 (1.1)	37.9 (0.9)	37.8 (1.5)
CONT.	35.6 (1.8)	36.6 (1.9)	35.9 (2.4)	35.5 (2.0)	33.5 (4.0)	36.8 (2.3)
BIN.	15.1 (3.0)	18.8 (5.0)	21.8 (4.2)	7.7 (3.8)	9.4 (4.7)	11.5 (5.5)

Uma análise semelhante, tendo por variável dependente a percentagem de decréscimo em ruído contralateral (também transformada), mostrou não haver quaisquer efeitos significativos\*. Assim, o resultado prévio de diferenças intergrupo na condição de ruído contralateral pode atribuir-se a diferenças pré-existentes no reconhecimento em monaural; a este propósito pode ver-se no Quadro 3, mais à frente, a reduzida variabilidade intergrupo no desempenho em ruído contralateral, quando são considerados apenas subgrupos emparelhados por exactidão em monaural.

#### Reanálise do efeito do ruído por emparelhamento na condição monaural

A segunda via consiste em criar grupos emparelhados quanto à exactidão em monaural e, tanto quanto possível, nas restantes variáveis (nomeadamente a idade). Foi possível constituir subgrupos de sujeitos com aproximadamente a mesma idade e o mesmo nível de desempenho em monaural, em número de oito por grupo na ordem I e onze por grupo na ordem II (cf. Quadro 3).

Uma análise de variância às respostas correctas na condição binaural, com grupo e ordem como factores intersujeito, replicou os resultados anteriores: os dois efeitos principais são significativos, para a ordem,  $F_{(1,51)}=55.98$ ,  $p<.0001$  e para o grupo,  $F_{(2,51)}=5.75$ ,  $p=.0056$ , não havendo interacção\*\*. Comparações de médias por testes Tukey ( $\alpha=.05$ ) mostraram que só a diferença entre letradas e iletradas é significativa; as semiletradas situam-se a meio caminho entre aqueles dois grupos, sem deles diferir significativamente. Uma análise análoga à exactidão em ruído

\*Para resultados pormenorizados, cf. Anexo IP.2/4.

\*\*Para pormenores, cf. Anexo IP.2/5.

**QUADRO 4:** Tempo médio (em segundos) necessário para a leitura de uma lista de palavras (Coluna 1, C<sub>1</sub>: monossílabos; C<sub>2</sub>: dissílabos) consoante o grau de escolarização. Desvio-padrão e amplitude são apresentados entre parêntesis (d.p.; mínimo-máximo).

	SEMILETRADAS		LETRADAS	
C <sub>1</sub>	17.0 (9.5; 7-40)	n = 24	8.7 (2.3; 6-14)	n = 26
C <sub>2</sub>	22.7 (16.1; 10-75)	n = 18	9.9 (2.1; 7-17)	n = 26
Total	37.7 (24.3; 18-110)	n = 18	18.6 (4.0; 13-30)	n = 26

contralateral não mostrou quaisquer efeitos significativos, replicando deste modo o resultado obtido por controlo dos *items*\*.

Estas reanálises dos dados permitem salientar dois resultados importantes. Primeiro, confirmou-se que o efeito do ruído binaural é mais pernicioso para as iletradas do que para as letradas. A posição das semiletradas não é tão clara; todavia, a maior parte das análises feitas sugere que o seu desempenho se aproxima mais do das letradas, nesta condição de ruído binaural. Assim, o facto de se saber ler ajuda a identificar palavras faladas em condições difíceis, como quando aparecem no meio de ruído. Segundo, nas duas análises em que o nível de desempenho na condição controlo foi explicitamente tomado em conta, não se observou efeito do ruído contralateral em nenhum dos grupos. Este resultado indica claramente que os sujeitos são capazes de dirigir a sua atenção para a palavra sem serem afectados pelo ruído, quando ambos surgem de fontes espaciais distintas. Em situações como esta, em que as mensagens em competição são diferentes quanto à natureza (fala *vs* ruído) e localização, não se encontra indicação de que a atenção selectiva dependa do nível de escolarização.

#### 4.3.3. Fluência de leitura

Conforme se pode observar no Quadro 4, além de as semiletradas demorarem consideravelmente mais tempo a ler a lista de palavras, elas apresentam resultados muito mais variáveis. De facto, a relativa heterogeneidade do grupo das semiletradas não é surpresa, ela constituiu justamente um dos motivos para se realizar este pré-teste.

\*Para pormenores, cf. Anexo IP.2/6.

**QUADRO 5:** Número médio de dígitos correctamente repetidos (desvio-padrão entre parêntesis) no teste de memória de dígitos, consoante o grau de escolarização. Na ordem directa os algarismos são repetidos tal como foram ouvidos, na ordem inversa são repetidos ao contrário.

	ILETRADAS n = 36	SEMILETRADAS n = 28	LETRADAS n = 32
ORDEM DIRECTA	4.0 (0.7)	5.0 (0.8)	6.4 (1.4)
INVERSA	2.9 (0.7)	3.4 (0.6)	4.8 (1.4)

A inspecção dos resultados individuais revela que as semiletradas com maior fluência de leitura são as que apresentam melhores resultados na condição de ruído binaural. De facto, existe uma correlação negativa,  $r=-.52$  ( $p=.02$ ), entre o tempo de leitura e o número de respostas correctas em binaural. Esta relação não ocorre nas letradas (o valor correspondente é de  $r=.08$ ); repare-se porém que o teste de leitura é, para estes sujeitos, pouco discriminativo. Estes resultados indicam que, entre determinados limites, a proficiência no uso da escrita pode ajudar a identificação de palavras orais.

#### 4.3.4. Amplitude de memória imediata

Diferenças consoante o grupo estão patentes no teste de memória de dígitos (cf. Quadro 5). Uma análise de variância, com grupo como variável intersujeito e tarefa (ordem directa *vs* inversa) como variável intrasujeito, confirmou que ambos os efeitos principais são significativos, não havendo interacção (para o grupo,  $F_{(2, 77)}=41.26$ ,  $p=.0001$ ; para a tarefa,  $F_{(1, 77)}=223.3$ ,  $p=.0001$ )\*. Comparações de médias através de testes Tukey ( $\alpha=.05$ ) revelaram que só as diferenças entre letradas e iletradas por um lado, e letradas e semiletradas, por outro, são significativas.

Verifica-se pois um efeito do nível de escolarização na amplitude de memória imediata de dígitos, que replica claramente os achados de Read e Ruyter (1985).

Devido à variação de idade intra e intergrupo, foram calculadas correlações entre os *scores* de memória de dígitos e idade. Só no grupo das letradas se obtiveram correlações significativas, negativas:  $r=-.52$  ( $n=18$ ,  $p=.009$ ) para a tarefa em ordem directa e  $r=-.41$  ( $n=18$ ,  $p=.04$ ) em ordem inversa. Os valores correspondentes são, respectivamente,  $-.02$  e  $-.12$  para as

\*Para pormenores, cf. Anexo IP.2/7.

iletradas e -.03 e -.04 para as semiletradas. Talvez este resultado traduza o facto de que as letradas, por um lado, e as semiletradas e iletradas, por outro, não recorrem aos mesmos processos para realizar a tarefa.

Segundo Read e Ruyter (*ibidem*), os piores resultados dos adultos pouco escolarizados dever-se-iam a uma dificuldade de recodificação fonológica (ou seja, estabelecer uma representação do estímulo em termos verbais). Não conseguindo criar e manter uma representação fonológica da palavra, temporalmente estável, estes sujeitos teriam piores resultados não só no teste de memória de dígitos, como noutros dependentes da existência dessa representação, por exemplo, em tarefas de análise consciente da fala. De facto, aqueles autores relatam correlações significativas entre os *scores* no teste de memória de dígitos e os resultados em tarefas de segmentação.

A relação entre a chamada memória verbal a curto prazo e a consciência da estrutura fonética da fala tem vindo a ser apontada por vários investigadores no contexto das dificuldades de aprendizagem da leitura: os "maus" leitores, em quem aquela consciência ("phonetic awareness") é deficiente, têm piores resultados em várias tarefas de memória verbal a curto termo (*e.g.*, Mann, 1985; Mann, Liberman e Shankweiler, 1980; cf. também Shankweiler e Crain, 1986). Uma interpretação proposta é que estas diferenças em tarefas de memória a curto prazo estejam por sua vez associadas a um uso menos efectivo da representação fonológica\* (Mann e Liberman, 1984). Na mesma linha, tem sido sugerido que dificuldades de processamento perceptivo fonético poderiam estar na base dos piores resultados em tarefas de memória a curto prazo (Brady, Shankweiler e Mann, 1983). Estes autores buscam em seu apoio o achado de Rabbitt (1968) de que factores que afectam a qualidade da estimulação exercem uma

\*De facto, tem sido proposto que as dificuldades de aprendizagem da leitura estão associadas a deficiências de recodificação fonológica (*e.g.*, Katz, 1986; Mark, Shankweiler e Liberman, 1977), mas este ponto de vista tem suscitado algumas objecções (Alegria, Pignot e Morais, 1982; Hall, Wilson, Humphreys, Tinzman e Bowyer, 1983).

influência negativa não só na identificação imediata como também na evocação retardada dessa estimulação.

Rabbitt (*ibidem*) verificou que é mais difícil lembrar dígitos quando eles são apresentados em ruído (Experiência 1); há também mais erros na evocação de dígitos nas primeiras posições de uma lista quando os dígitos subsequentes são apresentados com ruído (Experiência 2); um efeito análogo verifica-se com trechos em prosa: dão-se mais erros ao responder a perguntas sobre o conteúdo factual da primeira metade do trecho se a segunda metade tiver sido apresentada com ruído (Experiência 3). A conclusão é que a maior dificuldade em reconhecer a fala provocada pelo ruído interfere com as actividades necessárias para reter essa informação na memória (p. 241). Estes resultados conferem efectivamente alguma plausibilidade à ideia de que problemas a nível de elaboração da representação fonética se prolongariam em dificuldades de manutenção dessa mesma representação.

Supondo que, para identificar palavras embebidas em ruído, o sujeito tem de elaborar uma representação fonética a partir da estimulação ambígua, e que as diferenças no teste de memória de dígitos estão associados à elaboração e manutenção dessa representação, então deveríamos encontrar uma associação entre estas duas medidas. Calculámos pois as correlações entre o desempenho na condição de ruído binaural e os *scores* de memória de dígitos por grupo e ordem. Observou-se uma correlação significativa apenas nos sujeitos iletrados que fizeram a condição binaural no princípio,  $r=.48$ ,  $p=.03$  para o teste em ordem directa e  $r=.44$ ,  $p=.04$  para o teste em ordem inversa\*. Portanto, estes resultados não são consistentes com a interpretação acima proposta. A única conclusão que deles se extrai é que a variação individual quanto à memória imediata de dígitos não esteve

\* Para pormenores, cf. Anexo IP.2/8.

em geral associada à melhor ou pior *performance* na tarefa de identificação das palavras em ruído. Na falta de um quadro interpretativo sólido, limitamo-nos a referir a excepção à regra, as iletradas que ouviam as palavras sem familiarização prévia.

Além da recodificação fonética, têm sido apontados outros factores explicativos quanto às diferenças de amplitude de memória imediata (cf. Pinto, 1986); sem dúvida valeria a pena averiguar se as diferenças entre indivíduos com nula ou reduzida escolarização por um lado e indivíduos escolarizados, por outro, estariam associados à utilização de estratégias activas de manipulação dos *items*, como o agrupamento, ou à velocidade de identificação dos próprios *items*, ou ambos.

#### 4.3.5. Análise das respostas incorrectas

Analisaremos os erros observados na identificação das palavras em ruído binaural, considerando separadamente os resultados obtidos na ordem I - palavras previamente ouvidas em monaural, e na ordem II - sessão experimental iniciada pela audição binaural. Esta estratégia justifica-se porque, além do nível de desempenho ter sido muito diferente consoante a ordem, a familiarização prévia com as palavras poderá favorecer a ocorrência de certo tipo de erros em detrimento de outros.

Tipo de respostas incorrectas na identificação de palavras em ruído binaural com audição prévia noutras condições (ordem I)

O Quadro 6 mostra a distribuição a várias respostas incorrectas. Dispuseram-se primeiro aquelas em que o grau de aproximação à palavra apresentada é mínimo, as respostas de "ininteligível" (quando o sujeito

**QUADRO 6:** Incidência dos tipos de erro (percentagem média e desvio-padrão entre parêntesis) na identificação de palavra CVCV mascarada por ruído binaural, por grupo, na ordem I. As percentagens foram calculadas relativamente ao total de erros por sujeito.

	ILETRADAS n=16		SEMILETRADAS n=14		LETRADAS n=14	
ININT.	24.1	(18.3)	20.5	(12.9)	26.2	(20.9)
G1	13.7	(8.9)	5.1	(4.8)	8.0	(5.3)
C <sub>1</sub>	20.8	(8.8)	32.3	(13.2)	25.0	(11.2)
C <sub>1</sub> V <sub>2</sub>	4.4	(4.2)	6.4	(4.9)	9.6	(6.8)
C <sub>1</sub> C <sub>2</sub>	14.2	(4.7)	14.6	(5.3)	13.0	(10.3)
C <sub>1</sub> C <sub>2</sub> V <sub>2</sub>	4.9	(4.6)	3.2	(2.7)	5.4	(5.9)
C <sub>2</sub>	9.2	(4.4)	9.0	(6.9)	9.9	(6.4)
V <sub>1</sub> (V <sub>2</sub> )	0.4	(1.4)	1.9	(2.8)	0.6	(1.6)
C <sub>1</sub> V <sub>1</sub>	1.4	(1.8)	2.2	(3.8)	0.3	(1.0)
C <sub>2</sub> V <sub>2</sub>	5.5	(5.1)	4.0	(3.3)	1.8	(2.5)
Outros	0.3	(0.1)	0	-	0.3	(1.0)

Consideram-se as categorias ININT ("ininteligível"), G1 (erros globais, em pelo menos os três segmentos iniciais), C<sub>1</sub> (erros exclusivamente na primeira consoante), C<sub>1</sub>V<sub>2</sub> (acompanhados por erro na segunda vogal), C<sub>1</sub>C<sub>2</sub> (nas duas consoantes), C<sub>1</sub>C<sub>2</sub>V<sub>2</sub> (*idem*, com erro de segunda vogal), C<sub>2</sub> (erros de segunda consoante), V<sub>1</sub>(V<sub>2</sub>) na primeira vogal só, ou nas duas), C<sub>1</sub>V<sub>1</sub> e C<sub>2</sub>V<sub>2</sub> (na primeira e segunda sílabas). Os restantes erros foram agrupados na categoria Outros.

relata "não perceber nada") e os erros em pelo menos três segmentos: esta última categoria, denominada erros globais, compreende também os erros que não respeitaram a estrutura CVCV dos estímulos\*. Seguem-se os erros envolvendo maior ou menor grau de partição silábica, como erros na primeira consoante exclusivamente,  $C_1$ , ou extensivos à segunda vogal,  $C_1V_2$ , ou em ambas as consoantes,  $C_1C_2$ , etc. Finalmente apresentam-se os erros que preservam a integridade silábica, erros de primeira sílaba,  $C_1V_1$ , ou de segunda sílaba,  $C_2V_2$ .

Tal como nas experiências anteriores, considera-se não o número bruto de erros por categoria, mas sim a sua incidência tendo em conta o nível de desempenho. Podemos assim comparar mais facilmente a distribuição do tipo de erros em sujeitos com diferentes níveis de *performance*, e apreciar o peso relativo dos vários tipos de erro na caracterização do desempenho.

Calculámos a proporção de cada tipo de erro relativamente ao total de respostas incorrectas exibindo por cada sujeito. Uma opção alternativa teria sido a de considerar cada tipo de erro relativamente ao total de respostas erradas, excluindo as respostas de "ininteligível". Convém pois esclarecer melhor o que são estas respostas. Tipicamente, o sujeito dizia "desta vez é que não consegui entender mesmo nada", ou abanava negativamente a cabeça com um ar de impotência. Como a situação experimental previa que se pudesse voltar a ouvir o estímulo no caso de distração momentânea, e atendendo à própria reacção dos sujeitos, parecemos claramente injustificável atribuir estas respostas em geral a falta de atenção, e discutível falar de "ausência de resposta". Após trabalharmos nesta situação com dezenas de sujeitos, cremos mais conforme ao espírito do

---

\* Ao contrário da situação de audição dicótica, neste caso de audição de palavras embebidas em ruído não há possibilidade de observar fenómenos de integração interaural, como fusões fonológicas.

que se observou considerar o "não percebi!" como mais do que uma simples ausência de resposta.

Claro que pelo menos algumas destas reacções de "não percebi" esconderão situações onde os sujeitos se terão apercebido de um som qualquer, mas que terão preferido não relatar; enfim, supondo que seria exequível "obrigar" os sujeitos a dizer alguma coisa que imitasse o seu percepto, é razoável supor que estas "respostas" fossem de vários tipos, e se distribuíssem pelas restantes categorias consideradas. Nesta perspectiva, consideramos apenas dois grandes tipos de respostas: aquelas em que o sujeito identifica correctamente o estímulo; e aquelas em que isso não acontece, analisando a incidência relativa das várias categorias de respostas incorrectas tendo em conta o conjunto complementar das respostas correctas. Tentamos assim dar conta do conjunto das observações feitas: as primeiras (respostas correctas) foram já analisadas no ponto 3.1; as segundas (respostas in-correctas) são integradas num quadro interpretativo comum na análise que se seguirá. Esta opção, além de mais parcimoniosa, tem a vantagem pragmática de não reduzir à partida o número de observações disponíveis e assim de não tornar o cálculo das proporções demasiado afectado por casos extremos\*.

Para os três grupos, a maioria das respostas erradas concentra-se nas categorias "Ininteligível", e  $C_1$ ,  $C_2$ , e  $C_1C_2$ , ou seja, em erros de consonantes. Estas categorias dão conta de 68% dos erros nas iletradas, 76% nas semiletradas, e 74% nas letradas. Erros restringidos às vogais foram particularmente raros nos três grupos. Em contrapartida, identificações

---

\*Por exemplo, um sujeito com 15 respostas de "ininteligível" e 2 erros globais teria uma incidência de 100% de erros globais se optássemos por dar às respostas de "ininteligível" um estatuto equivalente ao das respostas correctas ou das respostas erradas. É mais parcimonioso descrever a *performance* deste sujeito como: 57.5% de respostas correctas; das restantes 42.5% de respostas, 88% são de "ininteligível" e 12% são erros globais.

erróneas que preservam apenas a primeira vogal (erros  $C_1C_2V_2$ ) foram relativamente frequentes.

A existência de erros ininteligíveis poderia ser devida à perda de acuidade auditiva associada à idade; todavia, a correlação entre idade e proporção de erros ininteligíveis foi praticamente nula:  $r=.02$ ,  $n=44$ ,  $p=.44^*$ . Em contrapartida existe uma forte correlação negativa entre a exactidão e a proporção destes erros: nas iletradas,  $r=-.69$ ,  $n=16$ ,  $p=.002$ , nas semiletradas,  $r=-.42$ ,  $n=14$ ,  $p=.065$  e nas letradas  $r=-.74$ ,  $n=14$ ,  $p=.001$  (fizeram-se os cálculos por grupo pois o nível de exactidão depende do grupo).

Noutras categorias de erro notam-se diferenças grupais. Optámos por agrupar os erros  $C_1V_1$  e  $C_2V_2$  numa categoria única, que referiremos como erros silábicos. Para algumas análises agruparemos os erros silábicos com os globais, designando esta categoria conjunta por "global". De facto, trata-se de erros que não envolvem partição silábica e que provavelmente se podem atribuir a uma semelhança global com o estímulo; além disso, este agrupamento permitir-nos-á comparar as observações da Experiência III com as do presente estudo. Os erros  $C_1$  e  $C_1V_2$  serão agrupados na categoria  $C_1(V_2)$ , ou de erros de primeira consoante. Este último agrupamento justifica-se pelas características do português em geral, e das palavras usadas em particular: a vogal final não é acentuada e mesmo na pronúncia de palavras isoladas ela tende a ser reduzida à vogal central neutra  $[\ə]$ \*\* . Se a "palavra" resultante de um erro  $C_1$  não tiver sentido (e.g., troca do /g/ por /k/ em "garra"), o sujeito terá tendência a errar também a vogal final, acusticamente fraca, ouvindo "carro". Supomos pois

\*O cálculo foi feito para todos os sujeitos por não haver diferença intergrupo quanto à proporção de erros ininteligíveis.

\*\*Análises espectrográficas de fala corrente mostram mesmo que a segunda vogal não tem existência acústica (cf. Prudêncio, Rebelo, Atalaia Costa, Lacerda Marques, Namorado e Delgado Martins, 1978).

que os erros  $C_1V_2$  são essencialmente erros  $C_1$  que por constrangimentos lexicais e ambiguidade acústica foram acompanhados por mudança da vogal final.

Uma análise de variância às proporções (transformadas\*) de erros  $C_1(V_2)$ , e erros silábicos mais globais, com grupo como variável intersujeito e tipo de erro como variável intrasujeito (segmental *vs* global) mostrou que a interação entre grupo e tipo de erro é significativa ( $F_{(2,41)}=6.99, p=.002$ \*\*). Análises separadas por tipo de erro confirmaram que os erros  $C_1(V_2)$  variam significativamente com o grupo ( $F_{(2,41)}=3.93, p=.027$ ); comparações através de testes Tukey ( $\alpha=.05$ ) mostraram que só atinge significância a diferença entre iletradas, com cerca de 25% de erros  $C_1(V_2)$  e semiletradas, com cerca de 39%. Apesar de as letradas terem uma proporção destes erros muito próxima das semiletradas, de cerca de 35%, devido à grande variabilidade intragrupo a diferença relativamente às iletradas não é significativa.

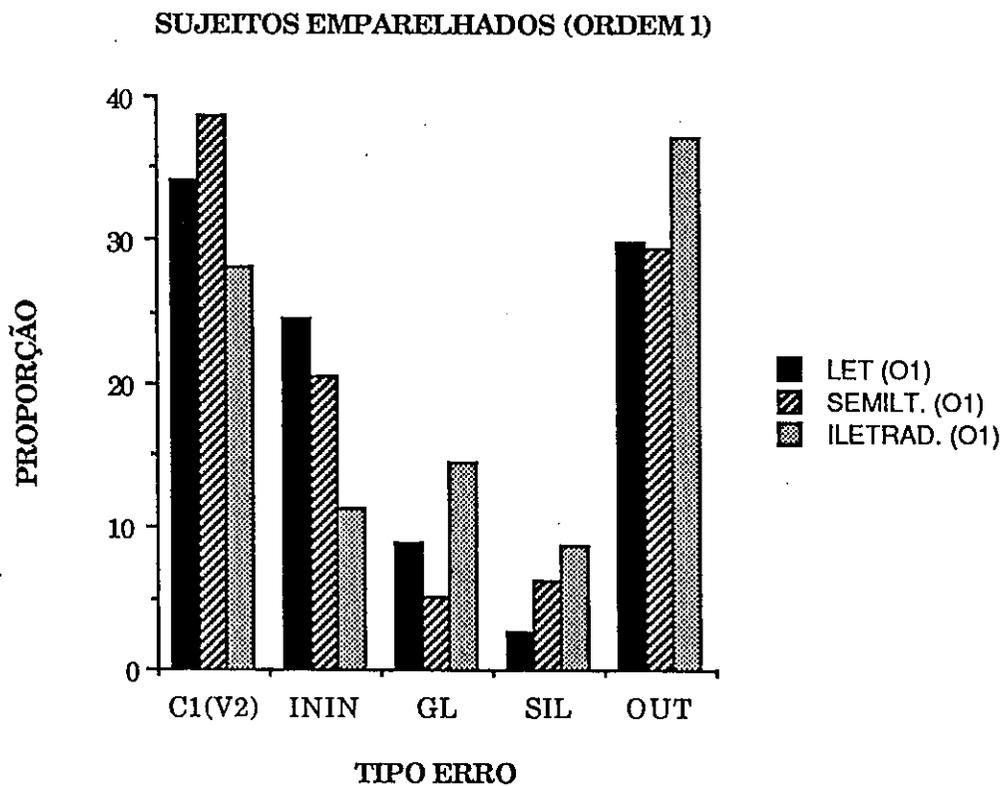
Quanto aos erros globais e silábicos, uma análise de variância análoga à anterior confirmou o efeito de grupo,  $F_{(2,41)} = 5.49, p = .007$ . Testes Tukey confirmaram que as iletradas dão significativamente mais erros deste tipo (21%) do que semiletradas (11%) e letradas (10%)\*\*\*.

O nível de desempenho está também associado à proporção de erros de primeira consoante: quanto melhor é o desempenho, maior é a

\*Excepto se indicado em contrário, todos os valores apresentados se referem a cálculos efectuados com base em valores transformados (arcoseno).

\*\*Para pormenores, cf. Anexo IP. 2/9; para as duas análises seguintes, cf. Anexos IP. 2/10 e IP2/11.

\*\*\*Os resultados são semelhantes considerando separadamente os erros globais ( $F_{(2,41)}=4.49, p=.017$ ) ou os erros  $C_1$  ( $F_{(2,41)}=4.02, p=.025$ ). Para os erros silábicos, propriamente ditos ( $F_{(2,41)}=4.75, p=.13$ ).



**Figura 3:** Proporção de erros de primeira consoante  $C_1(V_2)$ , ininteligíveis, ININT., globais, GL, silábicos, SIL, e outros OUT, em letradas ( $n = 7$ ), semiletradas ( $n = 14$ ) e iletradas ( $n = 7$ ) com o mesmo nível de desempenho ( $\approx 46\%$ ) na identificação de palavras embebidas em ruído binaural (ordem I).

incidência daqueles erros; nas iletradas,  $r = .47$ ,  $p = .031$ , nas semiletradas,  $r = .69$ ,  $p = .003$ , nas letradas,  $r = .70$ ,  $p = .003^*$ .

O facto de o nível de desempenho estar associado à preponderância de erros de primeira consoante, e de ambos variarem com o grupo, torna necessário que se comparem subgrupos com nível de desempenho semelhante. Infelizmente, a estrutura dos resultados dificulta esta comparação: foi possível emparelhar apenas as sete melhores iletradas com sete letradas com uma percentagem semelhante de respostas correctas (respectivamente, 45% e 46.8%, d.p. correspondentes 9.2 e 10.2).

As semiletradas ( $n=14$ ) tiveram um nível de exactidão praticamente idêntico ao deste subgrupo de sujeitos, cerca de 45%. A comparação dos seus resultados com os das letradas e iletradas não está pois enviesada pelo nível de desempenho, e a posição relativa das semiletradas face aos outros dois subgrupos é esclarecedora quanto ao papel do grau de escolarização (cf. Figura 3). Uma análise de variância análoga à anterior, com grupo como variável intersujeito, e tipo de erro, segmental,  $C_1(V_1)$  vs global (silábico mais global propriamente dito) como variável intrasujeito, replicou a interacção entre grupo e tipo de erro,  $F_{(2,25)}=3.55$ ,  $p=.043^{**}$ . Análises separadas por tipo de erro confirmaram que as iletradas dão mais erros globais e silábicos, cerca de 21%, do que as semiletradas e letradas, ambas 11%: para o efeito de grupo,  $F_{(2,25)}=4.14$ ,  $p=.027$ , comparações feitas por testes Tukey. Porém, o efeito de grupo não atingiu a significância quanto aos erros de primeira consoante,  $F_{(2,25)}=1.59$ ,  $p=.22$ . Tendo em conta o

\*Os resultados são semelhantes considerando apenas os erros  $C_1$  em vez dos erros  $C_1(V_2)$ :  $r=.36$ ,  $p=.08$ ,  $r=.64$ ,  $p=.007$  e  $r=.61$ ,  $p=.009$  nas iletradas, semiletradas e letradas, respectivamente. A correlação entre o nível de desempenho e a proporção de erros globais e silábicos é negativa e significativa no caso das semiletradas,  $r=-.62$ ,  $p=.009$ ; nas iletradas, é positiva e significativa,  $r=.43$ ,  $p=.046$ . Talvez este último facto, à primeira vista insólito, se deva a que as iletradas com melhores resultados são aquelas que "arriscam" mais, dizendo sempre "qualquer coisa".

\*\* Cf. Anexo IP.2/12. Para as análises seguintes, cf. Anexos IP. 2/13 e IP. 2/14.

**QUADRO 7:** Incidência dos tipos de erro (percentagem média e desvio-padrão entre parêntesis) na identificação de palavra CVCV mascarada por ruído binaural, por grupo, na ordem II. As percentagens foram calculadas relativamente ao total de erros por sujeito.

	ILETRADAS n=20		SEMILETRADAS n=14		LETRADAS n=18	
ININT.	57.7	(29.0)	31.4	(19.0)	36.8	(22.4)
GL	11.7	(13.4)	16.7	(11.4)	15.1	(9.0)
C <sub>1</sub>	8.8	(9.9)	13.2	(9.5)	17.8	(10.9)
C <sub>1</sub> V <sub>1</sub>	0.8	(1.9)	1.2	(2.1)	2.2	(3.3)
C <sub>1</sub> C <sub>2</sub>	6.5	(6.9)	12.5	(7.3)	9.0	(5.6)
C <sub>1</sub> C <sub>2</sub> V <sub>2</sub>	1.5	(2.2)	3.3	(3.1)	3.3	(2.7)
C <sub>2</sub>	4.6	(4.6)	5.4	(4.2)	5.0	(4.2)
V <sub>1</sub> (V <sub>2</sub> )	1.1	(2.1)	3.5	(4.8)	0.9	(1.9)
C <sub>1</sub> V <sub>1</sub>	0.5	(1.3)	2.6	(3.7)	2.6	(3.9)
C <sub>2</sub> V <sub>2</sub>	3.2	(4.5)	4.3	(2.2)	3.4	(3.7)
Outros	3.3	(4.3)	5.9	(3.7)	3.8	(3.9)

Consideram-se as categorias ININT (ininteligível), GL (erros globais, em pelo menos os três segmentos iniciais), C<sub>1</sub> (erros exclusivamente na primeira consoante), C<sub>1</sub>V<sub>2</sub> (acompanhados por erro na segunda vogal), C<sub>1</sub>C<sub>2</sub> (nas duas consoantes), C<sub>1</sub>C<sub>2</sub>V<sub>2</sub> (*idem*, com erro de segunda vogal), C<sub>2</sub> (erros de segunda consoante), V<sub>1</sub>(V<sub>2</sub>) na primeira vogal só, ou nas duas), C<sub>1</sub>V<sub>1</sub> e C<sub>2</sub>V<sub>2</sub> (na primeira e segunda sílabas). Os restantes erros foram agrupados na categoria Outros.

reduzido número de sujeitos considerados, e para obviar a falta de poder estatístico daí decorrente, fez-se o seguinte: depois de assegurarmos que entre letradas e semiletradas não havia diferenças significativas quanto a estes erros, ( $t_{(19)}=.65, p=.52$ ), efectuámos uma comparação de sujeitos iletrados com sujeitos letrados em sentido lato (*i.e.*, letradas e semiletradas) através do teste *t*. Esta comparação mostrou que, com o mesmo nível de desempenho, sujeitos que conhecem a representação escrita da fala têm uma proporção significativamente maior de erros de primeira consoante 37%, do que sujeitos iletrados, 28% ( $t_{(26)}=1.78, p=.044$ , unicaudal)\*.

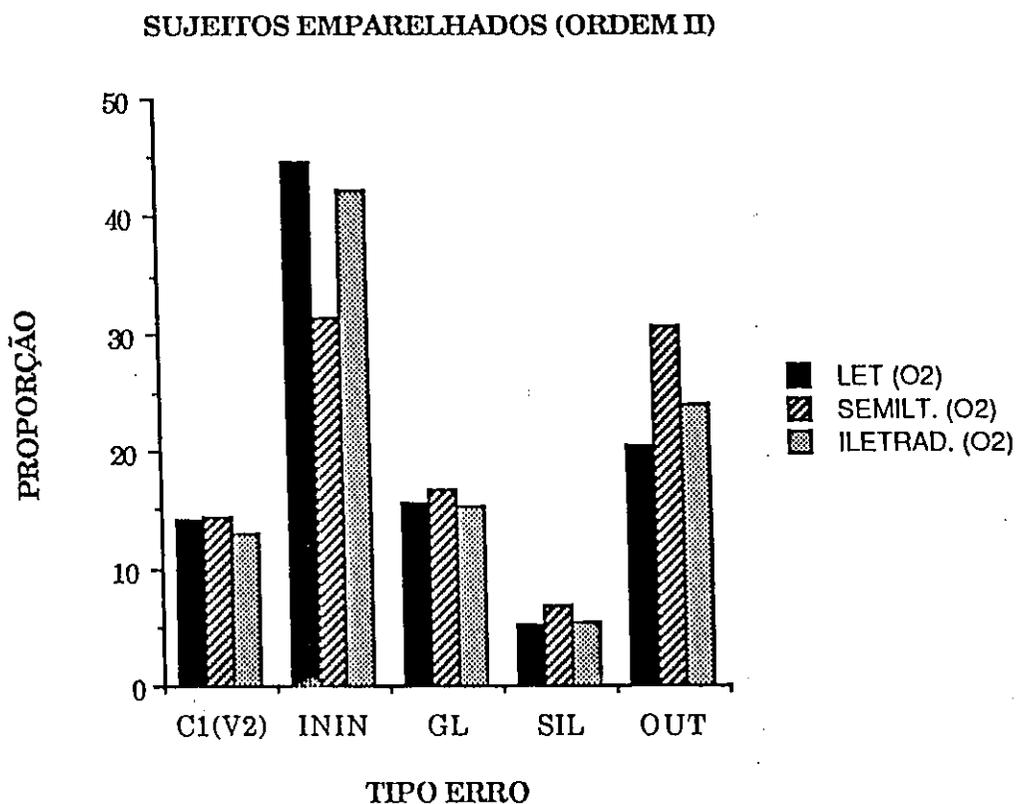
Tipo de respostas erradas na identificação de palavras em ruído binaural sem audição prévia (Ordem II)

A distribuição das respostas erradas na mesma tarefa, mas sem audição prévia dos estímulos (Ordem II), distingue-se pela evidente preponderância das respostas de "ininteligível" (cf. Quadro 7). Isto acontece em todos os grupos, sendo particularmente flagrante nas iletradas. A proporção destes erros não está associada à idade, em nenhum dos grupos\*\*, mas tem uma forte correlação negativa com a percentagem de respostas correctas (os valores de *r* são -.76, -.68 e -.61 para iletradas, semiletradas e letradas, respectivamente, todos com  $p \leq .01$ ). Existe igualmente uma clara associação entre exactidão e incidência de erros de primeira consoante,  $C_1(V_2)$ :  $r=.71, n=20, p=.001$  nas iletradas,  $r=.49, n=14, p=.036$  nas semiletradas e  $r=.85, n=18, p=.001$  nas letradas.

O facto de letradas e semiletradas darem cerca de menos 20% de respostas "ininteligível" do que as iletradas dificulta comparações

\* Considerando apenas os erros  $C_1$ , o resultado é semelhante:  $t_{(26)}=1.67, p=.053$ .

\*\* Nas iletradas,  $r=.03, n=20, p=.43$ , nas semiletradas,  $r=.25, n=14, p=.15$  e nas letradas  $r=.31, n=18, p=.13$ .



**Figura 4:** Proporção de erros de primeira consoante  $C_1(V_2)$ , ininteligíveis, ININT., globais, GL, silábicos, SIL, e outros, OUT, em letradas ( $n = 7$ ), semiletradas ( $n = 14$ ) e iletradas ( $n = 7$ ) com o mesmo nível de desempenho ( $\approx 24\%$ ) na identificação de palavras embebidas em ruído binaural (ordem II).

intergrupais quanto à incidência de outros erros; estes serão necessariamente menos frequentes neste último grupo. Além disso, diferenças inter-grupais quanto ao nível de *performance* levam-nos a circunscrever aquelas comparações a subgrupos de exactidão semelhante. Foi possível emparelhar 13 iletradas e 13 letradas por exactidão (percentagem de respostas correctas de 23.7 e 24%, respectivamente; os d.p. correspondentes são 9.6 e 10.8). Os vários tipos de erro têm uma incidência praticamente idêntica nos dois subgrupos (cf. Figura 4). Note-se que o nível de desempenho das semiletradas (grupo total) se situa na mesma faixa do dos subgrupos emparelhados, cerca de 22%, e os erros se distribuem também da mesma maneira. Assim, a única diferença intergrupais que emerge nesta situação é em termos de nível de desempenho. Sujeitos com melhores resultados dão menos respostas de "ininteligível" e mais erros de primeira consoante, independentemente de serem ou não alfabetizados.

#### 4.3.6. Observações adicionais

Respostas de tipo não-palavra ocorreram com relativamente pouca frequência em todos os grupos. Elas dão conta, em média, de cerca de 8% dos erros, considerando os três grupos e as duas ordens. Não são aparentes diferenças intergrupais quanto à incidência destas respostas e a inspecção dos resultados revela a existência de grande variação interindividual\*.

---

\*Os valores observados foram os seguintes: na ordem I, as letradas deram em média 1.5 não-palavra (d.p. = 1.51, entre 0 - 5), as semiletradas 1.6. (d.p. = 1.50, entre 0 - 5) e as iletradas 2.4 (d.p. = 2.63, entre 0 - 10). A proporção destas respostas relativamente ao total de erros é, em média e para cada grupo, 7%, 7% e 9%, respectivamente. Considerando apenas os subgrupos de iletradas e letradas com o mesmo nível de *performance* em binaural, a média de respostas não-palavra foi de 2.5 (d.p. = 2.6.) e 2 (d.p. = 1.7), respectivamente. Na ordem II, as letradas deram em média 1.8 não-palavras (d.p. = 1.9), as semiletradas, 2.4. (1.7) e as iletradas, 1.4 (2.0); estas respostas davam conta de 6%, 8% e 4% dos erros, respectivamente. Nos subgrupos de iletradas e letradas emparelhados por exactidão em binaural, as primeiras deram em média 2.2 respostas não-palavra (d.p. = 2.2) e as segundas 1.8 (d.p. = 2.7).

Uma análise sumária por *item*, considerando as respostas de todos os sujeitos na situação de ruído binaural e ordem I, mostrou que tende a haver mais erros nas palavras iniciadas por consoante sonora, em todos os lugares de articulação. Examinaram-se as respostas aos *items* cuja segunda consoante não era oclusiva\*. De salientar apenas um erro sistemático, observado nos *items* com a vibrante [R]: 47% dos erros consistiam na confusão com [z] ou [s]. Note-se que em fala corrente e em posição medial a vibrante dupla /r/ é realizada como fricativa. Trata-se pois de um erro quanto ao lugar de articulação, que ocorre frequentemente na audição de palavras embebidas em ruído (cf. Miller e Nicely, 1955; Sorin e Thouin-Daniels, 1983).

#### 4. Resumo e conclusão

Observámos como adultos letrados, semiletrados e iletrados identificaram palavras em três condições: sem ruído, com ruído contralateral, e misturadas em ruído. Em termos estritamente quantitativos, observou-se nas três condições um desempenho inferior por parte dos iletrados. Porém, as análises feitas permitem estabelecer que a diferença entre iletrados e letrados é mais pronunciada na condição de ruído binaural; elas indicam ainda que, nesta condição, o desempenho das semiletradas se aproxima mais do das letradas que das iletradas.

Os resultados sugerem pois um efeito da alfabetização nesta tarefa: sujeitos que dominam o código escrito conseguem perceber melhor palavras mascaradas por ruído. Encontra-se ainda indicação de que quanto melhor for aquele domínio, menos se é afectado pelo ruído (veja-se a correlação entre fluência de leitura e desempenho nas semiletradas; o facto

\*Um resumo detalhado encontra-se no Anexo IP.3/2.

de as letradas serem em geral superiores às semiletradas pode ter mais que ver com a própria experiência com a fala, por exemplo, maior riqueza lexical, do que com factores específicos da linguagem escrita).

O padrão de resultados observado quando os sujeitos identificavam a palavra apresentada a um ouvido, ignorando o ruído apresentado ao outro ouvido, permite concluir que (i) o efeito do ruído contralateral na inteligibilidade de palavras foi praticamente nulo; (ii) na medida em que o resultado anterior traduz o controlo efectivo da atenção, não se encontraram diferenças de alfabetização e escolarização a nível da atenção selectiva perante mensagens espacial e acusticamente distintas.

Finalmente, ficou claramente demonstrado que a experiência prévia - ou com as palavras, ou com a própria situação de escuta dicótica, com ou sem ruído - facilita o reconhecimento nas condições mais difíceis. Este efeito aconteceu nos três grupos e é evidente não só em termos de percentagem de respostas correctas, como também no padrão de respostas: o número bruto e a frequência relativa de respostas de "ininteligível" eram claramente inferiores se a identificação das palavras misturadas com ruído binaural fosse precedida pelas outras duas condições.

A análise do padrão de respostas incorrectas trouxe três esclarecimentos principais.

Primeiro, diferentes níveis de desempenho estão associados a diferentes padrões de erro: uma frequência relativamente elevada de erros silábicos e globais, e diminuta de erros segmentais, nomeadamente de primeira consoante, acontece em níveis de *performance* inferiores; melhores *performances* apresentam o padrão inverso.

Segundo, mesmo considerando níveis de desempenho semelhantes, verifica-se um efeito de alfabetização quando a situação de



reconhecimento não é demasiado difícil. Ou seja, em graus de dificuldade intermédia (cerca de 40%), observa-se uma dissociação entre iletradas, por um lado, e letradas e semiletradas, por outro: as primeiras dão mais erros globais e silábicos, e menos erros de primeira consoante, observando-se o contrário nas letradas e semiletradas. Este resultado é consistente com a hipótese de que a aprendizagem da leitura promove o desenvolvimento de uma estratégia de atenção aos segmentos da fala.

Finalmente, ficou também claramente demonstrado que o efeito de alfabetização na identificação de palavras em ruído se circunscreve a determinadas condições. Não se observou nenhum efeito na condição de maior dificuldade, cujo nível de *performance* era de cerca de 20%. Propomos para já que o efeito de alfabetização se observa apenas quando a *performance* não seja limitada pelos dados ("data-limited"), e as exigências da tarefa dêem azo à utilização de estratégias opcionais, como a atenção segmental. Esta proposta deverá ser testada tão directamente quanto possível em futuros trabalhos empíricos.

## 5. Discussão geral e conclusão

Este conjunto de experiências destinou-se a averiguar os efeitos da alfabetização na identificação de palavras faladas em condições difíceis. Escolhemos situações em que a análise das características da resposta fosse esclarecedora quanto ao processo perceptivo subjacente. Recorremos à técnica da audição dicótica, que permite observar fenômenos de análise e integração de propriedades dos estímulos em competição (cf., e.g. Studdert-Kennedy e Shankweiler, 1970), e à introdução de ruído, que, tornando a identificação mais difícil, aumenta as exigências de tarefa. Procurámos dissociar o efeito da idade, escolarização e alfabetização, através da observação de sujeitos que não dominassem a escrita, adultos iletrados e crianças pré-letradas, e de sujeitos de diferentes níveis de escolarização e idade.

Em geral, observámos que a identificação de palavras em condições difíceis é afectada pela idade e escolarização: sujeitos menos jovens e menos escolarizados têm um nível de desempenho inferior. Além disto, encontramos um efeito específico da alfabetização: para um mesmo nível de eficácia, mais ou menos intermédio, sujeitos que não disponham da possibilidade de representar a fala ortograficamente ou através de uma sequência de fones exibem um padrão de erros distinto do de sujeitos letrados, caracterizado por uma incidência relativamente menor de erros segmentais (Experiências III, IV e V) e relativamente maior de erros globais (Experiência V).

Os resultados de Brady, Shankweiler e Mann (1983) fornecem evidência convergente para a ideia de que o domínio do código escrito e o conhecimento ortográfico podem ser úteis ao ouvir linguagem falada. Numa situação experimental semelhante à que usámos na Experiência V, Brady *et*

*al. (ibidem)* verificaram que o efeito do ruído na identificação de palavras faladas era mais marcado nos chamados maus leitores do que nos bons leitores. Os sujeitos eram crianças de cerca de 8 anos, do terceiro ano de escolaridade; foram selecionadas 30 crianças de quocientes intelectuais semelhantes, mas com diferentes níveis de leitura, 15 bons leitores e 15 maus leitores. Estes indivíduos ouviram palavras de alta e baixa frequência em duas condições, primeiro mascaradas por ruído (razão sinal-ruído de 0 dB) e depois sem ruído. Apesar de os dois grupos terem o mesmo nível de *performance* nesta última condição, os maus leitores tinham piores resultados do que os bons leitores quando as palavras estavam mascaradas por ruído. Que este resultado não se deve a uma deficiência geral em tarefas de percepção auditiva ficou claramente demonstrado através de uma experiência adicional (*ibidem*, Experiência 3), em que não houve diferenças entre bons e maus leitores numa tarefa de identificação de sons ambientais (bater à porta, telefone a tocar, etc) mascarados por ruído. O facto do efeito pernicioso do ruído ser mais marcado nos maus leitores é pois específico, só ocorre com material verbal. É importante salientar que os piores resultados dos maus leitores se verificaram tanto para as palavras de baixa frequência, como para as muito frequentes. Não se pode pois atribuí-los a uma hipotética pouca familiaridade com as palavras.

A interpretação sugerida por Brady *et al. (ibidem)* é que os piores resultados na identificação de palavras em ruído se deixam explicar por dificuldades de recodificação fonética: "recall performance of poor readers for words presented auditorily suffers as a result of faulty phonetic coding" (p. 365). Esta sugestão filia-se naquela outra mais geral, defendida por vários investigadores dos Laboratórios Haskins, segundo a qual crianças com dificuldades de aprendizagem de leitura e maus leitores usariam menos a recodificação fonética (para uma explicitação recente deste ponto de vista, cf. Katz, 1986). O principal argumento em favor desta ideia é que os maus

leitores são menos afectados pela semelhança fonética dos estímulos em tarefas de memória a curto termo (não se observa o efeito de rima, *e.g.*, Shankweiler, Liberman, Mark, Fowler e Fischer, 1979).

Porém, outros investigadores têm obtido efeito de rima também em maus leitores (Hall, Ewing, Tinzmann e Wilson, 1981; Hall, Wilson, Humphreys, Tinzmann e Bowyer, 1983; Morais, Cluytens, Alegria e Content, 1986). O efeito de rima observa-se também em letrados e ex-letrados (Morais *et al.*, 1986), apesar de terem sido observadas diferenças substanciais quanto à memória imediata de dígitos nestes sujeitos (Morais, comunicação pessoal; ver também Experiência V, 4.3.4.; Read e Ruyter, 1985). Assim, as diferenças entre bons e maus leitores quanto à memória imediata para material verbal poderão não ser causadas pelo uso da recodificação fonética. A existência de diferenças mnésicas daquele tipo entre adultos muito e pouco escolarizados levanta a suspeita de que o próprio nível de desenvolvimento da linguagem escrita possa, em certa medida, afectar a *performance* em tarefas de memória imediata para material verbal (Morais, Alegria e Content, 1987-b).

Enfim, se fosse possível atribuir ao mesmo factor, a recodificação fonética, o pior desempenho na identificação de palavras em ruído e na evocação imediata de material verbal, dever-se-ia observar uma correlação entre as *performances* nestas duas tarefas. Os resultados da Experiência V não vão de encontro àquela hipótese: observou-se uma correlação positiva entre a amplitude de memória de dígitos e a percentagem de respostas correctas na identificação de palavras em ruído apenas num caso. Apesar de valer a pena explorar se, em condições de inteligibilidade particularmente difíceis, a maior ou menor facilidade em elaborar uma representação fonética é fonte importante da variabilidade inter-individual, parece-nos mais plausível uma interpretação alternativa à de Brady *et al.*

(1983). As diferenças entre indivíduos mais ou menos proficientes na leitura, e indivíduos iletrados, quanto à identificação de palavras em condições difíceis podem atribuir-se, pelo menos em parte, ao recurso a estratégias atencionais dirigidas para a estrutura fonética da palavra.

O papel das estratégias atencionais no reconhecimento da palavra falada tem vindo a ser assinalado noutros trabalhos. Investigando a ilusão de restauração fonémica, Nusbaum, Walley, Carrel e Ressler (1982) verificaram que sujeitos letrados conseguiam aprender a distinguir entre adição de ruído a um fonema, e a substituição do próprio fonema pelo ruído, se atendessem à estrutura fonémica de palavra. Mais recentemente, Samuel e Ressler (1986) obtiveram um efeito semelhante se fornecessem ao sujeito pistas sobre a identidade e localização do fonema crítico, e a palavra fosse previamente conhecida.

Os nossos resultados em situações onde é difícil identificar a palavra devido à competição dicótica (Experiência III e IV) ou ao ruído (Experiência V), e onde há menos constrangimentos lexicais por a incerteza ser sobre a palavra e não sobre um dos seus segmentos, propiciam uma sugestão mais forte: a atenção aos segmentos pode contribuir para a própria identificação da palavra, e não só para a identificação de um ou outro dos seus constituintes.

A interpretação segundo a qual o efeito de alfabetização, que observámos nas três experiências, é mediado pela estratégia de atenção segmental repousa em evidências correlacionais e indirectas. Convém submeter esta interpretação a um teste mais exigente. Iniciámos algum trabalho nesta direcção, e dispomos já de observações relevantes que vem a propósito referir.

**QUADRO 1:** Resumo dos resultados do estudo-piloto com estudantes universitários (cf. Texto)

	GRUPO CONTROLO (n=8)	GRUPO EXPERIMENTAL (n=10)
%R.C.	53.8 (30.0)	54.8 (21.3)
Inint.	26.2 (26.0)	28.7 (17.6)
Gl.	23.0 (13.9)	10.9 (8.4)
Sil.	3.9 (5.4)	3.5 (4.3)
C <sub>1</sub>	20.9 (12.3)	27.5 (10.2)
C <sub>1</sub> V <sub>2</sub>	4.3 (8.1)	4.0 (4.5)

Apresenta-se a percentagem de respostas correctas, seguida da proporção das várias categorias de erro relativamente ao total dos erros exibidos por cada sujeito: Inint., ou erros em que o sujeito diz que "não percebe nada", Gl, ou erros em pelo menos os três segmentos iniciais, Sil., ou erros na primeira e na segunda sílabas (C<sub>1</sub>V<sub>1</sub> + C<sub>2</sub>V<sub>2</sub>), C<sub>1</sub> ou erros restringidos à primeira consoante, C<sub>1</sub>V<sub>2</sub>, ou erros na primeira consoante acompanhados por erro na segunda vogal.

Estudantes universitários identificaram palavras em ruído binaural em duas condições. Na condição controlo, as instruções eram apenas de "se concentrarem bem e tentar identificar as palavras". Na condição experimental, tentava induzir-se uma estratégia de atenção segmental: explicava-se que todas as palavras tinham a estrutura CVCV, começavam por /b, d, g, p, t, k/ e eram palavras relativamente comuns; os sujeitos deveriam "esforçar-se por prestar atenção a cada som e em particular, a cada segmento e não só ao som global da palavra". Além destas instruções, os sujeitos na condição experimental recebiam "feedback". Nos dois grupos, apresentavam-se primeiro as palavras sem ruído, e após uma breve pausa fazia-se a identificação em ruído. O material foi o mesmo da Experiência V; apenas se eliminou a condição intermédia de ruído contralateral.

Os sujeitos eram estudantes do 2º ano de Psicologia\*; a sua idade rondava os 20 anos. Dez ouviram as palavras na condição experimental, e oito na condição controlo. Os resultados (cf. Quadro 1) mostram claramente que as instruções e o "feedback" não exerceram grande influência na exactidão, nem na proporção de respostas de tipo "ininteligível"; isto não deixa de ser um tanto surpreendente e mostra como as variáveis de estímulo são importantes.

Em todos os tipos de resposta há uma acentuada variabilidade interindividual. Notam-se porém algumas diferenças grupais: os sujeitos que fizeram a experiência na condição controlo têm em média mais erros globais e menos erros de primeira consoante do que os sujeitos da condição experimental.

---

\* Ano lectivo de 1987/88, estudantes das turmas de Psicologia da Linguagem e Cognição. Gostaríamos de agradecer a todos os estudantes que colaboraram na realização desta experiência e concordaram com a utilização dos dados para fins científicos.

Calculámos uma análise de variância aos valores transformados das proporções dos erros, com grupo como variável intersujeito e tipo de erro, globais *vs* segmentais, como variável intrasujeito. Consideramos, como nos estudos anteriores, a categoria global constituída pelos erros silábicos e pelos erros globais propriamente ditos; para a categoria segmental tomámos os erros de tipo  $C_1$  e  $C_1V_2$ , ou erros de primeira consoante.

Esta análise mostrou que a interacção entre grupo e tipo de erro é significativa,  $F_{(1,16)}=4.94$ ,  $p=.04$ . O efeito de grupo não é significativo,  $F_{(1,16)}=1.29$ ,  $p=.27$ , e o do tipo de erro não chega a atingir a significância,  $F_{(1,16)}=4.14$ ,  $p=.058$ . Para elucidar a interacção entre grupo e tipo de erro, comparámos separadamente a incidência de erros globais e segmentais. A diferença entre os dois grupos atingiu a significância apenas para os erros globais,  $t_{(16)}=2.33$ ,  $p<.02^*$ ; para os erros segmentais,  $t_{(16)}=1.21$ ,  $p=.12$ . Apesar destas comparações estarem à partida limitadas pelo reduzido número de sujeitos observados, vale a pena salientar alguns pontos.

Numa situação de audição de palavras em ruído, sujeitos universitários com um nível de desempenho equivalente exibiram um padrão de erros contrastante, conforme tivessem ou não sido instruídos a prestar atenção aos segmentos das palavras, e lhes tivessem sido dadas algumas indicações sobre a estrutura interna dos estímulos. Este resultado sugere que é possível manipular o modo como se "ouvem" palavras em condições difíceis, através de instruções relativamente simples. A instrução utilizada visou, de um modo um tanto grosseiro, induzir o sujeito a prestar atenção à estrutura interna das palavras, ou seja, a exercer uma atenção segmental.

---

\*Teste  $t$  para medidas independentes, calculados com os valores transformados; os valores da probabilidade referem-se a testes unicaudais.

Observámos uma dissociação quanto ao padrão de erros exibido pelos sujeitos submetidos a indução segmental, e pelos sujeitos que executaram a tarefa sem nenhuma instrução especial. Esta dissociação é semelhante à que se observa entre sujeitos letrados e iletrados no desempenho da mesma situação. Encontrámos pois evidência convergente para a hipótese de que a atenção segmental é em larga medida responsável pelas diferenças observadas entre letrados e iletrados em situações de identificação de palavras faladas em condições difíceis.

## CAPÍTULO VI

IDENTIFICAÇÃO DE ESTÍMULOS DIFERINDO QUANTO AO  
VOZEAMENTO

No presente capítulo exploraremos o papel da experiência com a linguagem oral e escrita na identificação de categorias fonémicas, nomeadamente quanto à distinção sonoro/surdo. Far-se-á um apanhado geral dos trabalhos procurando elucidar factores ontogenéticos na percepção da fala, através da comparação do desempenho de crianças de faixas etárias diferentes e de adultos. Apontar-se-á a relação encontrada por alguns investigadores entre percepção da fala e a aprendizagem da leitura.

O conjunto destes trabalhos, juntamente com observações informais de adultos iletrados, inspirou a realização de uma experiência sobre a nomeação de *continua* ao longo de uma dimensão foneticamente relevante, por sujeitos de diferentes graus de escolarização e idade. Uma tal experiência, em que se comparassem iletrados e letrados, teria a vantagem de fornecer dados relevantes para a questão teórica sobre que nos debruçamos, a de situar o nível a que se dá a influência da alfabetização na percepção da fala. Apresentaremos pois uma experiência de nomeação de *continua* de vozeamento por sujeitos iletrados, semiletrados e letrados, e ainda por um grupo adicional de estudantes universitários (Experiência VI).

Um objectivo adicional desta experiência é o estudo da distinção sonoro/surdo na língua portuguesa.

## 1. Introdução

### 1.1. Ontogénese e percepção de sons da fala

A ideia de que o modo como se percebe a fala depende de certas características do auditor tem recentemente vindo a ser submetida a teste experimental. Particular atenção tem merecido a investigação dos efeitos da ontogénese ligados à idade. Um número razoável de trabalhos tem demonstrado a existência de mudanças desenvolvimentais quanto à percepção de palavras, sílabas e frases, naturais ou sintéticas (*e.g.*, respectivamente: Elliot, Connors, Kille, Levin, Ball e Katz, 1979; Simon e Fourcin, 1978; Zlatin e Koenigsknecht, 1975; Elliot, Longinotti, Meyer, Raz e Zucker, 1981; Krause, 1982; Elliot, 1979; Orchik e Burgess, 1977; cf. também Experiência IV). Assim, por exemplo, verificou-se que adultos e crianças menos jovens conseguem identificar palavras a um nível sonoro mais fraco do que o necessário para que crianças mais jovens desempenhem a tarefa ao mesmo nível assintótico de quase 100%; concretamente, para ouvir sem erros palavras monossilábicas simples e conhecidas, como "milk" e "dog", crianças de 3 anos precisam que elas tenham uma intensidade quase 20 dB superior à exigida por crianças de 9 anos e adultos (Elliot, 1979; Elliot *et al.* 1979).

Estas mudanças desenvolvimentais quanto à fala parecem ser acompanhadas por um aumento de capacidades auditivas, como a sensibilidade a tons puros (*e.g.*, Yoneshige e Elliot, 1981) e a discriminação de frequência (*e.g.*, Eguchi, 1976). Não é porém verosímil que as mudanças ontogenéticas observadas quanto à fala, que veremos mais de perto já a seguir, se possam atribuir pelo menos exclusivamente a esta maior sensibilidade nos limiares absolutos e diferenciais para tons puros. Para tratar de estímulos acústicos com uma estrutura complexa, outras

operações estão envolvidas, como foi demonstrado por Spiegel e Watson (1981). Um factor que repetidamente tem vindo a ser discutido ou proposto para dar conta destas alterações ligadas à idade é a experiência, nomeadamente a experiência com a linguagem oral (Elliot *et al.*, 1981; Spiegel e Watson, 1981).

Certamente que um dos aspectos mais básicos da percepção da fala é a identificação de diferentes categorias fonéticas e a discriminação entre elas. O papel da experiência linguística precoce naquelas funções foi já abordado no contexto da percepção categórica (cf. Capítulo II). Concentremo-nos agora na faixa etária superior aos dois, três anos, e vejamos que alterações relacionadas com a idade e a experiência têm sido relatadas na percepção de sons da fala.

Elliot *et al.* (1981) averiguaram se a identificação e a discriminação de um *continuum* de /ba/ a /ga/ seriam de algum modo afectadas pela idade dos auditores. Estes eram sujeitos com audição normal, repartidos por três faixas etárias: 3 anos, 6 anos e adultos (cerca de 24 anos). Não se verificaram diferenças entre estes grupos quanto à capacidade de nomear sílabas sintéticas em comparação com sílabas naturais. Quanto à nomeação dos estímulos observaram-se algumas diferenças conforme a idade. Os autores analisaram o número de sujeitos que em cada faixa etária atingiram o critério de um mínimo de 81% de identificações correctas de /ba/, /da/ e /ga/. Apesar de algumas variações menores, conforme características dos estímulos - naturais ou sintéticos, com ou sem o ruído da explosão inicial - houve menos crianças do que adultos a atingir aquele critério. No caso dos estímulos naturais, a diferença entre as crianças de 6 e 10 anos era também significativa. Quanto à discriminação, os resultados indicaram que à medida que a idade aumentava, eram necessárias menos diferenças para se discriminarem os estímulos (*ibidem*).

Baseando-se na análise do número de ensaios necessários para atingir o critério, os autores frisam que não é correcto atribuir as diferenças etárias observadas a uma falta de atenção ou à maior variabilidade de resposta por parte dos sujeitos mais jovens. O facto de o desempenho das crianças de 10 anos se aproximar do dos adultos leva-os a concluir pelo carácter ontogenético dos efeitos observados, e a considerar o papel da experiência com a fala como um factor explicativo plausível.

### 1.2. Identificação de categorias fonémicas quanto ao vozeamento

Outros trabalhos têm apreciado o efeito da idade na identificação dos estímulos, independentemente da discriminação propriamente dita. O paradigma é semelhante ao da tarefa de identificação característica dos estudos de percepção categórica: usa-se um *continuum* de estímulos variando ao longo de uma dimensão acústica foneticamente relevante, que o sujeito deve nomear. Trata-se pois de tarefas de nomeação de *continua*. Tanto quanto sabemos, os trabalhos que procuraram verificar se existiam efeitos de idade na identificação de sons da fala usaram *continua* de vozeamento.

O primeiro foi levado a cabo por Zlatin e Koenigsknecht (1975), que observaram crianças de 2 e 6 anos, e adultos (idade média de 19.5 anos). Os estímulos eram palavras sintéticas iniciadas por oclusivas, variando o VOT de modo a serem constituídos os seguintes *continua* de vozeamento: "bees/peas", "bear/pear", "dime/time" e "goat/coat". Os sujeitos adultos identificavam as palavras escrevendo a consoante inicial que tinham percebido, e as crianças pressionavam um botão correspondente à imagem representando cada das palavras. A partir das funções de identificação, foram calculados os valores seguintes: a fronteira fonémica (tempo de vozeamento para o qual 50% das respostas são surdas e outras 50% são



sonoras), limite superior e inferior da fronteira fonémica (pontos do *continuum* correspondendo a 75% de respostas surdas e sonoras, respectivamente) e amplitude da fronteira fonémica (diferença em milisegundos entre os limites superior e inferior). Para estas três medidas verificaram-se diferenças significativas quanto à idade.

Quanto ao local da fronteira fonémica, as diferenças ocorreram apenas para o *continuum* velar. As crianças mais jovens tinham a fronteira mais tardia, em média, +77 milisegundos de "voicing lag",\* contra cerca de +64 milisegundos nas crianças de 6 anos e adultos; a distribuição dos sujeitos era tal que a maioria das crianças de 2 anos tinha valores de fronteira fonémica superiores aos dos outros sujeitos. A amplitude da fronteira fonémica foi significativamente maior nas crianças de 2 anos do que nos adultos, nos quatro *continua* utilizados. Por outras palavras, a zona de perceptos instáveis era maior para as crianças mais jovens do que para os adultos (cerca de duas vezes maior nos *continua* labial e dental, 14 milisegundos, e três vezes maior no velar, 21 milisegundos). A diferença entre as crianças de 6 anos e os adultos só ocorreu no *continuum* velar, e foi no mesmo sentido.

Estes resultados sugerem que a diferença de VOT necessária para se distinguirem consistentemente oclusivas surdas de sonoras diminui com a idade (*ibidem*, p. 548). Esta interpretação em termos de um refinamento na identificação de fonemas, em que cada categoria ganha limites mais claramente definidos, faz lembrar um fenómeno análogo na produção de fala. A variabilidade de certos aspectos da produção diminui com a idade. Mais especificamente, ao produzir vogais em vários contextos consonantais, a sua duração varia; ora, independentemente do contexto consonântico, o desvio-padrão da duração de várias realizações da mesma vogal diminui

\* Trata-se evidentemente de fonemas ingleses, em que as surdas se caracterizam por um atraso relativamente maior do vozeamento em relação à libertação da oclusão ("voicing lag").

com a idade, particularmente entre os 3 e os 6 anos (DiSimoni, 1974). Este e outros resultados na mesma linha indicam que há um desenvolvimento de precisão articulatória (Smith, 1978), talvez análogo ao desenvolvimento de precisão perceptiva.

Resultados coerentes com esta hipótese de um desenvolvimento da precisão perceptiva são os de um estudo translinguístico sobre nomeação de *continua* de vozeamento por crianças e jovens franceses e ingleses entre os 2 e os 14 anos de idade (Simon e Fourcin, 1978). Estes investigadores construíram três *continua*, entre palavras previsivelmente conhecidas dos sujeitos, "toto-dodo", "coat-goat" e "Paul-ball", através da manipulação de dois parâmetros acústicos, o VOT e a transição de primeiro formante. Quer para os sujeitos franceses, quer para os ingleses, ao longo da idade observou-se o seguinte padrão de resposta: as crianças mais jovens nomeiam os estímulos com VOT intermédios quase ao acaso; crianças mais velhas e os jovens identificam os estímulos de modo descontínuo; e crianças de idades intermédias apresentam uma transição gradual entre a proporção de respostas sonoras e surdas. As idades concretas para estes vários padrões de resposta diferem nos ingleses e franceses. Nos primeiros, temos o padrão aleatório aos 2 anos, e o categórico aos 4. Nos segundos, só aos 9 anos aparece o padrão categórico; entre 5 e 9 anos há a transição gradual, e até aos 5 anos persiste a inconsistência de resposta para VOTs intermédios.

Estas diferenças entre franceses e ingleses devem-se provavelmente às características acústicas dos estímulos em questão; com efeito, um dos índices da distinção entre sonora e surda era a transição de primeiro formante, que parece ser mais relevante na língua inglesa que na francesa. Devendo basear as suas distinções nas diferenças de VOT, que oscilavam apenas entre -30 e +30 milisegundos para /dodo-toto/, a tarefa

seria mais difícil para os franceses que para os ingleses (cf. Caramazza, Yeni-Komshian, Zurif e Carbone, 1973).

A variação do padrão de resposta observada por Simon e Fourcin (*ibidem*) poderia ocorrer não tanto porque os sujeitos identificam os estímulos mais ou menos consistentemente, mas devido à existência de grande variabilidade em cada grupo. Uma análise aos dados individuais mostrou que a distribuição dos sujeitos em cada faixa etária traduzia efectivamente um aumento do número de indivíduos que identificavam os estímulos de modo categórico ao longo da idade. Conforme o seu padrão de resposta, cada sujeito foi classificado num de três tipos de nomeação: categórica (se houvesse uma transição abrupta no número de respostas surdas ou sonoras), progressiva (se aquela transição fosse de tipo 0 1 1 2 3 ..., por exemplo) e dispersa ou "scattered" se não se observasse uma preponderância clara de respostas surdas ou sonoras conforme o valor de VOT dos estímulos (*ibidem*, p. 931). Havia uma nítida preponderância de nomeação dispersa nos grupos mais jovens e de nomeação categórica nos mais velhos. Talvez estes resultados traduzam a existência de um processo de aprendizagem ao nível das estratégias de nomeação (*ibidem*, p. 933). A influência da aprendizagem revela-se não só em termos de idade, mas também em termos de língua: como dizem Simon e Fourcin (*ibidem*, p. 934), o modo como o sujeito 'organiza os padrões acústicos em categorias linguísticas' varia de língua para língua.

Finalmente, um estudo de Krause (1982) encontrou diferenças ligadas à idade na nomeação de um tipo diferente de *continua* de vozeamento. É conhecido que um dos índices de contraste de vozeamento é, no inglês americano, a duração da vogal precedente; ou seja, à medida que a vogal antecedendo uma oclusiva dura mais, aumenta a probabilidade de que aquela oclusiva pós-vocálica seja percebida como sonora (e.g., Raphael,

1972). Krause (*ibidem*) utilizou três *continua* sintéticos variando a duração da vogal inter-consonântica em /bip-bib/, /pot-pod/ e /back-bag/, que foram identificados por sujeitos de três grupos etários: 3 anos, 6 anos e adultos (idade média de 20 anos).

Houve dois resultados principais, um sobre a fronteira fonémica e outro sobre o declive da função de nomeação. Quanto à fronteira fonémica observou-se um efeito significativo da idade: à medida que a idade aumenta, são necessárias durações vocálicas cada vez mais curtas para que a consoante seguinte seja percebida como sonora. Por exemplo, para o contraste de vozeamento labial, a fronteira fonémica é de 198 milisegundos (ms) nas crianças de três anos, 168 ms nas crianças de 6 anos e apenas 153 ms nos adultos. Nos *continua* alveolar e velar observa-se o mesmo padrão de resposta.

Quanto à transição entre respostas sonoras e surdas, ela foi mais abrupta nos adultos do que nas crianças. Este facto traduz-se quantitativamente no declive das funções de nomeação, que variou de modo significativo conforme a idade: as crianças de 3 e 6 anos não são tão consistentes quanto os adultos na utilização de uma determinada duração vocálica como indício de vozeamento da consoante seguinte. Estes resultados são semelhantes aos observados por Zlatin e Koenigsknecht (1975) quanto à amplitude de fronteira fonémica; ambos sugerem um refinamento perceptivo no sentido de uma maior capacidade discriminativa: são necessárias menos diferenças físicas ou acústicas para sinalizarem uma distinção fonémica. No caso específico do papel da duração vocálica na percepção do vozeamento, a influência da aprendizagem por experiência linguística aparece de modo particularmente claro. É que observações em bebés entre os 6 e 12 meses de idade revelaram ser-lhes muito difícil discriminar sílabas sonoras e surdas através da duração vocálica (cf. Eilers,

Wilson e Moore, 1977). A utilização desta variação acústica em termos linguísticos poderá pois dever-se, se não exclusivamente, pelo menos em grande parte, à experiência com a fala numa língua onde aquela variação esteja associada a distinções fonémicas (cf. Krause, 1982).

Há pois uma série de indicações consistentes de que a maneira como se identificam categorias fonémicas é, em certos aspectos, afectada pela experiência com a língua, nomeadamente com a fala. A questão que agora colocamos é de saber até que ponto um tipo específico de experiência linguística, a experiência com a escrita, não terá também alguma influência a esse nível.

## 2. Nomeação de *continua* de vozeamento velar e labial por letradas, semiletradas e iletradas (Experiência VI)

### 2.1. Introdução

A relação entre percepção da fala e linguagem escrita tem sido abordada no contexto da aprendizagem da leitura (cf. Brady, Shankweiler e Mann, 1983; Mann, 1984; Shankweiler, Liberman, Mark e Fowler, 1979). Foi sugerido que crianças com dificuldades naquela aprendizagem teriam problemas subtis a nível de identificação e discriminação de sons da fala: as categorias fonológicas seriam menos robustas, o que se traduziria numa identificação menos categórica e numa discriminação menos fina de estímulos variando numa dimensão foneticamente relevante (Godfrey, Syrdal-Lasky, Millay e Knox, 1981; Werker e Tees, 1987; para uma conclusão discordante, cf. Brandt e Rosen, 1980).

Estes trabalhos não permitem estabelecer claramente uma relação causa-efeito entre a percepção categórica de sons da fala e o sucesso na aprendizagem da leitura (cf. Stanovich, 1983; Tallal, 1980); mostram porém

que certos aspectos da percepção da fala, por um lado, e do domínio e utilização da linguagem escrita, por outro, não são independentes. É obviamente importante esclarecer esta relação. Para já, a sua mera existência confere plausibilidade ao argumento de que a experiência com a escrita, na medida em que envolve a consciencialização da estrutura da fala\*, poderá favorecer um refinamento na percepção de categorias fonológicas (categorias mais estáveis e claramente diferenciáveis). Uma hipótese semelhante foi levantada por Godfrey *et al.* (1981). Estes investigadores referem que a consciência da estrutura fonética da fala pode suscitar uma aprendizagem perceptiva que por sua vez desenvolveria a capacidade ("ability", *ibidem*, p. 422) de fazer distinções acústico-fonéticas subtis. Parece-nos todavia preferível não colocar para já o problema em termos de capacidade ou "ability", deixando em aberto a questão de, a existir essa tal aprendizagem, ela se centrar a nível de estratégias atencionais.

O estudo que se segue pretende averiguar se o conhecimento e experiência com a linguagem escrita afectam a maneira como são percebidos *continua* de vozeamento. Comparar-se-á a maneira como letradas, semiletradas e iletradas de idade semelhante identificam uma série de estímulos variando entre "pala-bala" ou entre "cola-gola". A mesma tarefa de identificação foi completada por um grupo de jovens estudantes universitários; estendendo as observações a um grupo etário diferente, dispomos de um ponto de comparação adicional que poderá ser útil na interpretação de possíveis diferenças grupais. Procurou-se minimizar factores como a variação dialectal: todos os sujeitos foram recrutados na zona Norte do país e as distinções críticas, /k-g/, /p-b/, não sofrem variações diatópicas (cf. Cunha e Cintra, 1987; Cintra, 1983). Supomos pois que

---

\* Subentendemos a escrita num sistema alfabético.

eventuais diferenças intergrupais se poderiam atribuir a aspectos mais específicos da experiência com a língua.

Eventuais efeitos da experiência com a língua poderão ser de dois tipos principais: ou provenientes do conhecimento e experiência com a escrita, ou relacionados com a experiência e o uso da linguagem oral. Na sequência dos resultados revistos ao longo desta Introdução, esperar-se-ia que, a haver efeitos da experiência com a escrita na nomeação de *continua*, eles seriam do seguinte teor: sujeitos letrados seriam mais consistentes não só na atribuição de uma etiqueta fonémica a um determinado estímulo ou grupo de estímulos, como também na atribuição de uma etiqueta diferente a outro estímulo ou grupo de estímulos diferentes. Isto traduzir-se-ia em funções de nomeação com zonas claramente definidas por uma preponderância de respostas ora surdas, ora sonoras, e uma descontinuidade abrupta entre estas duas zonas. A posição relativa das estudantes e semiletradas no contexto dos resultados dos outros grupos permitiria destrinçar um efeito da alfabetização independente da idade, de um eventual efeito de grupo independente da alfabetização.

Se se verificar uma dissociação entre, por um lado, estudantes, letradas e semiletradas e, por outro, iletradas, teremos indicação que a experiência com a escrita é relevante para a categorização de sons da fala. Se pelo contrário a dissociação obedecer a um padrão diferente, em que iletradas e letradas tenham resultados semelhantes, teremos encontrado evidência empírica em favor de um efeito de grupo possivelmente ligado à experiência com a linguagem oral.

Escolheram-se palavras em fala natural para que os estímulos se aproximassem tanto quanto possível da linguagem normal e obviar assim a possíveis problemas devido à utilização de material pouco familiar aos sujeitos. O português é uma língua que se presta particularmente bem à

constituição de *continua* de vozeamento com fala natural; na medida em que as oclusivas sonoras se distinguem das surdas homorgânicas pela existência de vozeamento anterior à libertação da oclusão, é possível através do corte ou adição de pré-vozeamento construir surdas ou sonoras, respectivamente, a partir de uma elocução natural.

Tanto quanto sabemos, trata-se do primeiro estudo com fala portuguesa a abordar a distinção perceptiva entre oclusivas surdas e sonoras através da manipulação do pré-vozeamento. Um objectivo adicional da experiência que se segue é uma primeira contribuição para o estudo da percepção de contrastes de vozeamento na língua portuguesa.

## 2. 2. Método

### 2.2.1. Sujeitos

Participaram nesta experiência 65 sujeitos do sexo feminino, distribuídos pelos seguintes grupos: jovens universitárias (n=14), adultas letradas, semiletradas e iletradas (n=18x3). As universitárias frequentavam o 1º ano do Curso de Psicologia\* e a sua idade variava entre 19 e 28 anos (média de 21, desvio-padrão de 2.8). As adultas letradas tinham completado um curso superior e trabalhavam como docentes do ensino secundário\*\*. Tinham em média uma idade de 49.5 anos (desvio-padrão de 8.6), variando entre 36 e 66 anos. As semiletradas tinham um grau de escolarização de, no máximo, quatro anos. Embora sabendo ler e escrever com variável grau de fluência, não usavam quotidianamente estas competências. Trabalhavam

---

\* Ano lectivo de 1985/86, Faculdade de Psicologia e de Ciências da Educação, Universidade do Porto.

\*\* Zona do Porto e Grande Porto, Escola Secundária Clara de Resende, Aguas Santas e Gondomar.

em actividades agrícolas ou domésticas\* e a sua idade variava entre 39 e 62 anos (média de 49.3, desvio-padrão de 8.2 anos). As iletradas não tinham frequentado a escola enquanto crianças ou adultas por razões sócio-económicas. Ocupavam-se em trabalhos domésticos ou agrícolas\*\*. A sua idade variava entre 40 e 64 anos (média de 51.2, desvio-padrão de 6.9 anos).

Nenhuma das participantes tinha uma história conhecida de deficiência neurológica, intelectual ou auditiva. Quando submetidas a um teste sumário de acuidade auditiva, composto por 12 palavras CVCV (6 em cada ouvido) a dois níveis decrescentes de intensidade (70 e 50 dB SPL), não se notaram perdas auditivas uni ou bilaterais.

### 2.2.2. Estímulos

Com base em emissões naturais de "pala", "bala", "cola" e "gola" foram construídos quatro *continua* de vozeamento, duas versões do *continuum* "pala-bala" e duas versões do *continuum* "cola-gola"\*\*\*. Cada *continuum* variava entre zero e 100 milisegundos (ms) de "voicing lead", ou pré-vozeamento, a intervalos de 20 ms; havia pois seis estímulos por *continuum*. Descrevem-se seguidamente as várias etapas de produção dos estímulos.

Gravaram-se num Ampex vários exemplares de pares mínimos pronunciados por uma falante portuguesa numa câmara acústica insonorizada. Os "melhores" destes exemplares, *i.e.*, os mais claros e

---

\* A maioria era proveniente de meio agrícola (Sanguedo e Lourosa, Vila da Feira ou Macieira, Vale de Cambra). Três trabalhavam no Porto em serviços domésticos.

\*\* A maioria era proveniente de meio agrícola (Sanguedo, Vila da Feira ou Macieira, Vale de Cambra).

\*\*\* Este trabalho foi realizado nos Laboratórios Haskins e contou com a preciosa ajuda de Bruno Repp.

auditivamente semelhantes quanto à entoação, intensidade e duração foram então digitalizados usando o sistema PCM ("Pulse Code Modulation") do computador VAX 750 dos Laboratórios Haskins, a uma frequência de amostragem ("sampling rate") de 20 KHz. Os sonogramas destes estímulos foram analisados e manipulados por meio do programa Wendy ("Wave Editing and Display") com o objectivo de criar a partir de um estímulo sonoro o seu correspondente surdo, e vice-versa. Por exemplo, o estímulo sonoro "bala" tinha um pré-vozeamento de 130.4 ms; extraindo porções maiores ou menores deste pré-vozeamento era possível modificar a percepção correspondente de [bal $\alpha$ ] para [pal $\alpha$ ]. O estímulo surdo "pala" não tinha pré-vozeamento; juntando-lhe porém porções variáveis do pré-vozeamento de [bal $\alpha$ ], era possível alterar o percepto de [pal $\alpha$ ] para [bal $\alpha$ ].

As manipulações acústicas (corte ou adição de pré-vozeamento) foram perceptivamente relevantes (indução de percepto surdo ou sonoro) para dois dos cinco pares mínimos iniciais, "cola/gola" e "pala/bala". Nos restantes, "galo/caló", "gana/cana" e "doca/toca", aquelas manipulações afectavam ora a clareza do percepto (na adição de vozeamento às surdas) ou eram inoperantes: mesmo sem pré-vozeamento, os exemplares truncados de "galo", "gana" e "doca" evocavam a percepção de sonora. Certamente isto ocorreu porque em português, como noutras línguas, haverá múltiplos índices da característica perceptiva de sonoro ou surdo ou, eventualmente, por efeitos de coarticulação. Foram pois construídos quatro *continua*, dois por corte de pré-vozeamento do estímulo original sonoro [bal $\alpha$ ] e [gol $\alpha$ ], em passos de 20 ms, (e.g., "bala" com 100 ms de pré-vozeamento = E<sub>10</sub>, "bala" com 80 ms = E<sub>8</sub>, ..., "bala" com 0 ms = E<sub>0</sub>); e outros dois por adição do pré-vozeamento da correspondente sonora ao exemplar surdo (e.g., "pala" original = E<sub>0</sub>; "pala" com 20 ms do pré-vozeamento de "bala" = E<sub>2</sub>; ...; "pala" com 100 ms do pré-vozeamento de "bala" = E<sub>10</sub>).

A selecção quer dos melhores exemplares dos estímulos, quer dos intervalos de pré-vozeamento, baseou-se essencialmente nas avaliações sobre a qualidade perceptiva dos estímulos feitas por uma falante portuguesa; em caso de ambiguidade e para teste recorreu-se à avaliação de um falante com grande experiência em fonética desconhecedor do português.\* Nos dois casos, os pontos extremos dos *continua* suscitaram percepções não-ambíguas e contrastantes; após tentativas com outros intervalos temporais, o de 20 ms parecia ser o mais apropriado para estabelecer um *continuum* entre surda e sonora que não fosse demasiadamente longo. Queríamos evitar uma situação de teste morosa e repetitiva, por receio que isso não fosse bem suportado por alguns sujeitos.

Os 24 estímulos finais (6 estímulos por *continuum*) foram definitivamente reconvertidos de digital para analógico, e gravados em fita magnética numa sequência aleatória com 5 replicações por estímulo e 3 segundos de intervalo inter-estímulo. A gravação em fita foi posteriormente passada a cassette num Revox A77, e apresentada aos sujeitos num gravador Phillips D 6920 MK2 e, nas sessões individuais, através de auscultadores Philips EM 6421.

### 2.2.3. Procedimento

Os universitários fizeram a experiência numa sessão grupal, com os estímulos apresentados por meio de dois altifalantes situados em frente aos sujeitos, cada um num canto da sala. Dispunham de uma folha de resposta em que escreviam a consoante inicial da palavra ouvida em cada

---

\* A falante portuguesa foi a autora, e o falante não-português foi B. Repp. Claro que idealmente ter-se-ia recorrido a pelo menos dois falantes portugueses para avaliação da qualidade dos estímulos. Infelizmente isso não foi possível durante a curta estada nos Laboratórios Haskins.

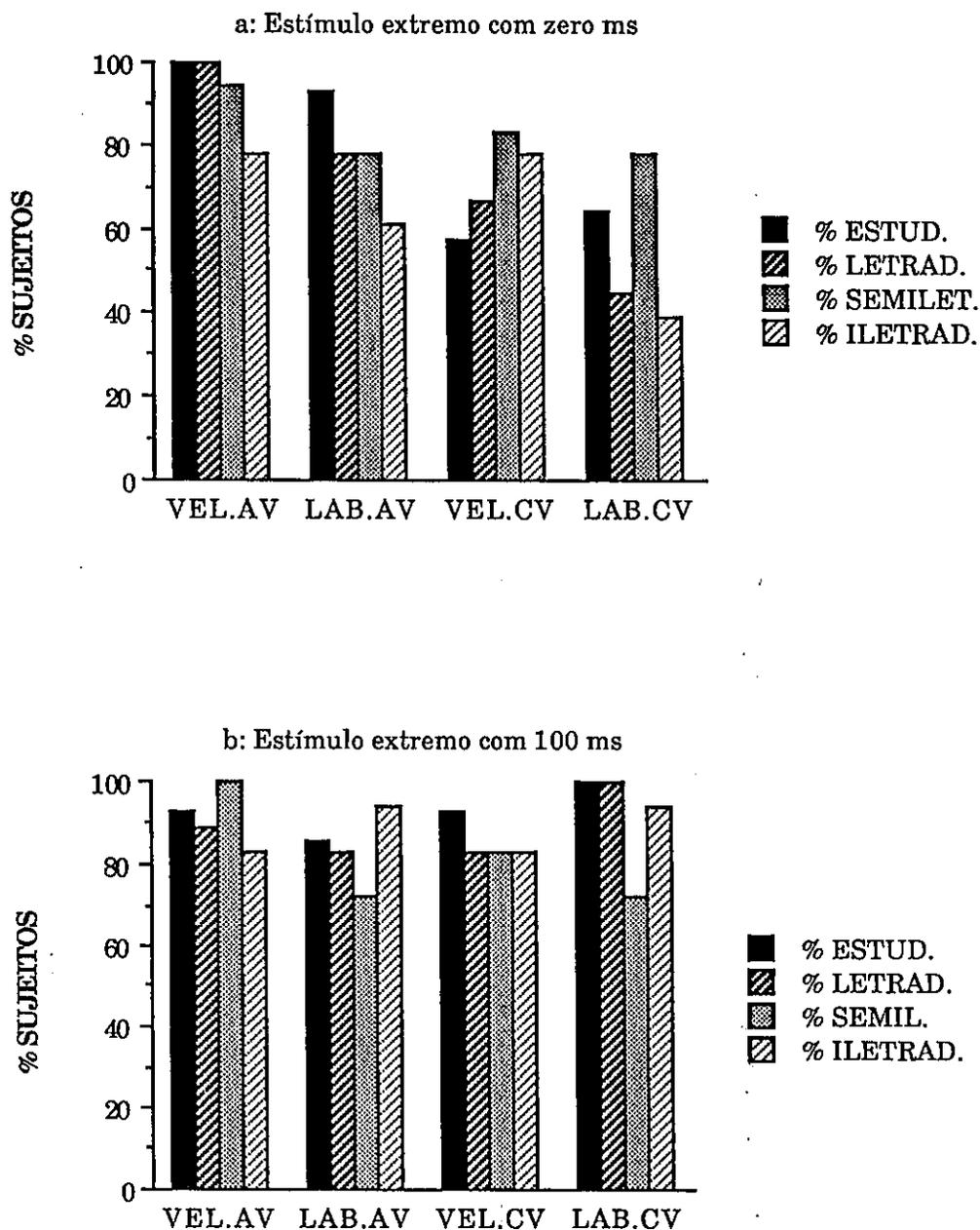
ensaio. Os restantes sujeitos foram testados em sessão individual. Os estímulos eram apresentados em monaural por meio de auscultadores e a resposta era dada oralmente, sendo anotada pela experimentadora.

As instruções eram idênticas para todos os sujeitos. Explicava-se-lhes que iriam ouvir vários exemplares de palavras portuguesas comuns; a tarefa consistia apenas em escrever/repetir o que iam ouvindo. Sempre que ouvissem uma palavra, deveriam responder; se tivessem dúvidas, era importante que mesmo assim escrevessem/repetissem exactamente como lhes tinha parecido. Antes dos ensaios experimentais, faziam-se oito ensaios de treino em que eram apresentadas as versões extremas ( $E_0$  e  $E_{10}$ ) dos quatro *continua*. Dava-se *feedback* e, antes de prosseguir com a experiência, esclarecia-se que a seguir iriam "ouvir várias vezes aquelas palavras, 'bala, pala, gola, cola'". Nenhum dos sujeitos teve dificuldade em identificar correctamente os estímulos extremos nos ensaios de treino.

Os estímulos eram apresentados a uma intensidade confortável, cerca de 70 dB SPL se através de auscultadores, e com o volume regulado de modo a que todos os sujeitos da sessão grupal achassem "ouvir bem".

Nas sessões individuais, antes de fazer a experiência propriamente dita, os sujeitos eram submetidos ao teste sumário de acuidade auditiva. Todos os estudantes fizeram o mesmo teste de acuidade auditiva em sessão individual, numa situação subsequente. Se fossem notadas perdas uni ou bilaterais, o sujeito era substituído.

Quase todos os sujeitos adultos realizaram a seguir outras tarefas de reconhecimento auditivo de palavras (experiência de inteligibilidade de palavras ou de fusões fonológicas), após um intervalo de alguns minutos.



**FIGURA 1:** Percentagem de sujeitos que responderam consistentemente aos estímulos extremos com zero ( $E_0$ ) e 100 ( $E_{10}$ ) milissegundos de pré-vozeamento, identificando nas 5 apresentações o primeiro como surdo e o segundo como sonoro. AV e CV são abreviaturas para adição e corte de pré-vozeamento. As percentagens foram calculadas relativamente ao total de sujeitos por grupo, 14 nos estudantes, e 18 nos restantes grupos. Apresentam-se os resultados por *continuum* e grupo.

### 2.3. Resultados

Importa primeiro verificar como os sujeitos responderam aos estímulos extremos dos *continua*. Com efeito, a comparação das curvas de identificação obtidas pelos vários indivíduos e grupos assenta no pressuposto de que os estímulos extremos são percebidos como exemplares de categorias "opostas"; a utilização de dados em que tal não aconteça compromete severamente uma interpretação clara das semelhanças e diferenças encontradas. Na Figura 1 pode ver-se que nem todos os participantes deram respostas invariáveis, principalmente aos estímulos dos *continua* labiais com zero ms de pré-vozeamento.

A inspecção dos resultados individuais mostrou que para um mesmo sujeito a consistência de respostas variava conforme o *continuum*; por exemplo, a letra n<sup>o</sup> 9 ouviu sempre como surdo o E<sub>0</sub> do *continuum* velar por adição de vozeamento, mas ouviu 2 vezes como sonoro o estímulo correspondente (também com zero ms de pré-vozeamento) do *continuum* velar por corte de pré-vozeamento. Houve também uma minoria de sujeitos que num caso ou outro não identificaram "correctamente" os estímulos extremos; por exemplo, três iletradas ouviram sempre "bala" nas várias apresentações dos E<sub>0</sub> do *continuum* labial por corte de pré-vozeamento, e duas semiletradas identificaram os extremos sonoros de ambos os *continua* labiais como surdos, em três das cinco apresentações. Resultados deste tipo são nitidamente discordantes mesmo no seu contexto grupal, e podem distorcer mais ou menos gravemente os valores médios. Eles não podem ser claramente atribuídos a deficiências auditivas dos sujeitos, pois um mesmo sujeito pode exibir um padrão atípico num *continuum*, e um padrão tipicamente médio noutra.

**QUADRO 1:** Número de sujeitos cuja identificação dos estímulos extremos,  $E_0$  e  $E_{10}$ , apresenta uma variação de no máximo 3 desvios-padrão relativamente à média grupal (cf. Texto), por grupo e tipo de *continuum*. De modo a que seja possível fazer análises sobre o efeito do modo de construção nos *continua* velares e labiais, verificou-se o número de sujeitos que obedecessem àquele critério nos dois casos.

	EST.	LET.	SLET.	ILET.
VEL. AV	14	17	17	17
VEL. CV	13	17	17	17
LAB. AV	14	16	16	15
LAB. CV	14	17	15	14
VELAR	13	16	17	16
LABIAL	14	16	15	13

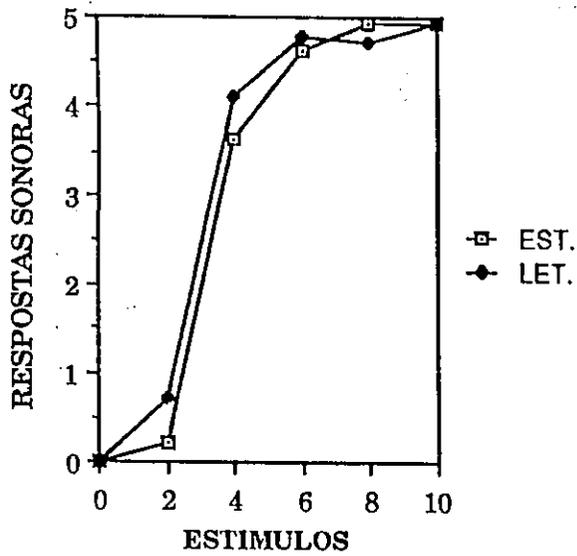
Usaram-se as seguintes abreviaturas: para o grupo, EST., estudantes, LET., letradas, SLET., semiletradas e ILET., iletradas; para os *continua*, VEL. refere-se a velar, LAB. a labial, AV a adição de pré-vozeamento e CV a corte de pré-vozeamento.

Independentemente da sua origem ser auditiva, atencional ou meramente idiossincrática, é conveniente eliminar os dados de sujeitos que tenham identificado os extremos dos *continua* de modo incorrecto. Como a noção de correcto e incorrecto neste contexto não é absoluta, usou-se o seguinte critério estatístico: se um sujeito responde aos estímulos extremos,  $E_0$  e/ou  $E_{10}$ , de um determinado *continuum* desviando-se da média do seu grupo três ou mais desvios-padrões, as suas respostas para esse *continuum* não são consideradas. Este critério apresenta a vantagem de eliminar as respostas atípicas respeitando a variabilidade característica de cada grupo; na prática, ele equivale a eliminar os sujeitos que identificam  $E_0$  e  $E_{10}$  da mesma maneira (*i.e.*, ambos surdos ou ambos sonoros), sempre ou a maioria das vezes. Isto foi relativamente raro para os *continua* velares, mas um tanto mais comum para os *continua* labiais, principalmente no de corte de pré-vozeamento (cf. Quadro 1).

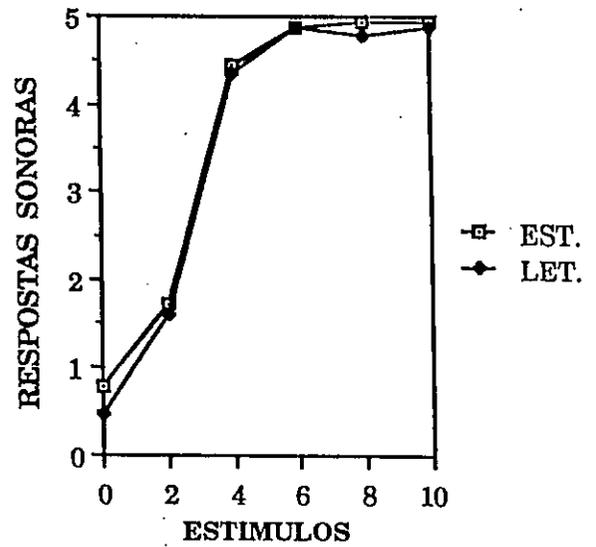
### 2.3.1. Universitárias e adultas letradas

Porque os modos de apresentação foram diferentes para os estudantes e restantes sujeitos, consideremos primeiro as respostas dos estudantes. Como se pode ver na Figura 2, nos dois *continua* por adição de pré-vozeamento os estímulos extremos foram praticamente sempre identificados como era esperado: sem pré-vozeamento como surdos ("cola" e "pa\_la"), e com 100 ms de pré-vozeamento como sonoros. Isto aconteceu com todos os sujeitos, sendo as excepções particularmente raras. No *continuum* "cola-gola", é particularmente evidente a descontinuidade da nomeação, que ocorre entre 20 e 40 ms ( $E_2$  e  $E_4$ , respectivamente) de pré-vozeamento: o primeiro evoca resposta sonora em média apenas 0.2 vezes em 5, enquanto o segundo é praticamente sempre ouvido como sonoro. No *continuum* "pala-

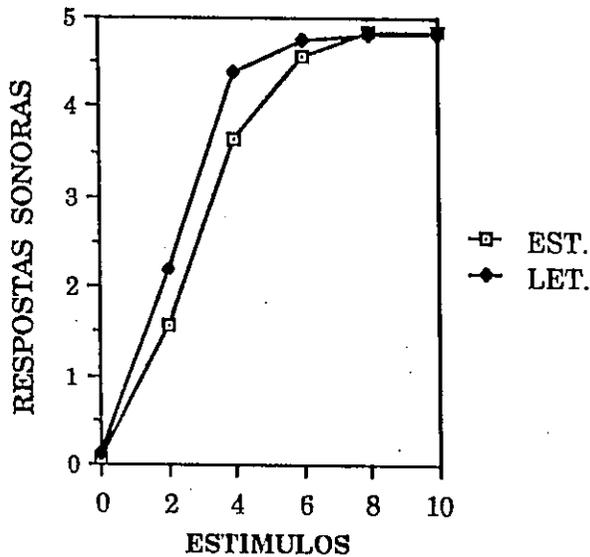
## VELAR (ADIÇÃO VOZ.)



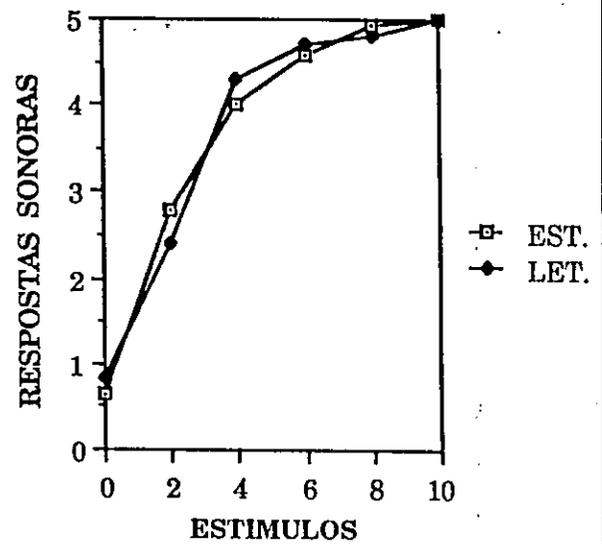
## VELAR (CORTE VOZ.)



## LABIAL (ADIÇÃO VOZ.)



## LABIAL (CORTE VOZ.)



**FIGURA 2:** Número médio de respostas sonoras a estímulos de zero a 100 milissegundos (ms) de pré-vozeamento, em passos de 20 ms, para estudantes universitários (UNIV.). Em ordenada, o número de respostas (máximo possível de 5) e em abcissa o estímulo (o subscripto indica a duração de pré-vozeamento em centésimos do segundo). Para fins de comparação, apresentam-se também os resultados das letradas adultas (LET.).

-bala", a fronteira fonémica situa-se também entre os estímulos 2 e 4, embora  $E_2$  suscitasse mais respostas sonoras (em média 1.6).

Os *continua* construídos por corte de pré-vozeamento (cf. Figura 2) oferecem resultados semelhantes quanto à configuração geral: respostas invariavelmente sonoras aos estímulos 6, 8 e 10, respostas maioritariamente surdas ao estímulo 0, e maior variabilidade nas respostas aos estímulos 2 e 4. Nota-se todavia que os estímulos supostamente surdos, os  $E_0$ , não foram ouvidos como tal de modo tão consistente como na situação anterior: há mais respostas sonoras, em média, e menos sujeitos que mantêm a resposta surda em todas as ocorrências de  $E_0$  (cf. Figura 2 e Figura 1, respectivamente).

Os resultados obtidos pelas adultas letradas estão também representados na Figura 2. É evidente a sua semelhança com os das estudantes, tanto mais notável quanto a técnica de apresentação foi distinta nos dois casos.

Para a análise estatística dos dados foi seguido um procedimento idêntico ao utilizado por Godfrey *et al.* (1981) e Werker e Tees (1987): fizeram-se análises de variância com estímulo como factor intrasujeito e grupo como factor intersujeito, além de outras variáveis de interesse.

O facto de termos apenas 6 estímulos não justifica que se procure quantificar quer a localização, quer a amplitude da fronteira fonémica, nem o declive da função de nomeação. Optámos por considerar separadamente as respostas aos *continua* velares e labiais, tomando o modo de construção (adição ou corte de pré-vozeamento) como variável intrasujeito. Deste modo podemos observar se haverá alguma interação entre o modo de construção e as outras variáveis (nomeadamente o grupo e o estímulo), sem nos expormos

a complicadas interacções triplas que poderiam ocorrer caso considerássemos o lugar de articulação (velar *vs* labial) como variável intrasujeito.

Com vista a analisar os resultados obtidos pelas universitárias e letradas calcularam-se duas análises de variância, com grupo como variável intersujeito, e estímulo e modo de construção como variáveis intrasujeito, uma para os *continua* velares e outra para os labiais. Como os dados sugerem, quer para os *continua* velares quer para os labiais o efeito do estímulo foi altamente significativo,  $F_{(5,27)}=456.94$   $p<.0001$  e  $F_{(5,28)}=277.76$ ,  $p<.0001$  respectivamente; não há efeito de grupo ( $F<1$  nos dois casos), nem interacção entre grupo e estímulo (para os *continua* velares,  $F_{(5,27)}=1.14$ ,  $p=.36$  e para os labiais,  $F<1$ ). Como suspeitávamos, o facto de se ter cortado ou adicionado porções de pré-vozeamento afectou o padrão de respostas: quer o efeito do modo de construção quer a sua interacção com o estímulo foram significativos, tanto nos *continua* velares, como nos labiais\*. Para os *continua* velares, não se verificou interacção entre modo de construção e grupo,  $F_{(1,27)}=1.8$ ,  $p=.19$ . Todavia, na identificação de "pala/bala" as letradas parecem ser menos afectadas pelo modo de construção do que as universitárias: a interacção é significativa,  $F_{(1,28)}=4.54$ ,  $p=.04$ . Não houve interacção tripla, nem entre grupo e estímulo ( $F<1$  nos dois casos).

Em suma, a despeito da diferença de idade e modo de apresentação, as letradas adultas e as estudantes universitárias identificaram de modo semelhante os estímulos extremos e os estímulos ambíguos de quatro *continua* de vozeamento. A única variação intergruppal significativa ocorreu quanto ao efeito do modo de construção dos estímulos labiais: as univer-

\* Os valores são os seguintes: nos *continua* velares, para o efeito principal do modo de construção  $F_{(1,27)}=26.95$   $p<.0001$ , e para a interacção com o estímulo  $F_{(5,27)}=4.22$ ,  $p=.007$ . Os valores correspondentes para os *continua* labiais são  $F_{(1,28)}=20.42$ ,  $p=.0001$  e  $F_{(5,28)}=20.42$ ,  $p=.0001$  e  $F_{(5,28)}=3.11$ ,  $p=.0262$ . Para pormenores sobre estas análises, cf. Anexos C.1/1 e C.1/2.

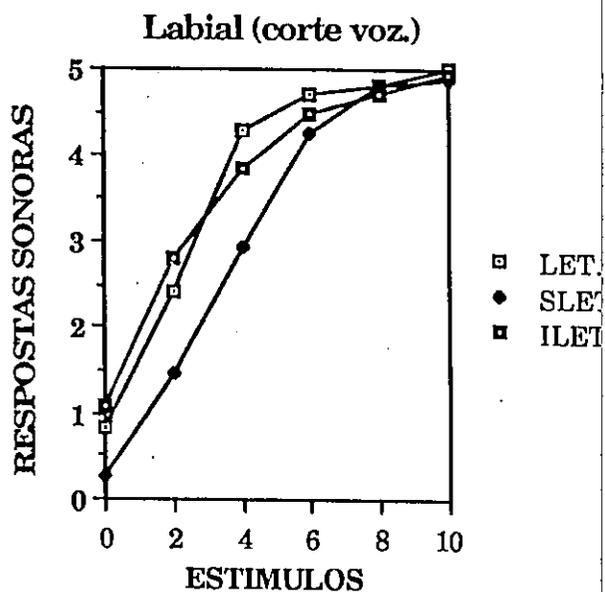
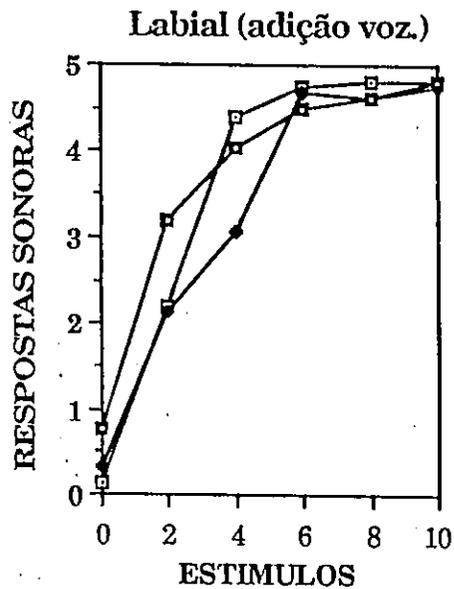
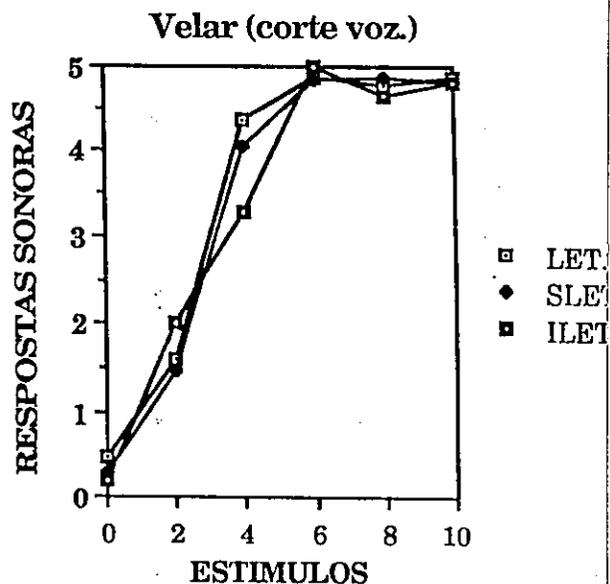
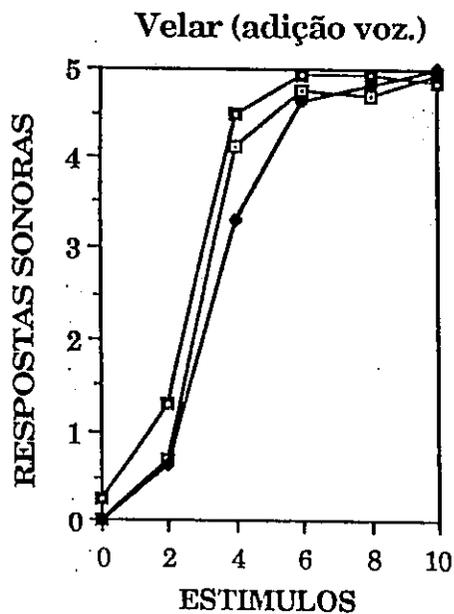


FIGURA 3: Número médio de respostas sonoras a estímulos de zero a 100 milissegundos (ms) de pré-vozeamento, em passos de 20 ms, no grupo das iletradas (ILET.) e semiletradas (SLET.). Em ordenada o número de respostas (máximo possível de 5) e em abcissa o estímulo (o subscripto a duração do pré-vozeamento em centésimos do segundo). Para fins de comparação, apresentam-se na mesma figura os resultados das letradas adultas (LET.).

sitárias parecem responder com mais sonoras se o *continua* tiver sido construído por corte de vozeamento, enquanto as letradas não são tão afectadas pelo facto de os estímulos terem o pré-vozeamento cortado ou adicionado.

### 2.3.2. Adultas iletradas, semiletradas e letradas

A configuração geral das respostas das semiletradas e iletradas obedece ao padrão observado nas jovens universitárias e letradas adultas: os estímulos com 60, 80 e 100 ms de pré-vozeamento são ouvidos quase sempre como sonoros; os estímulos sem pré-vozeamento são em geral percebidos como surdos, embora a consistência deste tipo de resposta seja variável; finalmente, os estímulos com 20 e 40 ms de pré-vozeamento provocam respostas muito menos consistentes, embora mais uma vez o grau de consistência da resposta varie segundo o *continuum* e, nalguns casos, segundo o grupo (cf. Figura 3).

Por questão de simplicidade, vejamos primeiro os resultados relativos aos *continua* labiais. Uma análise de variância com grupo como variável intersujeito, e estímulo e modo de construção como variáveis intrasujeito revelou uma interacção significativa entre grupo e estímulo,  $F_{(10,74)}=2.04$ ,  $p=.04$ . Nenhuma outra interacção atingiu a significância e só os efeitos principais do grupo e estímulo foram significativos\*. A interacção entre grupo e estímulo dever-se-á possivelmente a duas razões: por um lado as semiletradas respondem mais frequentemente com surda, principalmente para o estímulo 4; por outro, a transição entre preponderância de respostas surdas para preponderância de respostas sonoras parece ser mais abrupta nas letradas do que para os restantes sujeitos. Mais à frente trataremos este aspecto com mais pormenor.

\* Para pormenores, cf. Anexo C.1/3.

As funções de identificação destes *continua* caracterizam-se por uma frequência relativamente elevada de respostas sonoras aos estímulos com pouco pré-vozeamento: os E<sub>0</sub> e E<sub>2</sub> foram mais frequentemente ouvidos como "bala" do que os seus congéneres velares foram ouvidos como "gola". Verifica-se que, em geral, as semiletradas respondem mais vezes "pala" do que os outros dois grupos.

Voltemo-nos agora para os resultados obtidos com os *continua* velares. Uma análise de variância análoga à anterior revelou um quadro mais complicado, com o modo de construção a interagir significativamente com o estímulo e com o grupo, sendo também a interacção tripla altamente significativa,  $F_{(10,84)}=3.16$ ,  $p=.001^*$ . Por estas razões, fizeram-se análises separadas por *continuum*, com apenas duas variáveis, grupo e estímulo. Quanto ao *continuum* por adição de pré-vozeamento, os dois efeitos principais foram significativos, bem como a interacção: para o estímulo,  $F_{(5,42)}=542.77$ ,  $p<.0001$ , para o grupo,  $F_{(2,42)}=4.64$ ,  $p=.0014$  e para a interacção,  $F_{(10,74)}=2.99$ ,  $p=.0014^{**}$ . O efeito de grupo dever-se-á ao facto de, em geral, haver mais respostas sonoras nas iletradas e mais respostas surdas nas semiletradas. Quanto à interacção, ela poder-se-á atribuir essencialmente às respostas para os estímulos 2 e 4: a nomeação parece, em média, mudar mais abruptamente no grupo das letradas (68% de diferença entre a percentagem de respostas sonoras a E<sub>2</sub> e a E<sub>4</sub>) do que no das semiletradas (53% de diferença); o valor correspondente no grupo das iletradas é intermédio, 63% (cf. Figura 3).

Considerando agora o *continuum* velar feito por corte de pré-vozeamento, observa-se que as iletradas identificam os estímulos 2 e 4 relativamente próximo do nível do acaso, dando respectivamente 43 e 68% de

\* Para não sobrecarregar o texto, apresentam-se os valores apenas no Anexo C.1/4.

\*\* Para pormenores, cf. Anexo C. 1/5.

respostas sonoras; os valores correspondentes são, para as letradas, 36 e 87% e para as semiletradas, 30 e 84%. Estes resultados levantam a suspeita de que o hipotético\* declive da função da identificação é menos pronunciado nas iletradas do que nos outros dois grupos. De facto, uma análise de variância análoga à anterior revela uma interacção significativa entre grupo e estímulo,  $F_{(10, 74)}=2.99, p=.001^{**}$ . Contrariamente ao *continuum* anterior, não se observa aqui nenhuma "preferência" por parte de algum grupo em responder "gola" ou "cola": o efeito de grupo não é significativo ( $F<1$ ).

#### 2.4. Discussão e conclusão

Começamos por rever os resultados comuns aos vários grupos de sujeitos; com efeito, o interesse pelas possíveis fontes de variação inter-grupal não deveria levar-nos a negligenciar certas regularidades inter-individuais importantes num estudo deste tipo, em que se pretende também ganhar algum conhecimento sobre a distinção sonora-surda na língua portuguesa.

Quer jovens universitários quer adultos mais idosos de vários níveis de escolarização identificaram consistentemente estímulos com mais de 60 milisegundos de pré-vozeamento como sonoros; isso aconteceu tanto para a consoante velar /g/ como para a labial /b/, quer se tivesse cortado o pré-vozeamento de sonora original, quer se tivesse adicionado artificialmente pré-vozeamento à surda homorgânica. A resposta a estímulos com zero ou 20 milisegundos de pré-vozeamento parece ser mais dependente do *continuum*, nomeadamente conforme o lugar de articulação.

\* Dizemos "hipotético" porque não dispomos de medidas contínuas, ou sequer intermediárias, entre os estímulos 2 e 4.

\* \*Cf. pormenores no Anexo C.1/6.

De facto são mais frequentes as respostas sonoras perante 20 milissegundos de pré-vozeamento nos *continua* labiais do que nos velares.

Isto poderia acontecer devido a características específicas dos nossos estímulos, quer em termos acústicos, quer até lexicais. A palavra /pala/ poderá ser menos familiar do que /cola/, e isso talvez afecte o modo como os sujeitos categorizam os estímulos, quer a nível perceptivo propriamente dito (cf. Ganong, 1980; Connine e Clifton, 1987) quer a nível pós-perceptivo (cf. Fox, 1984). Todavia, não podemos deixar de salientar que a indicação de que a fronteira fonémica varia conforme o *continuum* seja velar ou labial é perfeitamente coerente com o que acontece noutras línguas. Por exemplo, é conhecido que em inglês a fronteira entre oclusivas labiais surdas e sonoras está próxima de 0 milissegundos VOT, enquanto a fronteira para as velares correspondentes é superior a +30 milissegundos de VOT (e.g. Abramson e Lisker, 1968; subentende-se que aqueles valores dizem respeito a situações em que não há variação de velocidade da fala ou do contexto).

O facto de o modo de construção afectar o padrão de respostas não constitui surpresa: a técnica de cortar períodos de pré-vozeamento obviamente não elimina outros índices do contraste sonora-surda, que certamente existirão. Curiosamente este efeito aparece mais nos sujeitos jovens, onde ocorre quer para os contrastes velares quer para os labiais; nos sujeitos adultos, ele só foi observado no *continuum* velar. Talvez isso tenha sucedido porque nos *continua* labiais as respostas dos adultos foram mais variáveis (veja-se o número de respostas sonoras, principalmente aos estímulos 0 e 2, Figura 3).

Convenhamos porém que não é clara a origem desta diferença grupal quanto ao efeito do modo de construção; teria sido conveniente usar intervalos de pré-vozeamento mais pequenos para apreciar efeitos relativamente subtis como este. Note-se que intervalos de 20 milissegundos

foram também usados por outros investigadores (Krause, 1982), apesar de ser mais comum o uso de intervalos menores (e.g., Simon e Fourcin, 1979, usaram 10 milissegundos, Zlatin e Koenigsknecht, 1975, 10 e 5 milissegundos). Obviamente a escolha da diferença inter-estímulos depende não só do objectivo da experiência como também do conhecimento prévio sobre os índices acústicos da distinção perceptiva. Como não dispunhamos de informação relevante a esse nível para os fonemas portugueses, recorreremos neste primeiro estudo a um intervalo que fenomenologicamente parecia apropriado, que não é certamente o ideal.

Concentremo-nos agora na questão do efeito do grupo. Sem dúvida o aspecto mais saliente é a similitude entre os resultados das universitárias e das letradas, apesar de terem nomeado os estímulos sob condições bem diferentes, contrastando com as variações, por pequenas que sejam, verificadas entre os três grupos de adultas mais idosas, que executaram as tarefas em idênticas condições. Quer nos *continua* labiais, quer nos velares se obteve uma interacção entre grupo (letradas *vs* semiletradas *vs* iletradas) e estímulo, que em nenhum dos casos se observou considerando apenas as letradas mais idosas e as universitárias jovens. A observação das funções de identificação sugere que aquela interacção nos grupos das adultas se poderia atribuir ao facto de, em geral, as letradas serem mais consistentes na identificação dos estímulos 2 e 4 do que os outros dois grupos, exibindo uma transição mais abrupta entre as categorias surda e sonora. As semiletradas têm por vezes resultados tão "categóricos" como os das letradas (cf. o *continuum* velar por corte de pré-vozeamento, Figura 3), ou tão "graduais" como os das iletradas (cf. *continua* labiais, na mesma figura).

Seria porém abusivo falar na existência de diferenças grupais quanto à robustez das categorias fonémicas sem apreciarmos a variabilidade inter-individual. Com efeito, resultados médios "categóricos" ou "graduais"

**QUADRO 2:** Exemplos de nomeação categórica, CAT., progressiva, PROG. e dispersa, DISP., segundo Simon e Fourcin, 1978. Apresentam-se exemplos daqueles autores (sujeitos A-E), em que houve 3 observações por estímulo e 8 estímulos, e exemplos dos nossos dados, em que houve 5 observações por estímulo e 6 estímulos.

Sujeitos	Estímulo nº						PADRÃO	
	1	2	3	4	5	6		
A	0	0	1	2	3	3	...	CAT.
B	0	0	0	2	3	3	...	CAT.
C	0	0	2	1	2	3	...	PROG.
D	0	1	1	2	3	3	...	PROG.
E	1	0	3	1	2	2	...	DISP.
Let. 14	0	1	3	4	5	5		CAT.
Ilet. 14	0	2	2	5	4	5		PROG.
Ilet. 10	0	4	4	5	5	4		PROG.

Os exemplos dos nossos dados referem-se às respostas para os estímulos do *continuum* velar por corte de pré-vozeamento.

**QUADRO 3:** Número de sujeitos com uma nomeação de tipo categórico (CAT.) ou progressiva (PROG.) - cf. texto - por grupo e *continuum*. Para facilitar comparações intergrupais, apresenta-se em percentagem a frequência relativa dos sujeitos com um padrão progressivo.

	CAT.	PROG.	% PROG.	
LET.	12	4	25%	VELAR
SLET.	16	1	6%	ADIÇÃO VOZ.
ILET.	13	3	19%	
LET.	12	4	25%	VELAR
SLET.	13	4	24%	CORTE VOZ.
ILET.	6	100	63%	
LET.	12	4	25%	LABIAL
SLET.	6	9	60%	ADIÇÃO VOZ.
ILET.	10	3	23%	
LET.	11	5	31%	LABIAL
SLET.	9	6	40%	CORTE VOZ.
ILET.	9	4	31%	

tanto podem dever-se a diferenças reais quanto à consistência com que os indivíduos identificam os estímulos, como podem ser apenas um artefacto proveniente do cálculo das médias para valores individuais muito variáveis. Ou seja, o padrão de resultados observado pode ser devido a uma maior variabilidade inter-individual no seio das semiletradas e iletradas. O problema das diferenças inter-individuais nem sempre é tido em consideração quando se pretendem esclarecer diferenças grupais, o que poderá criar certa confusão (cf. Simon e Fourcin, 1978, que referem esta lacuna no contexto da percepção da fala, ou Schwartz e Kirsner, 1984, que a discutem no contexto da lateralidade cerebral; como exemplo de uma interpretação baseada quase exclusivamente em medidas de tendência central veja-se Werker e Tees, 1987). Assim, por razões óbvias, apreciemos os resultados individuais.

Sigamos para já a estratégia adoptada por Simon e Fourcin (1978), já referida na Introdução. Trata-se de classificar o padrão das respostas de cada sujeito em três categorias: se o número de respostas sonoras variar de maneira abrupta, sem interrupções ou repetições, diz-se que há nomeação categórica; se a mudança de resposta se dá com interrupção ou patamares (e.g., 0 2 1 3..., ou 0 1 1 2...), a nomeação diz-se progressiva; se finalmente não se observar nenhum dos casos anteriores e as respostas exibirem um padrão quase aleatório, a nomeação é dispersa (cf. Simon e Fourcin, 1978, p. 931; para exemplos daqueles autores e dos nossos resultados, cf. Quadro 2). Aplicando estes critérios às respostas de cada sujeito por *continuum*, o resultado talvez mais saliente é que em nenhum caso se verificou um padrão disperso; todos os sujeitos - em cada *continuum* - se conformam a padrões categóricos ou progressivos, conforme se pode observar no Quadro 3. Salvo duas excepções, a maioria dos sujeitos exhibe um padrão categórico, não sendo evidentes discrepâncias grupais sistemáticas quanto à sua frequência.

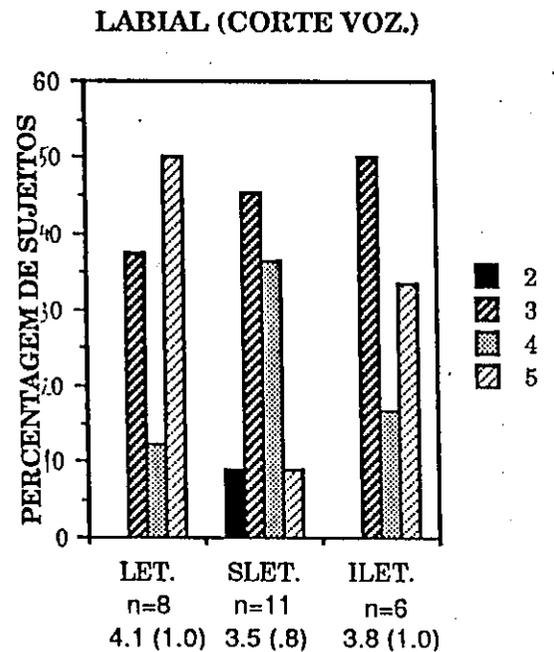
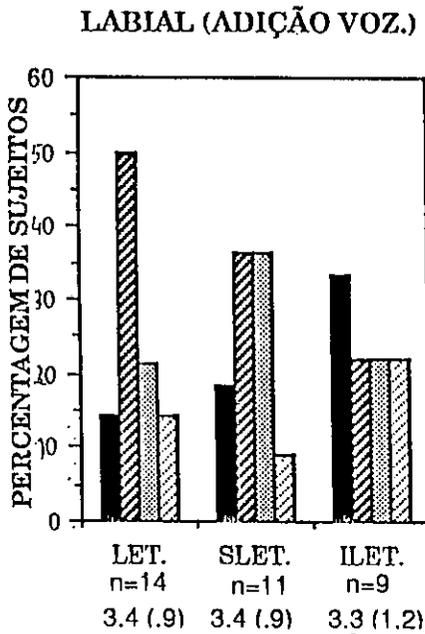
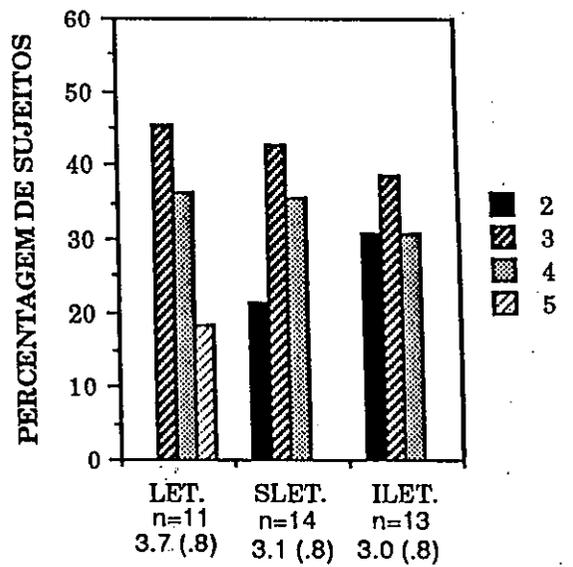
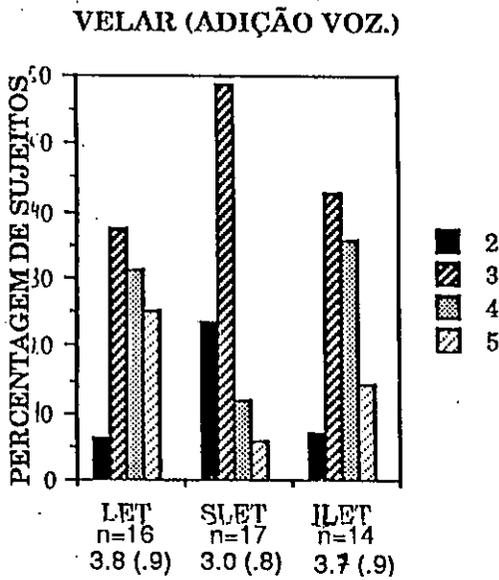
Enquanto examinávamos as respostas individuais para atribuição das categorias atrás definidas, verificámos que talvez elas se adequassem melhor à situação de Simon e Fourcin, onde havia apenas 3 observações por estímulo, do que à nossa, em que tínhamos 5 replicações. Com efeito, segundo os critérios daqueles autores, um sujeito que respondesse, por exemplo, 0 4 4 5 5 5 seria "progressivo", enquanto outro que respondesse 0 1 2 3 4 5 seria "categórico". Apesar de estes serem apenas exemplos fictícios, parecem evidentes as limitações impostas por um critério elaborado para um número máximo possível de três respostas por estímulo (veja-se 0 2 2 3 3 3 *vs* 0 1 2 3 3 3). Para os nossos resultados, o epíteto "progressivo" pode camuflar transições abruptas acompanhadas por alguma falta de atenção momentânea ou incerteza residual. Merecem todavia referência as considerações que se seguem.

Nas letradas, a frequência da nomeação progressiva mantém-se estável nos quatro *continua*, de cerca de 25-30%, um nível semelhante ao observado por Simon e Fourcin em jovens ingleses e franceses de 13 e 11 anos de idade, respectivamente. Exceptuando um caso, acontece o mesmo com as iletradas (frequência de nomeação progressiva entre 19 e 31%); a excepção ocorre no *continuum* velar por corte de pré-vozeamento, onde a função de identificação grupal também distinguia as iletradas (transição menos abrupta) dos restantes grupos. Finalmente, as semiletradas parecem constituir um grupo onde os padrões de resposta são particularmente variáveis.

Observando o número de respostas sonoras para cada um dos estímulos, ficava-se com a impressão de, em geral, haver duas zonas nítidas para cada sujeito, uma caracterizada por poucas ou nenhuma respostas sonoras, e outra constituída quase só por respostas sonoras. A transição parecia clara, mas a sua localização variava conforme o sujeito. Por outro

lado, apesar da rejeição prévia de sujeitos "atípicos", subsistia um número razoável de indivíduos que tinham dado uma ou até duas respostas sonoras aos estímulos com zero milisegundos de vozeamento. Isto deve-se a que o critério estatístico utilizado tolerava uma variabilidade relativamente grande quanto às respostas aos estímulos extremos. Alguns sujeitos já à partida respondiam mais frequentemente "bala" ou "gola", o que introduz diferenças sistemáticas entre os grupos (veja-se o efeito significativo de grupo quanto à nomeação dos *continua* labiais e no velar por adição de pré-vozeamento). Porém, o que agora interessa discutir é se, independentemente do lugar de fronteira fonémica, haverá diferenças intergrupais quanto à consistência de nomeação.

Para dispormos de dados tão pouco contaminados quanto possível por outros factores, restringimos as comparações seguintes aos sujeitos que, para cada *continua*, deram zero respostas sonoras ao estímulo com zero milisegundos de pré-vozeamento, e cinco (100%) respostas sonoras ao estímulo com 100 milisegundos de pré-vozeamento. Para cada sujeito, verificamos qual a diferença máxima entre as respostas sonoras a dois estímulos consecutivos, independentemente da sua localização. Por exemplo, no *continuum* velar por adição de pré-vozeamento, a letrada nº 18 respondeu aos seis estímulos com 0, 4, 5, 5, 5, 5 respostas sonoras, respectivamente; os valores correspondentes para a letrada nº 7 foram 0, 0, 4, 5, 5, 5; nos dois casos, a diferença máxima entre dois estímulos consecutivos foi de 4, o que equivale a dizer que 20 milisegundos de pré-vozeamento foram suficientes para provocar aquele aumento no número de respostas sonoras. A distribuição dos sujeitos por grupo e *continuum* conforme a magnitude da transição pode observar-se na Figura 4. Apresenta-se o número de sujeitos para cada caso sob forma de percentagem relativamente ao total de sujeitos por grupo, pois este varia conforme o *continuum*. A utilização do critério mais restritivo eliminou



**FIGURA 4:** Distribuição dos sujeitos em função da diferença máxima entre dois estímulos consecutivos (cf. Texto). Apresentam-se as médias e os desvios-padrão respectivos, por *continuum* e grupo, bem como o número de sujeitos considerados.

consideravelmente o número de sujeitos tomados para o *continuum* labial por corte de pré-vozeamento; também o número de iletradas obedecendo àquele critério quanto ao *continuum* labial por adição de pré-vozeamento sofreu uma redução razoável. Isto mostra mais uma vez que a resposta "bala" era relativamente mais frequente mesmo com nenhum pré-vozeamento; que isso tenha acontecido aparentemente mais vezes no *continuum* por corte de pré-vozeamento é consistente com a ideia de que haverá outra informação relevante para a distinção perceptiva entre sonora e surda para além do pré-vozeamento.

Nos dois *continua* labiais, cerca de 50% das iletradas e das semiletradas têm uma transição de 4 ou 5; há ligeiramente mais letradas com aqueles valores no *continuum* de corte de vozeamento, 63%, mas menos no *continuum* por adição de vozeamento, 36%. Assim, a distribuição dos sujeitos por magnitude de transição não aponta para diferenças grupais sistemáticas. Como se pode observar também na Figura 4, os valores médios e os desvios-padrões da diferença máxima são semelhantes para os três grupos; análises de variância com grupo como factor intersujeito mostraram que este efeito não é significativo (no *continuum* por adição de vozeamento  $F < 1$  e, no outro,  $F_{(2,24)} = 1.26, p = .30^*$ ).

Quanto aos *continua* velares, nota-se que para todos os grupos a transição mais frequente foi de 3. De um modo geral, salientam-se as distribuições das semiletradas, por a maioria dos sujeitos se concentrar naquela categoria (veja-se também os desvios-padrões respectivos na Figura 4). Convém porém ter presente que há relativamente poucos sujeitos por grupo. Por exemplo, no *continuum* velar por corte de pré-vozeamento aparentemente há mais iletradas com uma diferença máxima de apenas 2; trata-se porém de apenas uma iletrada ( $1/13 = 7.7\%$ ). Uma análise de

\* Para pormenores, cf. Anexo C. 1/7 e C. 1/8.

variância aos valores da transição máxima para esse *continuum* revela não existir um efeito significativo de grupo ( $F_{(2,38)}=1.29, p=.28$ )\*.

Finalmente, no *continuum* velar por adição de pré-vozeamento, a distribuição das letradas e iletradas é bastante semelhante, bem como as respostas médias e desvios-padrões. Uma análise de variância análoga às anteriores revela um efeito de grupo,  $F_{(2,44)} = 3.34, p = .04$ ; como os dados deixam antever, apenas é significativa a diferença entre semiletradas (transição média de 3) e as letradas (transição média de 3.8, cf. Figura 4).

Assim, analisando a magnitude máxima da transição independentemente de ela ocorrer entre os estímulos 2 e 4, ou 0 e 2, etc, não aparecem diferenças grupais sistemáticas. Apenas num caso, o do *continuum* velar por adição de vozeamento, se verificou uma transição mais abrupta nas letradas em comparação com as semiletradas. Poder-se-ia objectar que o critério escolhido, a diferença máxima entre dois estímulos consecutivos, permite que para determinado sujeito se tome um valor elevado - indicando uma transição abrupta - mesmo se para outros estímulos a nomeação fosse mais ou menos inconsistente, por exemplo, um padrão como 0 5 3 2 5 1. De facto, nunca se verificaram casos assim (que equivalem à nomeação dispersa de Simon e Fourcin, 1978), nem outras situações menos extremas, como por exemplo 0 4 5 3 5 5. Houve apenas casos em que a diminuição do número de respostas sonoras para estímulos com mais vozeamento foi de um (ou seja, por exemplo, 0 0 3 5 4 5\*\*).

Enfim, a reanálise efectuada indica que a interacção entre grupo e estímulo, observada quando sujeitos adultos de vários níveis de escolarização identificam *continua* de vozeamento, se pode atribuir ao facto de o lugar da fronteira fonémica ser mais variável nos grupos das iletradas e

\* Cf. Anexo C.1/9.

\*\* Resultado obtido pela letra n.º 12 no *continuum* velar por adição de pré-vozeamento.

semiletradas; os nossos dados não mostram uma nomeação menos categórica por parte dos sujeitos iletrados, nem a indicam claramente para sujeitos menos escolarizados. Para o estabelecimento mais firme daquelas asserções é importante testá-las directamente, empregando *continua* mais discriminativos, ou seja, com uma diferença entre estímulos consecutivos mais pequena, de 10 ou mesmo 5 milisegundos. Conseguiríamos assim observar mais finamente o comportamento dos sujeitos na zona da fronteira fonémica.

## CAPÍTULO VII

INTEGRAÇÃO DE INFORMAÇÃO FONOLÓGICA E ASPECTOS  
ESPECÍFICOS DO CONHECIMENTO DA ESCRITA

Encerraremos este nosso contributo para o esclarecimento do papel da alfabetização em certos aspectos do processamento da linguagem falada com a apresentação de uma experiência sobre a fusão fonológica. A análise aprofundada, a este fenómeno, efectuada ao longo do Capítulo III, permitiu salientar que ele é mais afectado por variáveis relacionadas com aspectos específicos à linguagem do que por variáveis de tipo acústico-auditivo. A experiência que será objecto deste capítulo pretende justamente averiguar um desses aspectos, a representação ortográfica, através da comparação entre sujeitos letrados e iletrados. Esta comparação não é relevante só para o esclarecimento da possível influência da escrita na fusão fonológica; na medida em que este fenómeno envolve processos relativamente centrais de integração de informação, uma tal comparação é importante para esclarecer a natureza das diferenças entre estes grupos de sujeitos, observadas noutras situações de processamento da fala.

Fusão fonológica e ortografia: comparação de letradas, semiletradas e iletradas (Experiência VII)

### 1. Introdução

Em algumas das experiências que efectuámos observaram-se incidentalmente respostas de fusão fonológica. Na medida em que elas

ocorreram com relativamente baixa frequência, e o próprio planeamento experimental não foi desenhado com esse objectivo, limitámo-nos a relatar que na situação de audição dicótica alguns sujeitos pareciam combinar os fonemas dos estímulos em competição, produzindo por exemplo, a resposta "balde" perante o par "bode + pala". Este tipo de resposta não se restringia a nenhum grupo em particular, quer em termos de idade, quer de escolarização. Notaram-se porém algumas variações grupais: sem pretender generalizar, os dados provenientes daquelas observações incidentais foram tal que o número bruto de fusões foi maior nas adultas com reduzida ou nenhuma escolarização (cf. Experiência III) e quer o número bruto, quer a sua proporção relativamente ao total de erros foram superiores em crianças pré-letradas e no início da escolarização (cf. Experiência IV). Aquelas diferenças poder-nos-iam por si levar a especular sobre o eventual papel da escolarização na tendência em integrar informação fonémica apresentada aos dois ouvidos.

Há porém uma razão mais forte para elaborar uma situação experimental com o propósito de estudar a incidência de fusões fonológicas em grupos de diferentes graus de escolarização. Ela prende-se com a distinção entre processos obrigatórios e opcionais de percepção da fala e a dissociação do efeito da alfabetização naqueles processos. Tínhamos proposto que o facto de ser letrado não afectaria os processos envolvidos na percepção da fala que se situassem a nível pré-atencional ou fossem de tipo modular. Um dos argumentos empíricos apresentados foi que a incidência de *blendings* ou fusões de traços fonéticos era semelhante para sujeitos letrados e iletrados. Ora a fusão fonológica é um fenómeno de integração perceptiva, em certos aspectos análoga aos *blendings*, mas talvez deles distinta quanto a outros aspectos (Cutting, 1976). Um dos critérios diferenciadores entre *blending* e fusão fonológica poderá ser justamente a sensibilidade a factores atencionais (cf. Sexton e Geffen, 1981; e também Cap. III). A comparação da incidência

de *blendings* e de fusões fonológicas em sujeitos letrados e iletrados poderá ser esclarecedora quanto a vários aspectos, que passamos a enunciar.

Em primeiro lugar, se se verificarem diferenças entre letrados, semiletrados e iletrados na própria incidência de fusões fonológicas, dado o facto de elas não ocorrerem quanto à incidência de *blendings* (cf. Experiência IV) teremos encontrado indicação de que os processos envolvidos naqueles dois fenómenos não são (integralmente, pelo menos) os mesmos; seguindo o trabalho teórico de Cutting (1976), teríamos assim um argumento em favor da distinção entre níveis diferentes de processamento linguístico, responsáveis respectivamente pela integração de traços fonéticos e pela integração de informação fonológica.

Em segundo lugar, conforme o padrão de eventuais diferenças entre letrados e iletrados poderemos ganhar algum conhecimento sobre os factores que, por um lado, determinam a fusão fonológica e, por outro, diferenciam os sujeitos ou grupos em questão. Encaremos para já duas possibilidades, a estratégia atencional e a representação ortográfica. Sexton e Geffen (1981) verificaram que a atenção dirigida a um só ouvido contribuía para uma ligeira redução de fusões fonológicas, como se o indivíduo fosse capaz de bloquear a informação fonológica do canal não-atendido. Porém, isto não acontecia sempre e, mesmo sem os sujeitos disso terem consciência, a informação do canal não-atendido era por vezes integrada com a do canal atendido desde que ambas fossem fonologicamente compatíveis. Propomos-nos efectuar uma manipulação da atenção semelhante à usada por aqueles autores. Visamos com isso um duplo objectivo: encontrar evidência convergente sobre o papel de factores atencionais nas fusões fonológicas; e explorar mais um aspecto da influência da alfabetização segundo a hipótese de que ela se circunscreve a fenómenos susceptíveis de controlo atencional.

Relativamente à exploração do papel da representação ortográfica, vejamos as seguintes considerações. Como se viu, a fusão fonológica parece ser essencialmente determinada pela compatibilidade fonológica entre dois estímulos diferentes. A questão que colocamos é até que ponto essa compatibilidade é estritamente fonológica, ou seja, envolve o acesso à representação fonológica das palavras, ou se haverá também uma activação da representação ortográfica.

Com efeito, se a fusão envolve um "processador" central em que é combinada informação fonológica, tal que essa combinação é afectada por variáveis como o estatuto lexical e, talvez, a frequência das palavras (Cutting, 1975, Cutting e Day, 1975), torna-se plausível que outros tipos de "conhecimento" sobre os estímulos candidatos a fusão sejam relevantes. A ortografia, ou a suposta representação gráfica interna, é um desses "conhecimentos" linguísticos, cuja influência em tarefas com linguagem oral foi já demonstrada (e.g., Seidenberg e Tanenhaus, 1979; cf. Cap. V.1.). Quanto ao seu papel na fusão fonológica, não foram feitas observações sistemáticas. Vejamos todavia se os trabalhos publicados permitem uma reanálise da fusibilidade dos vários pares de palavras atendendo à sua ortografia. A questão é se a palavra com encontro consonantal tem uma ortografia congruente com a das palavras componentes respectivas e se nesses casos há mais fusões; por exemplo, foi usado o par "bed + red" como estímulo para a fusão "bread", "cam + lamb" para "clam", "cane + rain" para "crane", etc (os dois primeiros em Cutting, 1975 e Cutting e Day, 1975; o último em Sexton e Geffen, 1981).

As observações disponíveis são descritas a seguir. No conjunto dos oito pares de palavras usadas por Cutting (1975; Experiência 1), podemos distinguir dois para os quais a fusão envolve uma grafia diferente da mera justaposição das palavras do par: são eles "cam + lamb = clam" e "bed + red

= bread". Admitamos que estes dois casos são bem distintos - no primeiro temos apenas uma consoante adicional, e no segundo uma grafia diversa para a vogal. Admitamos também, desde já, que o alcance desta reanálise só pode ser limitado. Uma das razões mais óbvias é a possibilidade de grafias alternativas, como o último caso deixa antever. Existem em inglês as palavras "bred" e também "red". Todavia, não terá sido por acaso que o autor apresentou os seus estímulos com a grafia "bed + red = bread" (*ibidem*, p. 107): estas representarão as palavras mais comuns para aquela fonologia. Ora foi justamente para estes dois pares de estímulos que se verificaram as frequências mais baixas de fusão, nomeadamente para as suas versões em fala natural: para "bed/red" observou-se 19%, e para "cam/lamb" 17%, de fusões. Os valores para os pares mais similares e ortograficamente congruentes foram de 33 e 29% (para "bed/led" e "cam/ram", respectivamente). Em versão sintética, as diferenças são muito menos acentuadas; mas, se as há, elas vão no mesmo sentido: 48% para "bed/led" vs 46% para "bed/red" e 63% para "cam/ram" vs 57% para "cam + lamb".

No estudo de Sexton e Geffen (1981), entre os dez pares fundíveis há três cuja ortografia não é perfeitamente congruente com a da fusão: "cane + rain = crane", "gate + rate = great" e "dead + red = dread". Eles encontram-se entre os pares de palavras com menos fusões. Finalmente, no trabalho de Cutting e Day (1975), a comparação entre fusões em posição inicial e final (Experiência 3), em que se encontrou muito menos fusões para *clusters* em posição final, todos os estímulos apropriados a fusão em posição final terminavam em /ed/ ("peeled; peered, paled, paired"), apesar de serem pronunciados /ld/ ou /rd/. Claro que todas estas comparações estão confundidas em larga medida com o tipo de *cluster*, com /l/ ou com /r/, e com a frequência das palavras. Elas são meramente indicativas de que a possibilidade de uma influência da representação ortográfica não é à partida desmentida por observações incidentais.

Um modo de avaliar se a representação ortográfica tem algum efeito facilitador ou inibidor na fusão fonológica consiste justamente na comparação entre sujeitos letrados, que têm acesso a essa representação, e sujeitos iletrados, que apenas usam a linguagem oral. A percentagem de fusão fonológica, consoante a ortografia da fusão seja apenas a justaposição da dos pares fundíveis, ou não, deveria ser diferente para sujeitos letrados e iletrados se a representação ortográfica afectar a fusão. O português, uma língua moderadamente fonética, apresenta alguns casos interessantes de *clusters* orais em fala corrente, cuja escrita separa em sílabas diferentes: [mlasu], [sleumα], [plu](ə)], etc. Se a ortografia afectar a fusão fonológica perante estímulos deste tipo, sujeitos letrados deveriam ter uma menor incidência de fusões do que sujeitos iletrados. Na medida em que se trata dos mesmos estímulos, é a comparação intergrupo que permite especificar o efeito em termos de ortografia; de outro modo, se sujeitos letrados dessem menos fusões a, e.g., "maço + laço", do que a "taça + raça", poder-se-ia tratar de diferentes graus de fusibilidade dos estímulos.

Na experiência a seguir apresentada, usámos palavras portuguesas mais ou menos correntes, que pudessem ser combinadas produzindo uma fusão. Todos os estímulos foram palavras para evitar possíveis artefactos devidos à tendência ou relutância em produzir não-palavras. Dos dez pares de palavras definitivamente usadas como estímulos, quatro tinham uma ortografia de duas sílabas enquanto a possível fusão tinha uma ortografia de três sílabas (cf. mais à frente o Quadro 1).

A tarefa dos sujeitos era dizer o que ouviam, como acontece na maioria dos estudos sobre fusões fonológicas. Os estímulos eram em fala natural. Como é sabido que há menos fusões em fala natural do que em fala sintética (e.g., Cutting, 1975), para maximizar a oportunidade de observar fusões os estímulos foram submetidos a algumas modificações prévias que

se explicam na rubrica Método. Os sujeitos realizaram a tarefa em duas condições, uma de atenção dividida, em que lhes era dito que prestassem atenção aos dois ouvidos, e outra de atenção concentrada, em que lhes era indicado que prestassem atenção só a um ouvido. Supunhamos que a incidência de fusões seria menor na condição de atenção concentrada, e que encontraríamos menos fusões nos sujeitos letrados do que nos iletrados para os pares de ortografia não congruente com a da palavra fundida. Esta diferença intergrupo ocorreria apenas, ou seria mais marcada, na condição de atenção concentrada.

## 2. Método

### 2.1. Sujeitos

Participaram nesta experiência 90 sujeitos do sexo feminino, com idades compreendidas entre os 36 e os 65 anos e com diferentes graus de escolarização. Vinte e oito tinham, no mínimo, doze anos de escolarização, sendo a maioria (22) licenciadas. Trabalhavam como docentes do ensino secundário\* e a sua idade era, em média, 46.8 anos (desvio-padrão de 8.1), variando entre 35 e 64 anos. Estes sujeitos constituem o grupo das letradas. O grupo das semiletradas é composto por sujeitos com, no máximo, 4 anos de escolarização. Embora lendo e escrevendo com variável grau de eficácia, não usam regularmente estas actividades no seu dia-a-dia. Trabalhavam em serviços domésticos e agrícolas,\*\* e a sua idade variava entre os 36 e 65 anos (média de 48.0, d.p. = 7.1). O terceiro grupo era constituído por 34 iletradas, que por razões sócio-económicas não tinham frequentado a escola

---

\* Zona do Porto e Grande Porto, Escolas Secundárias Clara de Resende, de Águas Santas e de Gondomar.

\*\* A maioria era proveniente do meio agrícola (Macieira, Vale de Cambra e Sanguedo, Vila da Feira).

enquanto crianças ou adultas. Faziam trabalhos domésticos e/ou agrícolas\* e a sua idade média era de 51.6 (d.p. = 5.9), variando entre 40 e 65 anos. Para minimizar diferenças de idade entre os grupos, optou-se por não integrar na análise os resultados das 6 iletradas mais idosas; ficamos assim com o mesmo número de sujeitos em cada grupo, 28, e com a idade média das iletradas a 49.8 anos (d.p.=4.9).

Todas as participantes foram submetidas a um teste prévio de acuidade auditiva, constituído por 12 palavras CVCV (6 a cada ouvido) em dois níveis decrescentes de intensidade (70 e 50 dB SPL); se fossem notadas perdas auditivas uni ou bilaterais, o sujeito era substituído. Nenhuma das participantes tinha história conhecida de problemas neurológicos, intelectuais ou auditivos.

## 2.2. Estímulos

Os estímulos consistem em 10 pares de palavras cuja sílaba inicial é de tipo CV, de modo a que mediante a junção das consoantes iniciais de cada par seja possível formar uma terceira palavra, cuja sílaba inicial é de tipo CCV\*\*. Estes estímulos foram selecionados com base num pré-teste e preparados como a seguir se descreve.

Escolheram-se várias tríades de palavras, do tipo /clara-cara-lara/ ou /prato-pato-rato/, constituídas por uma palavra, a que chamaremos total, e por um par de palavras de estrutura fonética mais simples, em que a total pode ser decomposta. Designaremos estas por pares componentes ou fundíveis. Foram pronunciados vários exemplares destas palavras por uma falante portuguesa numa frase-contexto, de modo a que os encontros

---

\* O seu meio era o mesmo das semiletradas.

\* \*Entendendo aqui a palavra falada, e não a sua representação ortográfica.

consonantais soassem o mais natural possível. Esta precaução foi tomada devido à natureza de algumas tríades, em que a pronúncia enquanto palavra isolada ou em contexto afecta a própria existência de encontro consonantal: repare-se, por exemplo, nas realizações de "feliz" em pronúncia cuidada, [fɔ'li] ou corrente, [fli].

A gravação das palavras foi feita numa câmara acústica isolada para um gravador Ampex. Os "melhores" exemplares de cada tríade (os mais claros e auditivamente parecidos quanto à intensidade, duração e entoação) foram digitalizados usando o sistema PCM do computador VAX 750 dos Laboratórios Haskins, com uma frequência de amostragem de 20 KHz. Os sonogramas destes sinais foram depois analisados e manipulados por meio do programa Wendy ("Wave Editing and Display"). Com o objectivo de tornar os pares componentes o mais fundíveis possível, reconstruiu-se artificialmente um elemento do par, mantendo a porção inicial de origem e substituindo a sua porção final pela porção final do outro elemento. Por exemplo, para o par /cara-lara/, tomou-se de /lara/ a porção final da primeira sílaba e a segunda sílaba, percebida como [(a)rα]; eliminou-se a porção equivalente de /cara/, substituindo-a pela extraída de /lara/. Fica-se assim com duas palavras distintas, [karα] e [larα], que são parcialmente idênticas do ponto de vista acústico. Este tipo de manipulação é usado quando por uma razão ou outra se pretende acentuar a semelhança acústica dos estímulos; dela resultam pares de palavras designados por parcialmente fusionados\* (cf. Repp, 1977-a; 1980).

---

\* Não se trata de pares quase completamente fusionados - designados por pares fusionados - em que se reduz ao mínimo a porção distintiva de cada elemento de modo a que apenas se perceba um estímulo (um elemento ou outro) se as palavras forem apresentadas dicoticamente (cf. Wexler e Halwes, 1982). No nosso caso, era possível por esforço de atenção perceber individualmente os elementos do par apresentado dicoticamente.

QUADRO 1: Palavras usadas na constituição dos testes de identificação monaural e dicótica, com vista à observação de fusões fonológicas.

A. Encontros consonantais com a lateral /l/

1. Clara	cara	+	lara	
2. Plena	pena	+	lena	
3. Pelar	par	+	lar	([plar])
4. Feliz	fiz	+	lis	([fli:])

B. Encontros consonantais com /r/

1. Prato	pato	+	rato	
2. Pregar	pegar	+	regar	
3. Grama	gama	+	rama	
4. Frita	fita	+	rita	
5. Grato ou Regato	rato	+	gato	([Rgat(u)])
6. Bruço ou Rebuço	ruço	+	buço	([Rbus(u)])

Outro tipo de manipulação efectuada visou a constituição de pares componentes coarticulados. Trata-se de, a partir do estímulo total, e.g., [klarα], constituir os estímulos componentes [karα] - extraíndo a porção ouvida como [l(ə)], e [larα] - extraíndo a porção ouvida como [k(l)]. Estes novos estímulos, versões truncadas do original, são *ipso facto* coarticulados; referi-los-emos sucintamente por pares coarticulados, distinguindo-os dos restantes pares ditos normais.

Finalmente, a terceira manipulação consistiu em sincronizar o início dos elementos de cada par, de modo a que em posterior audição dicótica aparecessem simultaneamente, um em cada ouvido.

Os pares resultantes de 5 das 17 tríades iniciais\* não suscitaram respostas de fusão quando ouvidos por dois auditores, ambos cientes do objectivo do teste\*\*. Selecionaram-se as versões normais e coarticuladas correspondentes às 12 tríades restantes, que se reconverteram em sinal analógico e gravaram dicoticamente numa fita magnética em sequência aleatória, com um intervalo inter-estímulo de 5 segundos. A gravação em fita foi passada a cassette num Revox A77.

Com base nos resultados de um pré-teste realizado com estudantes universitários\*\*\* eliminaram-se mais duas tríades ([mlasu - masu - lasu] e [trasα - tasα - rasα]), que não suscitaram quaisquer respostas de fusão. Os 10 pares definitivos (cf. Quadro 1) foram usados para constituir dois testes de identificação, um monaural e outro dicótico.

O teste monaural apresentava todas as palavras componentes, normais e coarticuladas, e ainda a palavra total, perfazendo 50 estímulos

\* Uma lista destas tríades pode ser vista no Anexo FF.1.

\*\* A autora e B. Repp.

\*\*\* Ano lectivo de 1985/86, estudantes do 1º ano do Curso de Psicologia da Faculdade de Psicologia e de Ciências da Educação, Universidade do Porto.

((10 x 2 x 2) + 10), em ordem aleatória. No teste dicótico, cada par de palavras normais e coarticuladas foi gravado duas vezes, cada elemento aparecendo uma vez num canal e outra vez no outro; fizeram-se três sequências aleatórias deste tipo, cada uma com 40 estímulos. A apresentação do material foi feita através de um gravador Phillips D 6920 MK2 e auscultadores Philips EM 6421, com um nível sonoro de cerca de 70 dB SPL.

### 2.3. Procedimento

A experiência foi feita em sessões individuais, realizadas numa sala sossegada\*. Após o teste sumário de acuidade auditiva, explicava-se ao sujeito que iria "ouvir palavras ditas por uma voz de mulher, através dos auscultadores"; a sua tarefa consistia "apenas em repetir o que vai ouvindo: cada vez que aparece uma palavra, por favor diga exactamente como ouviu, mesmo que não tenha a certeza ou lhe pareça disparate. Algumas palavras são um pouco estranhas; por exemplo, começam por [ɕa] ou [ɕα] ou [ɕi] ou [ɕu]. Se não perceber bem algumas palavras, não faz mal: eu repito e depois tornamos a ouvi-las. Vamos começar por ouvir as palavras só num ouvido". Seguiu-se o teste de identificação monaural, em que metade dos estímulos era apresentada a um ouvido, e outra metade ao outro ouvido (25 + 25). A ordem de ouvido estava contrabalançada dentro de cada grupo.

Após o teste monaural, explicava-se aos sujeitos que passariam a ouvir as palavras nos dois ouvidos. O teste dicótico era feito em duas condições, uma de atenção dividida, e outra de atenção concentrada. Na condição de atenção dividida, os sujeitos eram instruídos a prestar atenção aos dois ouvidos e a repetir o que ouvissem; se percebessem duas palavras,

---

\*Para a maioria dos sujeitos iletrados e semiletrados, a experiência desenrolou-se na sua própria casa, ou na casa de uma das participantes onde as outras se deslocavam.

deveriam relatá-las ambas. Na condição de atenção concentrada, instruía-se o sujeito a prestar atenção só a um ouvido, e repetir o que ouvisse apenas nesse ouvido. Todos os sujeitos prestavam atenção primeiro ao ouvido esquerdo e depois ao direito. A ordem das condições, atenção dividida ou concentrada, foi contrabalançada em cada grupo.

Entre o teste de acuidade auditiva e a experiência propriamente dita, alguns sujeitos participaram na experiência de nomeação de *continua* de vozeamento. Não se notaram diferenças entre as suas respostas e as dos restantes sujeitos, pelo que os dados foram analisados em conjunto.

### 3. Resultados

#### 3.1. Identificação monaural

A identificação das palavras apresentadas monauralmente não se fez com igual facilidade para todas as tríades. Os componentes coarticulados foram incorrectamente percebidos em muitos casos, pelo que trataremos separadamente os resultados obtidos para as palavras normais e as coarticuladas. Consideremos primeiro as normais.

##### 3.1.1. Estímulos normais

A observação do padrão de respostas correctas para os vários estímulos revela inequivocamente a existência de subgrupos de tríades com diferentes graus de inteligibilidade (cf. Quadro 2). Temos assim as tríades "clara", "pelar" e "grama", que são claramente percebidas: considerando todos os sujeitos, e os três estímulos de cada tríade houve apenas 0.7% de erros. A maior parte destes erros ocorreu em resposta ao estímulo iniciado

**QUADRO 2:** Resumo das respostas dos sujeitos aos vários estímulos de cada tríade em apresentação monaural. Apresenta-se o número de erros por grupo. Na última coluna mostram-se os erros, entre aspas, e a respectiva frequência, entre parêntesis. Indicam-se apenas os erros mais frequentes.

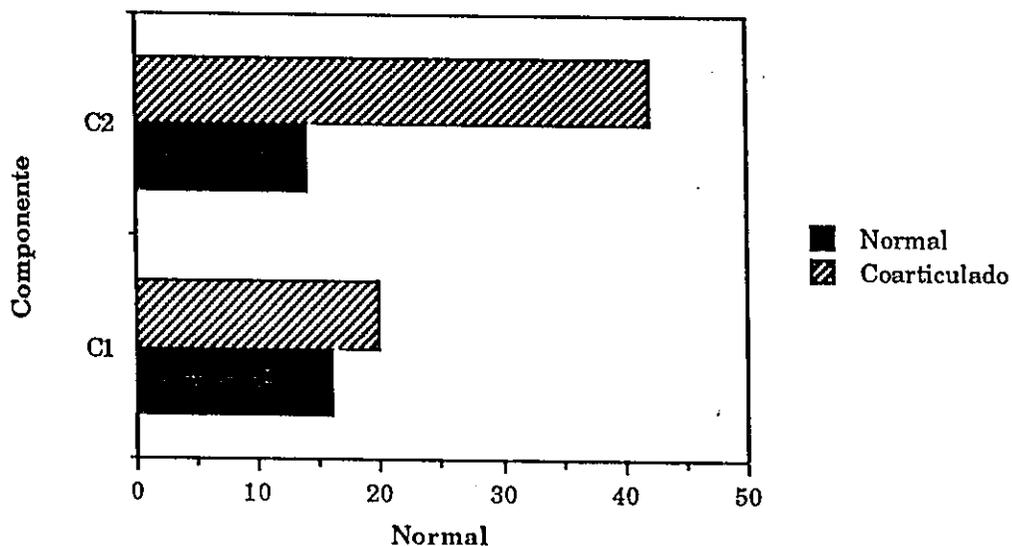
		LET. n=28	SLET. n=28	ILET. n=28	TOTAL	
T <sub>1</sub>	CLARA	0	0	0	0	-
	CARA	0	0	0	0	-
	LARA	0	1	2	3	"clara" (3)
T <sub>2</sub>	PELAR	0	0	0	0	-
	PAR	0	1	0	1	"pelar" (1)
	LAR	0	1	1	2	"pelar" (1)
T <sub>3</sub>	GRAMA	0	0	0	0	-
	GAMA	0	1	0	1	-
	RAMA	1	1	1	3	"grama" (1)
T <sub>4</sub>	PRATO	0	0	0	0	-
	PATO	0	0	0	0	-
	RATO	3+1	4+2	7+1	18	"gato" (14), "prato" (10)
T <sub>5</sub>	PLENA	0	9	2	11	"(he)lena"
	PENA	1	0	1	2	"lena"
	LENA	5	1+1	4	11	"plena" (10)
T <sub>6</sub>	FELIZ	1	0	0	1	"lis"
	FIZ	6+5	5	4	20	/riz/ (6)
	LIS	2	5	3	10	"feliz" (9)
T <sub>7</sub>	FRITA	0	0	0	0	-
	FITA	21	12	10	43	"rita" (33)
	RITA	0	0	0	0	-
T <sub>8</sub>	PREGAR	27	26	27	80	"pegar"
	PEGAR	1	4	5	10	"lagar", /dagar/
	REGAR	3	4	3	10	"pegar"
T <sub>9</sub>	REGATO	8+2	10	17	37	"gato"
	GATO	0	0	0	0	
	RATO	2	10	15	27	"gato"
T <sub>10</sub>	REBUÇO	3	2	7	12	"buço", "produto"
	BUÇO	5	14	12	31	/gusu/, /lusu/ etc
	RUÇO	18	20	25	63	"curso", /pusu/
<b>TOTAL</b>		<b>115</b>	<b>134</b>	<b>159</b>		

por líquida ou vibrante\*, e consistiu na percepção de palavra com encontro consonantal (ouvir /clara/ em vez de /lara/, por exemplo).

Um outro subgrupo de tríades apresenta pelo menos um elemento com uma percentagem de erros relativamente elevada, de cerca de 20%. É o caso de "prato", "plena", e "feliz". Em resposta aos estímulos constituintes da primeira tríade, "pato" e "prato" foram correctamente identificados por todos os sujeitos. Todavia, o estímulo "rato" foi frequentemente confundido confundindo-o com /gato/ ou, por vezes com /prato/. Na tríade seguinte, "plena" confundida com "lena" ou "helena" por cerca de 30% das semiletradas, e 7% das iletradas. O estímulo "lena" foi confundido com "plena" por cerca de 15% das letradas e das iletradas. Na tríade composta por "feliz", "fiz" e "lis", o estímulo total é praticamente sempre bem identificado. "Lis" foi confundido com "feliz" por cerca de 10% dos sujeitos. Identificações erradas foram particularmente frequentes para o estímulo "fiz", mormente nas letradas. Os erros consistiam sempre em trocas de primeira consoante, sendo /riz/ a resposta mais frequente, que ocorreu apenas nas letradas.

Das restantes quatro tríades há a notar uma inteligibilidade ainda mais reduzida em pelo menos um elemento. O estímulo "fita" é confundido com /rita/ por quase 40% dos sujeitos. Este resultado é semelhante ao observado com /fiz/, da tríade "feliz", em que também se verificou uma dificuldade em entender a fricativa [f] e a sua confusão com [R]. O estímulo "pregar" é praticamente sempre ouvido como "pegar" (cf. Quadro 2). Na tríade "regato", apenas o componente "gato" obteve 100% de respostas correctas. O componente "rato" foi ouvido como /gato/ nomeadamente pelas semiletradas e iletradas. Na última tríade, nota-se uma grande dificuldade

\* Por questão de simplicidade de exposição, referir-nos-emos ao batimento [r] em [caro] e à vibrante [R] em [carro] por vibrante.



**FIGURA 1:** Percentagem de identificações incorrectas para estímulos normais e coarticulados. Compara-se a percentagem de erros para os estímulos em que se removeu /l/ ou /r/ com a dos estímulos normais correspondentes (e.g., /c(l)ara/ vs /cara/), e a percentagem de erros para os estímulos em que se removeu a oclusiva ou fricativa com a dos estímulos normais correspondentes (/c)l ara/ vs /l ara/, /f)ri ta/ vs /ri ta/, etc. Designaremos os primeiros por Componente 1, C<sub>1</sub>, e os segundos por Componente 2, C<sub>2</sub>. As percentagens foram calculadas com base no número de erros observados relativamente ao total possível (10 estímulos x 28 sujeitos x 3 grupos).

na percepção de "ruço". A resposta mais comum é /curso/, que ocorre 56% das vezes, seguindo-se [pusu] (7% das respostas). O componente /buço/ tem também uma inteligibilidade relativamente baixa, cerca de 37% de identificação incorrectas. A palavra total nem sempre foi correctamente identificada, mormente pelas iletradas.

### 3.1.2. Estímulos coarticulados

De uma maneira geral, os erros foram mais frequentes ao identificar os estímulos coarticulados. Estes são de dois tipos: aqueles em que foi removida a líquida ou a vibrante, por exemplo, /c(l)ara/, /p(r)ato/, e aqueles a que foi removido o outro elemento do *cluster*, /(c)lara/, /(p)rato/, etc. Na Figura 1 compara-se a percentagem de erros para aqueles dois tipos de estímulos coarticulados com a percentagem de erros nas versões normais correspondentes.

Pode observar-se que há ligeiramente mais erros na identificação dos estímulos coarticulados em que foi removida a líquida ou a vibrante do que nos estímulos correspondentes normais. A diferença de inteligibilidade é bem mais pronunciada ao comparar os estímulos coarticulados por remoção de oclusiva ou fricativa com os estímulos normais correspondentes (e.g., /(f)rita/ e /rita/, respectivamente). Este aumento no número de identificação incorrectas deve-se, nalguns casos quase exclusivamente, a que muitos sujeitos ouvem o estímulo componente coarticulado como estímulo integral total. Isto acontece com /(c)lara/, /(f)rita/ e /(g)rama/. Os outros erros consistem geralmente em trocas de primeira consoante.

Podemos pois concluir que a manipulação efectuada, principalmente a remoção da oclusiva ou fricativa do *cluster* original, prejudicou a inteligibilidade dos estímulos. Possivelmente nalguns casos

não foi suficiente para retirar informação acústica indicadora do *cluster*; noutros, contribuiu para aumentar a ambiguidade do estímulo. Não se justifica portanto uma análise da incidência de fusões fonológicas com base na apresentação dicótica destes estímulos coarticulados.

### 3. 2. Apresentação dicótica de pares fundíveis

Como os estímulos coarticulados eram pouco inteligíveis e suscitavam frequentemente respostas com *cluster*, para analisar as fusões fonológicas consideramos apenas os pares normais.

Uma inspecção preliminar aos resultados revelou não existirem respostas de fusão para alguns pares, nomeadamente "fita/rita", "pegar/regar", "rato/gato" e "ruço/buço". Estes pares são justamente aqueles que tiveram piores resultados em apresentação monaural. O padrão de respostas foi idêntico nos três grupos: ao par "fita/rita", os sujeitos respondiam "rita"; ao par "pegar/regar" os sujeitos respondiam "pegar"; "rato/gato" era em geral ouvido como "gato". Estes resultados são consistentes com os observados sob apresentação monaural: repare-se que a resposta ao par dicótico consiste justamente na palavra mais frequentemente ouvida na identificação monaural dos elementos da tríade (e.g., "rato" e "regato" foram entendidos como "gato" por quase metade dos sujeitos). As respostas a "ruço/buço" eram mais variáveis ("curso", "ruço", "buço", etc), como já acontecera em apresentação monaural.

Nos seis pares restantes houve respostas de fusão, em grau variável conforme o grupo e o próprio par. Antes de passarmos à análise das fusões, convém resolver uma questão prévia: as respostas, nomeadamente das iletradas e semiletradas, incluíam além das fusões algum outro tipo de identificações incorrectas? Este é um ponto importante, pois dele depende o

**QUADRO 3:** Resumo dos erros observados perante apresentação dos pares fundíveis indicados, por grupo, Let., letradas, Slet. semiletradas, Ilet. iletradas. Consideram-se erros as identificações incorrectas com excepção da fusão.

	LET.	SLET.	ILET.	
Cara/lara	0	0	2	"mara"
Pena/lena	0	0	0	-
Par/lar	0	1	2	"mar"
Fiz/lis	1	0	1	/ris/
Gama/rama"	0	0	0	-
Pato/rato	3	4	4	"gato"*

**QUADRO 4:** Percentagem média de fusões em atenção dividida e concentrada, por grupo. As percentagens foram calculadas tendo em conta apenas os seis pares anteriores (cf. Quadro 3).

	LET.	SLET.	ILET.
ATENÇÃO			
DIVIDIDA	56.5	60.7	81.5
CONCENTRADA	55.1	60.4	85.4

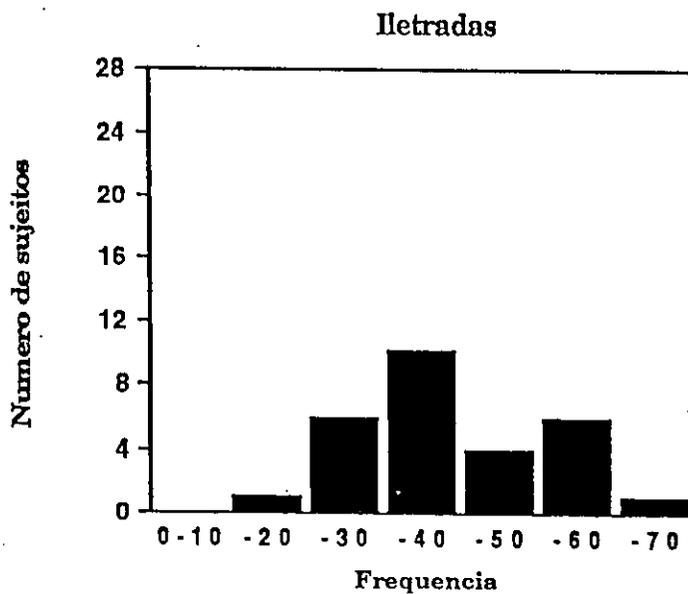
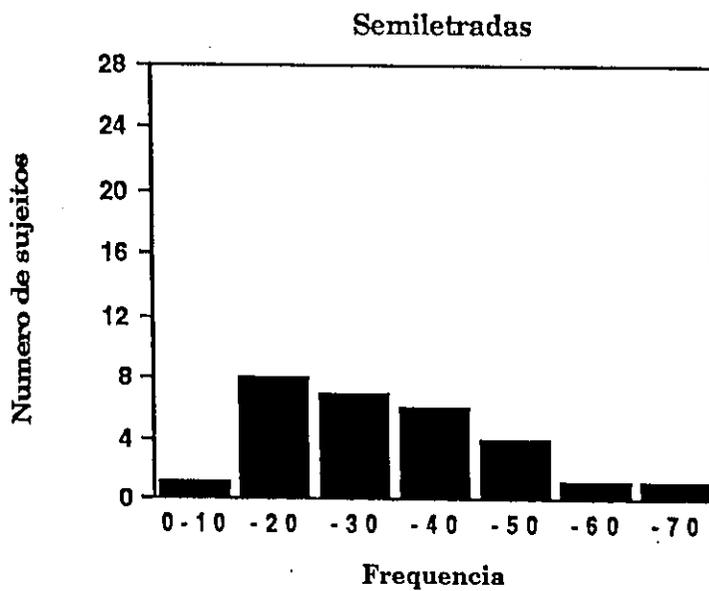
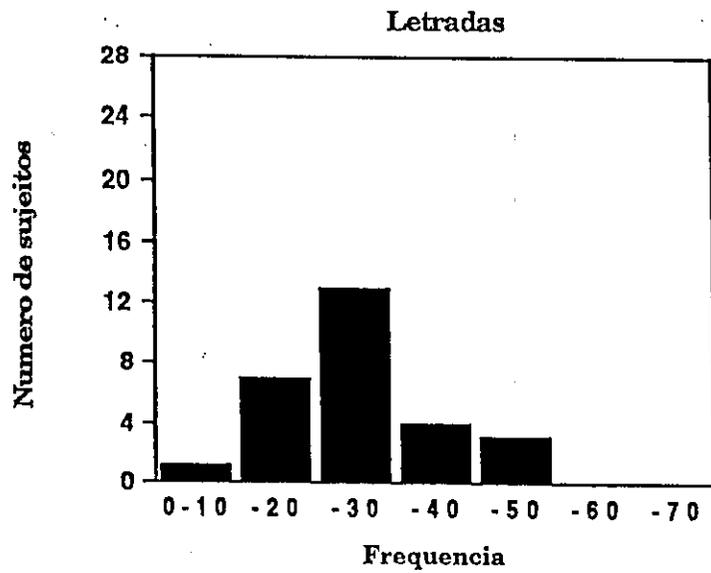
\* Estes sujeitos nunca responderam com fusão.

facto de se fazerem as análises subsequentes considerando o número de fusões observadas, ou a sua proporção relativamente ao total de erros. A inspecção dos resultados individuais para cada par mostrou que praticamente todas as respostas se distribuem por duas categorias: ou fusão, ou identificação de uma das palavras componentes, seja a iniciada por oclusiva ou fricativa, ou a iniciada pela líquida ou vibrante. Outras respostas foram raríssimas (cf. Quadro 3) e, em geral, observam-se em sujeitos que nunca ouvem fusão fonológica para o par a que dão tal resposta atípica. Podemos pois usar o número bruto de fusões nos tratamentos subsequentes, sem o receio de que, para cada sujeito, haja mais fusões só porque há mais identificações erradas deste ou daquele tipo.

A inspecção dos resultados sugeria a inexistência de um efeito da atenção, quer em geral, quer considerando cada par separadamente. Por esta razão fez-se uma análise preliminar ao número de fusões observadas em cada grupo, conforme a atenção fosse dividida ou concentrada; veja-se no Quadro 4 os valores respectivos, apresentados em percentagem por questão de clareza. Uma análise de variância com grupo como factor intersujeito e atenção como factor intrasujeito confirmou que só o efeito de grupo é significativo,  $F(2,83)=8.81$ ,  $p=.0003$ ; para o efeito da atenção e para a interacção atenção x grupo,  $F < 1$ .\*

Na condição de atenção concentrada, cada sujeito ora atendia ao ouvido esquerdo, ora ao direito. Havia pois a possibilidade de se encontrarem diferenças de ouvido quanto à percentagem de fusões. Uma análise de variância análoga à anterior, considerando os factores grupo e ouvido, mostrou não haver efeito de ouvido ( $F < 1$ ), nem interacção entre grupo e ouvido ( $F(2,81)=2.41$ ,  $p=.0955$ ). Mais uma vez, o efeito do grupo emergiu

\* Para pormenores, cf. Anexo FF.1.



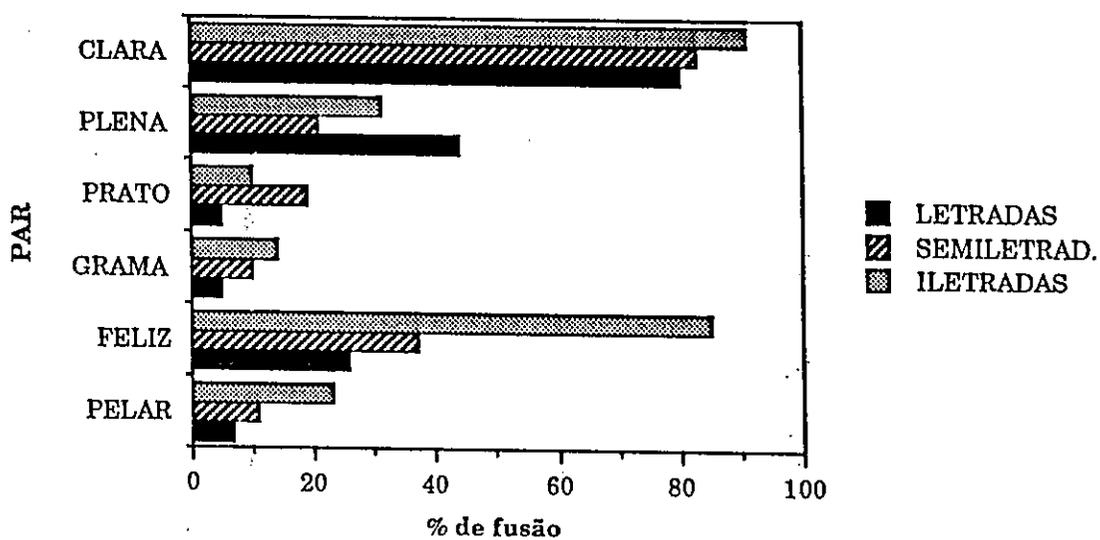
**FIGURA 2:** Número de sujeitos (total por grupo, n=28) com uma incidência de fusões compreendida entre 0-10%, ou 11-20%, e assim sucessivamente. (LET., letradas, SLET., semiletradas, e ILET., iletradas)

significativamente,  $F_{(2,81)}=9.01$ ,  $p=.0002^*$ . Independentemente de se ter instruído os sujeitos a focalizarem a sua atenção, e desta se dirigir à esquerda ou à direita, observam-se mais fusões no grupo das iletradas (cerca de 21%), e menos no grupo das letradas (cerca de 14%). Como os dados sugerem, testes Tukey para comparação de médias confirmam que é significativa ( $\alpha=.05$ ) a diferença entre iletradas e letradas por um lado, e iletradas e semiletradas por outro. Os resultados das semiletradas (cerca de 15% de fusões) não se diferenciam das das letradas de modo significativo.

Será esta diferença entre iletradas, e letradas em sentido lato devida à existência de subgrupos específicos como "high" e "low fusers" descritos por Day (1968)? Ou seja, distribuir-se-ão os sujeitos de modo dicotômico quanto à frequência de fusões? A Figura 2 mostra a distribuição de sujeitos por frequência de fusão. O que ressalta da observação daquele gráfico é a diferença quanto à distribuição das letradas e iletradas. Enquanto 75% das letradas se situam abaixo de 30% de fusões, com as iletradas aconteceu o contrário: 75% delas têm mais de 30% de fusões. Além disto, é apenas entre as iletradas que aparece um subgrupo de 7 sujeitos (25%) com uma incidência relativamente mais pronunciada de fusões, entre 51-69%. Como já vem sendo familiar, a distribuição das semiletradas embora aproximando-se mais da das letradas que da das iletradas, apresenta-se como que de transição.

Os nossos resultados não mostram pois uma distribuição dicotômica de "high" vs "low fusers". Revelam antes uma diferença entre letradas e iletradas, as primeiras com um subgrupo de sujeitos que ouviram poucas fusões ("low fusers"), e as segundas com um subgrupo de "high fusers", inexistente nas letradas.

\*Para pormenores, cf. Anexo FF.2.



**FIGURA 3:** Percentagem média de fusões observadas em cada grupo, por par de estímulos.

Tendo apreciado os resultados na sua globalidade, passemos agora a uma análise por *item*, com vista a esclarecer as seguintes duas questões principais: há pares mais fundíveis que outros? A preponderância de fusões no seio das iletradas observa-se em todos os casos?

Como se pode observar na Figura 3, em que estão representadas as percentagens de fusões por *item* e grupo, há três pares de estímulos claramente mais fundíveis que outros: para "cara/lara" há cerca de 85% de fusões, para "fiz/lis" cerca de 49% e para "pena/lena" cerca de 32%. O restante *item* com uma líquida, "lar/par", obteve menos fusões, cerca de 14%; nos dois pares com a vibrante, a percentagem de fusões foi ainda menor, cerca de 10 e 11% para "gama/rama" e "pato/rato", respectivamente. Assim, os pares mais fundíveis parecem ser os que envolvem a líquida como elemento do encontro consonantal; a fusão com a vibrante ocorre menos frequentemente.

Para apreciar melhor a variabilidade intra e intergrupo, e avaliar a possível influência ortográfica, calculou-se uma análise de variância ao número de fusões com os seguintes factores: grupo (intersujeito) e tipo de par (intrasujeito). Consideraram-se três tipos de par: com líquida ("cara/lara" e "pena/lena"), com batimento ("gama/rama" e "pato/rato") e de ortografia não-congruente ("fiz/lis" e "par/lar"). Esta análise revelou uma interacção altamente significativa entre grupo e tipo de par,  $F_{(4,160)}=11.43$ ,  $p<.0001$ \*. Para esclarecimento desta interacção, analisou-se separadamente o efeito do grupo em cada tipo de par.

Considerando os pares com líquida, "cara/lara" e "pena/lena", o efeito de grupo não chega a atingir a significância,  $F_{(2,81)}=1.82$ ,  $p=.16$ . Observando o padrão de respostas a cada um destes pares, verifica-se que ele

\* Para pormenores cf. Anexo FF.3

varia conforme o grupo. Enquanto em "cara/lara" há cerca de 10% mais fusões nas iletradas do que nas letradas (cf. Figura 3), para "pena/lena" acontece praticamente o contrário. Como este é o único caso em que se observa uma maior incidência de fusões nas letradas do que nos outros dois grupos, detenhamo-nos sobre ele mais pormenorizadamente. As letradas responderam 44% de fusões, contra 31% nas iletradas e 21% nas semiletradas. Uma análise de variância ao número destas fusões, com grupo como variável intersujeito, revelou que o efeito do grupo é significativo,  $F_{(2,81)}=3.301$ ,  $p=.0419$ ; só a diferença entre letradas e semiletradas é significativa (comparações por testes Tukey com  $\alpha=.05$ ). O facto de as semiletradas responderem menos vezes "plena" poderá estar relacionado com o seu padrão de resposta a estes estímulos em apresentação monaural. De facto, 9 semiletradas identificaram a palavra "plena" como "lena" ou "helena", o que não aconteceu com nenhuma letrada e só com 2 iletradas. Além disto, o estímulo "lena" foi percebido como "plena" por cinco letradas e quatro iletradas, o mesmo acontecendo com apenas uma semiletrada. Devido a estes resultados, só nos parece justificado fazer comparações entre letradas e iletradas, cujo padrão de resposta em monaural foi semelhante. Apesar de a diferença observada não ser significativa, o facto de ela ser na direcção oposta a todas as outras leva-nos a não pôr de parte a possibilidade de um eventual efeito da familiaridade de palavras: talvez as letradas tenham mais tendência do que as iletradas a fundir "pena" com "lena" porque a palavra "plena" lhes é mais familiar.

Considerando os pares com vibrante, o efeito de grupo foi significativo,  $F_{(2,81)}=3.2$ ,  $p=.0428$ . A percentagem de fusões das iletradas foi de 12%, e não diferia significativamente dos outros grupos. Só a diferença entre letradas (com cerca de 5% de fusões) e as semiletradas (com cerca de 15%) é significativa (testes Tukey  $\alpha=.05$ ). Este resultado poderá certamente atribuir-se às respostas das semiletradas perante o par "pato/rato". Este é o

único caso em que as semiletradas dão mais fusões do que as letradas (19 vs 10%, respectivamente). Talvez os resultados na identificação em monaural possam esclarecer este padrão: é que "rato" foi percebido por três semiletradas (contra nenhuma letrada e uma iletrada) como "prato". Portanto, a única diferença significativa encontrada neste caso, entre letradas e semiletradas, poderá dever-se a idiosincrasias destes últimos sujeitos.

Finalmente, considerando os dois pares restantes, de ortografia não congruente com a fusão, verifica-se que o efeito de grupo é significativo,  $F_{(2,81)}=23.59, p <.0001$ . Testes Tukey ( $\alpha=.05$ ) mostraram que as iletradas dão mais fusões que as letradas e que as semiletradas, respectivamente 53%, 24% e 16%. Entre estes dois grupos, não há diferença significativa.

Debrucemo-nos primeiro sobre as respostas aos estímulos "par/lar". A diferença entre letradas e iletradas não é interpretável à luz das respostas àqueles estímulos em identificação monaural; com efeito, os erros foram praticamente inexistentes: só uma semiletrada ouviu [plar] em vez de "par", e uma iletrada ouviu [plar] em vez de "lar". Não parecem pois existir viés grupais que à partida privilegiem este ou aquele percepto. Convém referir que algumas das respostas de fusão, [pl...], faziam uma transformação na consoante final, sendo o percepto [pla] ou [plar]; isto aconteceu com 3 iletradas e uma letrada. Outra observação digna de nota é que frequentemente as letradas e semiletradas identificavam os dois estímulos, "par de um lado, e lar do outro", diziam\*; isto ocorreu com doze letradas e sete semiletradas. Estas respostas duplas foram bem menos frequentes no grupo das iletradas, onde se observaram em apenas dois casos.

\* Na contagem de respostas de um ou outro tipo, se porventura o sujeito dava resposta dupla optámos por considerar apenas a primeira resposta.

A elevada incidência de respostas "feliz" ao par "fiz/lis" por parte das iletradas é um tanto surpreendente. Na identificação em monaural não apareceu nenhuma dificuldade especial por parte destes sujeitos em ouvir os estímulos componentes. Como se referiu atrás, nem sempre "fiz" foi correctamente percebido: confusões com [si], [i] e outras foram observadas em onze letradas, cinco semiletradas e quatro iletradas. Também na identificação de "lis" se deram algumas confusões com [fli]; estas todavia ocorreram tão frequentemente nas letradas quanto nas iletradas (dois casos cada). Portanto, à partida não se observa nenhuma preferência por parte das iletradas em responder [fli] aos estímulos "fiz" e "lis". Como seria de esperar, dada a má qualidade sonora ou a dificuldade em perceber a fricativa inicial em "fiz", em todos os sujeitos e grupos as respostas [li] são muito mais frequentes; de facto, as respostas [fi] constituem apenas 4% do total das respostas das letradas, 7% das semiletradas e 0% das iletradas.

As respostas duplas ocorreram mais raramente nas iletradas: houve apenas uma que percebeu uma vez [fi] de um lado e [fli] do outro. Respostas duplas [li] e [fli] apareceram em quatro letradas e duas semiletradas.

#### 4. Discussão e conclusão

O presente estudo, usando uma tarefa de identificação oral de pares fundíveis apresentados dicoticamente, revelou diferenças importantes no grau de fusibilidade de cada par de estímulos e, para alguns deles, diferenças intergrupais quanto à incidência de fusões fonológicas. Não se encontraram efeitos da estratégia atencional. Apreciemos mais detalhadamente estes resultados, situando-os no contexto das outras investigações sobre fusão fonológica.

A predição segundo a qual a percentagem de fusões seria reduzida pela orientação da atenção a só um ouvido, baseada em achados prévios de Sexton e Geffen (1981), não se verificou. A questão que se coloca é então se o efeito da estratégia atencional ocorre apenas em tarefas semelhantes à usada por aqueles autores, ou se o facto de não o termos observado se poderá atribuir a algum aspecto metodológico da nossa experiência. Com efeito, Sexton e Geffen usaram uma tarefa de detecção do alvo: o sujeito respondia sob pressão temporal, não reproduzindo o(s) estímulo(s), mas pressionando uma tecla quando achava que tinha ouvido uma palavra previamente determinada, cuja ocorrência ele tinha de detectar. Numa tarefa como esta, os estímulos aparecem em sucessão rápida (naquele caso, um em cada 750 milisegundos), enquanto na nossa tarefa de identificação a frequência era de um par de estímulos cada 5 segundos.

O facto de os sujeitos terem de responder rapidamente pode levar a que o processamento da informação fonológica se fique pelas exigências mais pregnantes da tarefa ("ouvir só o ouvido direito e bloquear o esquerdo"), enquanto numa situação menos constrangedora em termos temporais a instrução de prestar atenção só a um lado não será tão efectiva. A ideia é que o percêpto pode variar conforme o tempo - e as operações - de processamento (cf. *e.g.*, Treisman, no prelo, e Ward, 1983), e que o bloqueio da informação do canal rejeitado pode ser mais eficaz num curto intervalo de tempo. Os nossos dados não permitem todavia confirmar ou infirmar esta suposição, na medida em que não medimos a latência de resposta.

Outro factor que poderá estar na base do facto de não termos observado efeito da atenção são as próprias características dos estímulos usados, nomeadamente o facto de serem parcialmente fusionados. É sabido que se os estímulos apresentados a cada ouvido tiverem porções acústicas semelhantes, conforme estas sejam maiores ou menores, eles podem ser

fundidos numa única imagem acústica (cf. Repp, 1980; Wexler e Halwes, 1982), tornando-se irrelevantes os esforços atencionais para separá-la em duas mensagens. Quer ao preparar o teste, quer ao longo da experiência piloto com os estudantes, não parecia que os estímulos por nós construídos se fundissem num percepto único, imune ao controlo voluntário da atenção: através da concentração ora num ouvido, ora no outro, distinguíam-se perceptos diferentes. Pode todavia ter acontecido que o grau de separação acústica das mensagens não fosse suficiente para que os sujeitos desta experiência conseguissem voluntariamente distinguir o estímulo do ouvido esquerdo do do direito. De facto, muitos deles relataram "ser a mesma coisa quando prestavam atenção a um só ouvido ou aos dois em conjunto". O facto de isto não acontecer com todos os sujeitos mostra que não é só uma característica da estimulação que está em jogo. Alguns sujeitos não só deram respostas duplas, como identificavam a sua localização.

Convém a este propósito referir que, na ausência de qualquer instrução explícita sobre a atenção, é usual haver respostas duplas só quando os estímulos dicóticos diferem em entoação, timbre vocálico, duração ou em vários fonemas constituintes (cf. Day, 1968; 1970-a; Cutting e Day, 1975). Como nenhuma destas variações estava presente nos nossos estímulos, o facto de os sujeitos terem produzido respostas duplas mesmo em atenção dividida sugere que o grau de identidade acústica não era tão forte que impedisse algum controlo atencional.

Enfim, somos obrigados a aceitar que os estímulos por nós usados não terão sido os melhores: o seu grau de semelhança acústica leva a que o percepto correspondente seja extremamente vulnerável a diferenças individuais na proficiência da atenção. Se grande parte dos sujeitos não conseguia, mesmo fenomenologicamente, prestar atenção a só um ouvido, está à partida gorada a possibilidade de averiguarmos o efeito da atenção nas

fusões fonológicas, e do seu papel nas eventuais diferenças entre letradas e iletradas.

Passemos agora em revista outras observações comuns aos vários grupos. Ainda relativamente aos nossos estímulos, verificou-se consistentemente uma baixa inteligibilidade das fricativas e da vibrante, mesmo em audição monaural. Este facto alerta para a necessidade de um cuidado adicional na preparação destes estímulos; o processo de digitalização permite um grande controlo nomeadamente quanto ao alinhamento temporal dos estímulos dicóticos. Pode porém afectar negativamente a qualidade dos estímulos, como aconteceu neste caso.

Tal como noutras investigações, obteve-se uma percentagem de fusão muito variável consoante o par de estímulos. De uma maneira geral, as fusões foram mais frequentes com a líquida do que com /r/. Trata-se de um resultado análogo ao encontrado com fonemas ingleses, em que a grande maioria de fusões consistem em *clusters* com /l/, sendo menos frequentes as fusões com a líquida inglesa /r/. Também em português se observam menos fusões em que o segundo elemento do encontro consonantal é /ʀ/. Talvez à partida esta situação fosse previsível, pois aquele fone não existe em posição inicial. Na nossa língua, para fundir, e.g., "pato" com "rato" é preciso transformar o [R] em [ʀ]. O facto de isto acontecer é consistente com as observações de Cutting e Day (1975, Experiência 4), segundo as quais a fusão ocorre com pares dicóticos fonemicamente dissemelhantes. Os estímulos apenas têm de ser compatíveis a nível fonológico, no sentido de que os fonemas iniciais quando combinados devem produzir um *cluster* admissível; os nossos resultados indicam que podem mesmo ser efectuadas transformações nos fonemas iniciais obedecendo a regras fonotácticas, de modo a produzir um *cluster* admissível.

Outras três observações são consistentes com o descrito na literatura. Nos pares contrastando oclusiva com líquida, era mais comum a resposta com a palavra iniciada por oclusiva do que com a líquida, situação análoga à descrita por Cutting (1975). Não observámos um efeito análogo ao de substituição por /l/ (cf. pp. 91-93), que parece acontecer mais com estímulos sintéticos (cf. Cutting e Day, 1975). Tal como Sexton e Geffen (1981), não observámos efeitos de ouvido. O facto de não haver mais fusões no ouvido esquerdo é consistente com a ideia de que a fusão fonológica não é um mero "erro" de identificação dos estímulos.

De uma maneira geral, as iletradas ouviram consistentemente mais fusões do que as letradas e semiletradas. A única excepção ocorreu perante o par "pena + lena", que cerca de 50% das vezes foi percebido como "plena" pelas letradas, contra cerca de 40-20% pelos outros dois grupos. A questão está obviamente em compreender o porquê e o como destas diferenças. Um primeiro aspecto é se as diferenças encontradas são gerais ou específicas, ou seja, se os sujeitos iletrados têm mais tendência em fundir independentemente de características específicas dos estímulos, ou se pelo contrário aquelas diferenças se observam apenas em determinadas circunstâncias. Tomemos dois factores que poderiam mediar as diferenças: a familiaridade das palavras e o facto da ortografia do *cluster* ser ou não congruente com o das palavras componentes.

A fusão parece ocorrer mais frequentemente se as palavras resultantes são mais familiares (Cutting, 1975; Cutting e Day, 1975). Para avaliarmos da familiaridade das palavras usadas neste estudo temos que recorrer à nossa experiência de falantes e ao conhecimento algo intuitivo do que será o vocabulário mais comum nos diferentes grupos considerados. Admitamos que os vocábulos "clara, feliz, grama, prato" são tão familiares para letrados como para iletrados e semiletrados; admitamos ainda que

"plena" poderá ser mais conhecido por letrados. Sobre "pelar" parece-nos algo arbitrário estar a supor diferenças grupais; teríamos tendência a crer que este termo é mais usado pelos habitantes do meio agrícola (iletradas e algumas semiletradas), onde é corrente ouvir a frase "vou pelar umas batatas", mas trata-se apenas de uma impressão que pode estar enviesada pelo facto de se conhecer pessoalmente alguns desses sujeitos. No conjunto das nossas observações, parece bastante plausível que a variação intergrupar quanto à percentagem de fusão "plena" se deixe explicar pelo diferente grau de familiaridade daquela palavra nos grupos em questão, nomeadamente nas letradas e iletradas. Pensamos que este é o único aspecto dos nossos resultados interpretável à luz de um efeito da familiaridade.

Voltemo-nos agora para a ortografia. Os resultados obtidos vão de encontro à nossa predição, segundo a qual sujeitos que não conheçam a ortografia fundiriam mais quando a fusão resulta numa palavra com grafia distinta da justaposição da grafia das componentes. De facto, a única situação em que se observou diferença significativa entre letradas em sentido lato e iletradas foi justamente para as respostas "feliz" e "pelar" perante os pares "fiz/lis" e "par/lar".

Uma possível objecção à interpretação das diferenças intergrupais em termos puramente ortográficos situa-se a nível das respostas "feliz". Dada a fraca qualidade da fricativa, as respostas no par "fiz + lis" concentravam-se nas categorias "feliz" e "lis". Talvez o último vocábulo fosse menos familiar para as iletradas. Apesar de, em rigor, não podermos eliminar esta possibilidade, não parece claro por que razão a palavra "lis" já seria mais conhecida pelas semiletradas. Assim, as diferenças observadas entre letradas, em sentido lato, e iletradas são consistentes com a ideia de que a fusão pode ser afectada pela representação ortográfica.



Resta debater se a variabilidade intergrupar não será ainda determinada por algum outro factor importante. De facto, apesar das únicas diferenças significativas entre letradas e iletradas serem interpretáveis à luz de um efeito ortográfico, o facto de consistentemente se ter observado uma maior percentagem de fusões nas iletradas levanta a suspeita de que poderá haver em jogo outros factores.

Uma primeira observação é que respostas duplas foram menos frequentes nas iletradas\* do que nos outros sujeitos. Isto pode ser uma consequência trivial da compreensão da tarefa; apesar de as instruções referirem que se houvesse dois perceptos diferentes ambos deviam ser relatados, as iletradas poderiam ter negligenciado aquele pedido. Todavia, não sabemos até que ponto a raridade das respostas duplas traduzirá uma tendência real em integrar a informação num percepto único. Este é seguramente um aspecto a ter em conta no planeamento de investigações futuras.

Outro aspecto a merecer consideração é o que segue. Como em geral se observa que sujeitos iletrados dão mais erros em situações de reconhecimento auditivo de palavras, poderia acontecer que as respostas que descrevemos como "fusões" sejam, pelo menos nalguns casos, não verdadeiras combinações de fonemas de mensagens diferentes, mas "confusões": perante um percepto mais ou menos ambíguo, "parecido" com "fiz" ou "lis", o sujeito responderia "feliz", uma palavra mais "agradável", talvez até "a mesma que se ouviu antes". De facto, algumas iletradas diziam "é outra vez ...", mostrando assim que estavam conscientes da reprodutividade da sua resposta. Poderia acontecer um fenómeno de *priming* por repetição: tendo um determinado par de estímulos sido

---

\* Não se notaram diferenças intergrupais quanto às respostas duplas conforme o par de estímulos.

entendido uma, duas vezes de certa maneira, em qualquer nova apresentação o percepto seria idêntico. Este possível efeito de *priming* por repetição poderia ser mais importante nos iletrados, e em certas palavras com valor emocional, como "feliz".

O planeamento desta experiência não permite resolver este problema. A precaução inicial de familiarizar os sujeitos com as várias respostas possíveis, também sugerida por Poltrock e Hunt (1977), nada garante, obviamente, quanto ao efeito posterior da repetição dos estímulos. Para averiguar se os iletrados são menos discriminativos, e as respostas de "fusão" são apenas respostas parecidas com o que se ouviu, seria necessário usar ensaios-rasteira, em que fossem apresentados estímulos distractivos, foneticamente semelhantes, ou até a própria palavra componente nos dois ouvidos ("lis + lis", "fiz + fiz"). Poder-se-ia então comparar a discriminabilidade entre a fusão propriamente dita (resposta "feliz" perante "fiz + lis") e outros erros (*e.g.*, resposta "feliz" perante "fiz + fiz" ou, por exemplo, "voz + loz"), recorrendo eventualmente a uma análise em termos de detecção do sinal.

Mesmo que fenómenos como os descritos (confusão, maior susceptibilidade das iletradas a um *priming* por repetição) sejam verídicos, não parece crível que lhes pudéssemos atribuir por completo os efeitos observados. Porque seriam as confusões ou erros justamente as palavras que englobam a informação fonémica recebida dicoticamente? Porque se estabilizaria a resposta nas palavras tipo fusão, e não noutras - no nosso caso, por exemplo, as palavras "cara", "par", "lar", "pena", "pato" e "rato" são certamente pelo menos tão familiares como algumas das fusões correspondentes. Sem dúvida é mais justificável, no contexto do conhecimento actual sobre estes fenómenos, supor que "the linguistic

processor attempts to combine all available linguistic features when forming a linguistic unit from confusing stimuli" (Poltrock e Hunt, 1977; p. 73).

Em suma, os resultados do presente estudo revelaram que sujeitos iletrados dão significativamente mais respostas tipo fusão do que sujeitos letrados, quando a representação ortográfica do percepto não é idêntica à de um encontro consonantal. Apesar de algumas questões levantadas por estes resultados só poderem ser esclarecidas em futuros trabalhos experimentais, propomos para já que, além dos factores linguísticos referidos por outros autores, a representação ortográfica das palavras pode afectar a integração de informação fonológica exercendo um efeito inibidor quando ela divide "cluster" em sílabas diferentes.

A existência de um efeito de ortografia, e a influência de factores atencionais na incidência de fusões fonológicas observada por Sexton e Geffen (1981) são argumentos em favor da separação entre a fusão de traços fonéticos (*blendings*) e a fusão fonológica em níveis diferentes do processamento da fala, como tinha sido tentativamente sugerido por Cutting (1975).

## CONCLUSÃO GERAL

Propusémo-nos elucidar alguns aspectos da relação entre fala e escrita do ponto de vista dos processos psicológicos envolvidos. O objecto deste trabalho tomou uma das vertentes deste problema, menos explorada, a da possível influência do conhecimento da escrita no processamento da linguagem falada, nomeadamente na percepção da fala. Foi levado a cabo um conjunto de experiências visando diferentes níveis, pelo menos conceptuais, do processamento da fala, com sujeitos que dominam em maior ou menor grau a escrita, e com sujeitos que por razões sócio-económicas ou de idade apenas dominam a linguagem falada. Os resultados desse trabalho experimental foram já discutidos com pormenor, pelo que agora apenas salientaremos a sua relevância no contexto do estado do conhecimento sobre estas matérias.

Quanto ao suporte biológico da linguagem, as experiências I e II mostraram que a identificação de palavras se faz em geral relativamente melhor no ouvido direito do que no esquerdo, independentemente de se conhecer ou não a escrita; não se observou nenhuma relação específica entre alfabetização e magnitude da vantagem do ouvido direito. Este último resultado vem, em nossa opinião, esclarecer uma controvérsia existente na literatura sobre a influência da alfabetização nos efeitos de lateralidade e, supostamente, na organização cerebral para a linguagem. Na medida em que os efeitos de lateralidade auditiva permitem inferir sobre a actividade hemisférica, os resultados das experiências I e II constituem evidência convergente para a asserção de que o grau de dominância do hemisfério esquerdo para tratar a linguagem falada não é afectado pela alfabetização *per se*. No contexto da literatura sobre este problema, é importante salientar que os resultados contraditórios obtidos com adultos iletrados (uma inversão da vantagem de ouvido para palavras foneticamente semelhantes, por

Damáσιο *et al.*, 1979, e uma assimetria mais pronunciada para palavras foneticamente dissemelhantes, por Tzavaras *et al.*, 1981) não foram replicados, nas condições metodológicas mais estritas em que se desenrolaram as nossas experiências.

Para perceber a fala, é provavelmente fundamental identificar as categorias de sons linguisticamente relevantes, e discriminá-las claramente de categorias próximas do ponto de vista acústico, mas distintivas do ponto de vista semântico. Esta questão tem sido investigada experimentalmente através da identificação, acompanhada ou não por discriminação, de *continua* acústicos ao longo de uma dimensão fonética e fonemicamente relevante. Como se têm observado efeitos da experiência linguística nestas tarefas, é importante esclarecer se os efeitos da alfabetização na percepção da fala ocorrem já a este nível relativamente precoce da categorização de estímulos acusticamente semelhantes. Os resultados da Experiência VI (nomeação de *continua* de vozeamento) indicam que tanto sujeitos letrados como iletrados dividem a variação acústica gradual em categorias perceptivamente distintas, observando-se descontinuidades a nível da nomeação. Sujeitos com um elevado nível de escolarização parecem estabelecer estas descontinuidades aproximadamente nos mesmos locais dos *continua*, independentemente da idade (jovens universitários ou adultos de cerca de 40 anos); porém, os sujeitos menos escolarizados fazem-no de modo mais variável. Enfim, os resultados da Experiência VI sugerem a alfabetização não afecta o modo categórico como é feita a nomeação de *continua* acústicos, pelo menos nos que usam índices temporais de pré-vozeamento; possivelmente a experiência linguística de sujeitos com elevados níveis de escolarização contribui para uma maior uniformidade inter-individual quanto às estratégias de nomeação. Todavia, um esclarecimento cabal destes aspectos exige novos trabalhos empíricos.

Um dos processos que parece estar envolvido na identificação dos sons da fala é a análise em termos de traços fonéticos. Os resultados das Experiências III e IV permitem estabelecer que um fenómeno indicativo da extracção desses traços, a fusão de traços fonéticos ou *blendings*, ocorre também em adultos iletrados e pouco escolarizados. O facto de adultos iletrados responderem por vezes com *blendings* é duplamente importante; por um lado, mostra a generalidade do fenómeno, contribuindo assim para estabelecer a "realidade psicológica" da noção de traço fonético; por outro, mostra que a fusão de traços fonéticos decorre de um mecanismo relativamente fixo, não susceptível às influências de variáveis ambientais.

Em determinadas condições é possível observar erros de percepção que consistem na combinação de fonemas de palavras apresentadas aos dois ouvidos. Estes erros de fusão fonológica observam-se quer em indivíduos letrados, como iletrados e pré-letrados (Experiências III, IV, VII), o que sugere que a informação ao nível de fonema é relevante perceptivamente, sem estar dependente da alfabetização. Este resultado é particularmente importante, pois confirma experimentalmente a distinção entre representação consciente do percepto e atributos do estímulo usados nos processos perceptivos: apesar de os iletrados não terem, em geral, acesso a uma representação consciente da fala como uma sequência de fonemas, o fonema parece desempenhar um papel relevante na identificação da fala em geral, e das palavras em particular.

Identificámos assim alguns fenómenos, cuja mera ocorrência é independente da alfabetização: a integração de traços fonéticos ou *blendings*, a nomeação categórica de *continua* acústicos numa dimensão linguisticamente relevante, e a integração de fonemas. Na medida em que estes fenómenos traduzem processos subjacentes de, respectivamente, análise fonética, categorização "abrupta" de sons da fala, e extracção de

fonemas, os nossos resultados sugerem que estes processos ocorrem a um nível relativamente precoce, sendo em certa medida, automáticos e relativamente independentes de controlo atencional. Note-se que no caso da fusão fonológica tratamos agora unicamente o facto de ela ocorrer, e não aspectos mais específicos como a sua incidência consoante tipos de palavras ou grupos de sujeitos.

Ao longo das experiências efectuadas, além dos resultados comuns aos vários grupos, discutidos anteriormente, encontrámos também de modo consistente uma dissociação entre letrados e iletrados para os aspectos que passamos a referir. Em geral, sujeitos iletrados dão mais erros de identificação de palavras orais apresentadas através de auscultadores. Talvez alguns desses erros possam ser explicados por factores mais ou menos triviais, relacionados, entre outros, com a familiaridade da audição por auscultadores, concentração em atributos funcionais dos estímulos (responder "tacho" a "cacho", por exemplo). Valerá a pena, sem dúvida, observar o comportamento destes sujeitos em tarefas simples de repetição de palavras e não-palavras, pronunciadas *in loco* pelo experimentador.

Além do nível de desempenho, observou-se consistentemente uma variação no padrão de respostas em função da alfabetização. Mesmo a um nível de *performance* semelhante, quer na audição de palavras apresentadas dicoticamente, quer na audição de palavras mascaradas por ruído (Experiência V, Ordem I), sujeitos infantis (Experiência IV) ou adultos (Experiências III e IV) com conhecimento da escrita exibiam um padrão de resposta caracterizado por uma incidência relativamente maior de erros segmentais, enquanto crianças pré-letradas (Experiência IV) e adultos iletrados (Experiência III) se caracterizavam por uma incidência relativamente maior de erros globais, em toda a palavra ou que não envolvessem partição silábica.

A interpretação proposta para estes resultados é que eles são o reflexo de uma estratégia atencional dirigida aos segmentos da fala. Esta interpretação é sem dúvida consistente com o que se conhece a propósito da relação entre alfabetização e consciência da estrutura fonética da fala. Um primeiro passo para a submetermos a um teste mais directo, a experiência piloto que apresentámos sucintamente (cf. pp. 244 - 246), leva a crer que a diferença entre sujeitos letrados, e iletrados ou pré-letrados é de facto mediada por factores estratégicos, e não por um qualquer processo automático de reconhecimento de palavras. A propósito da ideia segundo a qual estes factores estratégicos se prendem com a atenção aos segmentos da fala vale a pena referir dados recentes de Nittrouer e Studdert-Kennedy (1986).

Trata-se de uma investigação sobre os efeitos da coarticulação, em que crianças entre 3 a 7 anos e adultos identificavam dois *continua* entre /s/ e /ʃ/, um com a vogal /i/ e outro com a vogal /u/. Os resultados mostraram uma dissociação entre as crianças pré-letradas, por um lado, e as crianças mais crescidas e os adultos, por outro. Os efeitos de coarticulação foram mais marcados nas crianças mais jovens, o que sugere que a organização perceptiva da fala se vai tornando mais segmental e menos global com o desenvolvimento (*ibidem*). O que os nossos resultados sugerem é que a aquisição da leitura pode ser uma experiência crítica para o desenvolvimento desta "organização segmental", como a designam Nittrouer e Studdert-Kennedy.

Um factor relacionado com a atenção segmental, mas não necessariamente idêntico, e que poderia mediar os efeitos da alfabetização, é a representação ortográfica. Examinámos até que ponto ela poderia afectar a fusão fonológica (Experiência VII). A observação de que a fusão fonológica é menos frequente nos sujeitos letrados do que nos iletrados, quando a

representação ortográfica da palavra correspondente envolver uma sílaba adicional ([fli] vs "feliz") é importante por duas razões. Em termos da relação entre alfabetização e percepção da fala, ela sugere que a representação ortográfica pode, em certas situações, mediar as diferenças encontradas entre letrados e iletrados. Em termos dos processos envolvidos na fusão fonológica, o facto de esta ser afectada pela representação ortográfica permite distingui-la da fusão de traços fonéticos, no sentido de que é mais sensível às representações cognitivas, e portanto mais central.

Em suma, apesar de múltiplas questões a responder (vejam-se as discussões das experiências III - VII), este trabalho aponta para um modelo de percepção da fala que incluiria processos pré-atencionais e automáticos, independentes da alfabetização, e processos estratégicos, opcionais, sensíveis ao efeito daquela variável. Entre os primeiros, contar-se-iam os seguintes: extracção da informação relativa a traços fonéticos, a sua integração em representações intermediárias de tipo fonémico, a organização destas representações em categorias bem definidas e contrastantes que se sobreporiam à variação acústica gradual. Os processos opcionais contribuiriam para uma "interpretação" mais fina do produto destes primeiros processos, e seriam afectados em determinada medida pela disponibilidade e acessibilidade de representações conscientes, como a da estrutura fonética da fala. Neste contexto é importante esclarecer até que ponto a representação ortográfica, nos indivíduos que dela dispõem, cabe dentro dos processos que denominamos pré-atencionais ou é opcional.

Uma possibilidade interessante é que os processos pré-atencionais estejam dependentes de um sistema biológico específico com características modulares (cf. Fodor, 1983; Liberman e Mattingly, 1985). A variabilidade do funcionamento deste sistema não será provavelmente uma fonte importante de diferenças individuais a nível de percepção de fala, nomeadamente

quanto à exactidão e ao padrão de resposta ao identificar palavras. Pelo contrário, a variabilidade destes comportamentos, quer individual, quer grupal, seria largamente dependente dos processos não-modulares opcionais.

## REFERÊNCIAS

- Abramson, A. S. e Lisker, L. (1968). Voice timing: Cross-language experiments in identification and discrimination. *Journal of the Acoustical Society of America* 44, 377 (A).
- Abramson, A. S. e Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. In *Proceedings of the 6<sup>th</sup> International Congress of Phonetic Sciences* (Prague, 1967). Prague: Academia.
- Alegria, J. e Morais, J. (1979). Le développement de l'habilité d'analyse consciente de la parole et l'apprentissage de la lecture. *Archives de Psychologie* 183, 251-270.
- Alegria, J., Pignot, E. e Morais, J. (1982). Phonetic analysis of speech and memory codes in beginning readers. *Memory and Cognition* 10, 451-456.
- Allport, D. A. (1968). Phenomenal spontaneity and the perceptual moment hypothesis. *British Journal of Experimental Psychology* 59, 395-406.
- Anisfeld, M. (1984). *Language Development from Birth to Three*. New Jersey: Lawrence Erlbaum Associates.
- Ashton, R. e Beasley, M. (1982). Cerebral laterality in deaf and hearing children. *Developmental Psychology* 18, 294-300.
- Aslin, R. N., Pisoni, D. B., Hennessy, B. L. e Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development* 52, 1135-1145.
- Bakker, D. J., Van Der Vlugt H., e Claushuis, M. (1978). The reliability of dichotic ear asymmetry in normal children. *Neuropsychologia* 16, 753-757.
- Barbosa, J. M. (1983). *Études de Phonologie Portugaise*. Évora: Universidade de Évora.
- Bastian, J., Eimas, P. e Liberman, A. M. (1961) Identification and discrimination of a phonemic contrast induced by a silent interval. *Journal of the Acoustical Society of America* 33, 842. (Sumário)
- Beaumont, J. G. (1983) *Introduction to Neuropsychology*. New York: Guilford Press.
- Bentin, S. e Mann, V. (1983). Selective effects of masking on speech and nonspeech in the duplex perception paradigm. *Haskins Laboratories Status Report on Speech Research SR-76*, 65-85.

- Bertelson, P. (1972). Listening from left to right vs. right to left. *Perception* 1, 161-165.
- Bertelson, P. (1986). The onset of literacy: liminal remarks. *Cognition* 24, 1-30.
- Bertelson, P. (Ed.) (1987). *The onset of literacy: cognitive processes in reading acquisition*. Cambridge, Mass: MIT Press.
- Bertelson, P. e De Gelder, B. (no prelo). Learning about reading from illiterates. In A. M. Galaburda (Ed.) *From Neurons to Reading*. Cambridge, Mass.: Mit Press.
- Bertelson, P., Morais, J., Alegria, J. e Content, A. (1985). Phonetic analysis capacity and learning to read. *Nature* 313, 73-74.
- Bertelson, P., Morais, J., Cary, L. e Alegria, J. (1986). Interpreting data from illiterates: Reply to Koopmans. *Cognition* 24, 113-115.
- Best, C. T., Hoffman, H. e Glanville, B. B. (1982) Development of infant ear asymmetries for speech and music. *Perception & Psychophysics* 31, 75-85.
- Best, C. T., Morrongiello, B. e Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics* 29, 191-211.
- Bever, T. G. (1971). The nature of cerebral dominance in speech behavior of the child and adult. In R. Huxley e E. Ingram (Eds). *Language Acquisition: Models and Methods*. London: Academic Press.
- Blumstein, S. E. (1974). The use and theoretical implications of the dichotic technique for investigating distinctive features. *Brain and Language* 1, 337-350.
- Blumstein, S. E., Goodglass, H. and Tartter, V. (1975). The reliability of ear advantage in dichotic listening. *Brain and Language* 2, 226-236.
- Blumstein, S. E., Isaacs, E. e Mertus, J. (1982). The role of the gross spectral shape as a perceptual cue to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America* 72, 43-50.
- Blumstein, S. E. e Stevens, K. N. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *Journal of the Acoustical Society of America* 66, 1001-1017.
- Blumstein, S. E. e Stevens, K. N. (1980). Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America* 67, 648-662.
- Blumstein, S. E. e Stevens, K. N. (1981). Phonetic features and acoustic invariance in speech. *Cognition* 10, 25-32.

- Borden, G. J. e Harris, K. S. (1982). *Speech Science Primer: Physiology, Acoustics and Perception of Speech*. Baltimore: Williams and Wilkins.
- Borowy, T. e Goebel, R. (1976). Cerebral lateralization of speech: the effects of age, sex, race and socio-economic class. *Neuropsychologia* 14, 363-370.
- Bradley, L. e Bryant, P. E. (1983). Categorizing sounds and learning to read: a causal connection. *Nature* 301, 419-421.
- Bradshaw, J. e Nettleton, N. (1981). The nature of hemispheric specialization in man. *The Behavioral and Brain Sciences* 4, 51-91.
- Bradshaw, J. e Nettleton, N. (1983). *Human Cerebral Asymmetry*. New Jersey: Prentice - Hall Inc. (pp. 228 - 241)
- Brady, S., Shankweiler, D. e Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology* 35, 345-367.
- Brandt, J. e Rosen, J. J. (1980). Auditory phonemic perception in dyslexia: Categorical identification and discrimination of stop consonants. *Brain and Language* 9, 324-337.
- Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology* 47, 191-196.
- Bryden, M. P. e Sprott, D. A. (1981). Statistical determination of degree of laterality. *Neuropsychologia* 19, 571 - 581.
- Burns, E. M. e Ward, W. D. (1978). Categorical perception - phenomenon or epiphonemenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America* 68, 456-468.
- Cameron, R. F., Currier, R. D. e Haerer, A. F. (1971). Aphasia and literacy. *British Journal of Disorders of Communications* 6, 161-163.
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. G. e Carbone, E. (1973). The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54, 421-428.
- Cary, L. (1988). *Análise explícita das unidades da fala nos adultos não alfabetizados*. Dissertação de candidatura ao grau de Doutor. Faculdade de Psicologia e de Ciências da Educação, Universidade de Lisboa.
- Cary, L. e Morais, J. (1979). A aprendizagem da leitura e consciência da estrutura fonética da fala. *Revista Portuguesa de Psicologia* 14/15/16, 97-106.

- Castro, M. S. L. V. (1983). *Aspectos psicológicos da assimetria hemisférica: Síntese da investigação comportamental e estudo exploratório*. Monografia de mestrado. Porto: Faculdade de Psicologia e de Ciências da Educação.
- Chomsky, N. e Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Cintra, L. F. L. (1983). *Estudos de Dialectologia Portuguesa*. Lisboa: Sá da Costa Editora.
- Cole, R. A. e Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.) *Perception and Production of Fluent Speech*. Hillsdale, N. Y.: Erlbaum. (pp. 133-163)
- Cole, R. A. e Scott, B. (1974). Toward a theory of speech perception. *Psychological Review* 81, 384-374.
- Connine, C. M. e Clifton Jr., C. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 13, 291-299.
- Content, A. (1984). L'analyse phonétique explicite de la parole et l'acquisition de la lecture. *L'Année Psychologique* 85, 73-99.
- Content, A. (1985). Le développement de l'habileté d'analyse phonétique de la parole. *L'Année Psychologique* 85, 73-99.
- Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M. e Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America* 24, 597-606.
- Cooper, F. S., Gaitenby, J. H. e Nye, P. W. (1984). Evolution of reading machines for the blind: Haskins Laboratories' research as a case history. *Journal of Rehabilitation Research and Development* 21, 51-86.
- Cooper, W. A. Jr. e O'Malley, H. (1975). Effects of dichotically presented simultaneous synchronous and delayed auditory feedback on key tapping performance. *Cortex* 11, 206 - 216.
- Corballis, M. C. e Morgan, M. J. (1978). On the biological basis of human laterality: I. Evidence for a maturational left-right gradient; e II. The mechanisms of inheritance. *The Behavioral and Brain Sciences* 1, 261-336.
- Crowder, R. (1982). A common basis for auditory sensory storage in perception and immediate memory. *Perception & Psychophysics* 31, 477-483.
- Cunha, C. e Cintra, L. F. L. (1987). *Nova Gramática do Português Contemporâneo*. Lisboa: Edições João Sá da Costa.

- Cutting, J. E. (1975). Aspects of phonological fusion. *Journal of Experimental Psychology: Human Perception and Performance* **1**, 105-120.
- Cutting, J. E. (1976). Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review* **83**, 114-140.
- Cutting, J. E. (1982). Plucks and bows are categorically perceived, sometimes. *Perception & Psychophysics* **31**, 462-476.
- Cutting, J. E. e Day, R. S. (1975). The perception of stop-liquid clusters in phonological fusion. *Journal of Phonetics* **3**, 99-113.
- Cutting, J. E. e Rosner, B. S. (1974). Discrimination functions predicted from categories in speech and music. *Perception & Psychophysics* **20**, 87-88.
- Damáσιο, A., Castro-Caldas, A., Grosso, J. e Ferro, J. (1976). Brain specialization for language does not depend on literacy. *Archives of Neurology* **33**, 300-301.
- Damáσιο, H., Damáσιο, A. R., Castro-Caldas, A. e De S. Hamsher, K. (1979). Reversal of ear advantage for phonetically similar words in illiterates. *Journal of Clinical Neuropsychology* **1**, 331-338.
- Darwin, C. (1976). The perception of speech. In E. Carterette e M. Friedman (Eds.) *Handbook of Perception, Vol. III*. New York: Academic Press. (pp. 175-226)
- Day, R. S. (1968). Fusion in dichotic listening. (Doctoral dissertation, Stanford University). *Dissertation Abstracts International* **29**, 2649B (University Microfilms No 69-211).
- Day, R. S. (1970-a). Temporal-order judgments in speech: are individuals language-bound or stimulus-bound? *Haskins Laboratories Status Report on Speech Research* **SR- 21/22**, 71-87.
- Day, R. S. (1970-b). Temporal-order perception of a reversible phoneme cluster. *Journal of the Acoustical Society of America* **48**, 95 (Abstract). Também in *Haskins Laboratories Status Report on Speech Research* **SR-24**, 47-56.
- Day, R. S. (1973-a). On learning "secret languages". *Haskins Laboratories Status Report on Speech Research* **SR-34**.
- Day, R. S (1973-b). Digit span memory in language-bound and stimulus-bound subjects. *Journal of the Acoustical Society of America* **54**, 287 (Abstract).
- Denenberg, V. H. (1981). Hemispheric laterality in animals and the effects of early experience. *The Behavioral and Brain Sciences* **4**, 1-49.

- Denis, M. e Whitaker, H. A. (1977). Hemisphere equipotentiality and language acquisition. In S. J. Segalowitz e F. A. Gruber (Eds.). *Language Development and Neurological Theory*. N. York: Academic Press.
- DiSimoni, F. G. (1974). Influence of consonant environment on duration of vowels in the speech of three-, six-, and nine-year-old children. *Journal of the Acoustical Society of America* 55, 362-363.
- Dorman, M. e Geffner, D. (1974). Hemispheric specialization for speech perception in six year old black and white children from low and middle socio-economic classes. *Cortex* 10, 171 - 176.
- Efron, R. (1970). The relationship between the duration of a stimulus and the duration of perception. *Neuropsychologia* 8, 37-55.
- Eguchi, S. (1976). Difference limens for the format frequencies: normal adult values and their development in children. *Journal of the American Audiological Society* 1, 145-149.
- Ehri, L. C. e Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology* 71, 26-40.
- Eilers, R. E., Wilson, W. R. e Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research* 20, 766-780.
- Eimas, P. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by Young infants. *Perception & Psychophysics* 18, 341 - 347.
- Eimas, P., Siqueland, E. R., Jusczyk, P. e Vigorito, J. (1971). Speech perception in infants. *Science* 171, 303-306.
- Eling, P., Marshall, J. C., Galen, G. (1981). The development of language lateralization as measured by dichotic listening. *Neuropsychologia* 6, 767-773.
- Elliot, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *Journal of the Acoustical Society of America* 66, 651-653.
- Elliot, L. L., Connors, S., Kille, E., Levin, S., Ball, K. e Katz, D. R. (1979). Children's understanding of monosyllabic nouns in quiet and in noise. *Journal of the Acoustical Society of America* 66, 12-21.
- Elliot, L. L., Longinotti, C., Meyer, D., Raz, I. e Zucker, K. (1981). Developmental differences in identifying and discriminating CV syllables. *Journal of the Acoustical Society of America* 70, 669-677.
- Elman, J. L. e McClelland, J. L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. Lass (Ed.) *Speech and Language Vol. 10*. New York: Academic Press.

- Elman, J. L. e McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language* 27, 143-165.
- Ferro, J., Castro-Caldas, A., Martins, J. e Salgado, V. Verbal and visuo-spatial disturbances: influence of literacy. Manuscrito não publicado.
- Fitch, H. L., Halwes, T., Erickson, D. M. e Liberman, A. M. (1980). Perceptual equivalence of two acoustic cues for stop-consonant manner. *Perception & Psychophysics* 27, 343-350.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, Mass.: MIT Press.
- Fodor, J. (1985). Précis of "The Modularity of Mind." *The Behavioral and Brain Sciences* 8, 1-42.
- Foss, D. J. e Hakes, D. T. (1978). *Psycholinguistics: An Introduction to the Psychology of Language*. Englewood Cliffs, N. Y.: Prentice-Hall, Inc.
- Fox, B. e Routh, D. K. (1980). Phonemic analysis and severe reading disability. *Journal of Psycholinguistic Research* 9, 115-119.
- Fox, R. A. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance* 10, 526-540.
- Frauenfelder, V. H. e Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition* 25, 1-20.
- Fromkin, V. A., Krashen, S., Curtiss, S., Rigler, D. e Rigler, M. (1974). The development of language in Genie: A case of language acquisition beyond the Critical Period. *Brain and Language* 1, 81 - 107.
- Fry, D. B., Abramson, A. S., Eimas, P. D. e Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech* 5, 171-189.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6, 110-125.
- Gazzaniga, M. S. e Ledoux, J. E. (1978). *The integrated mind*. New York: Plenum Press.
- Geffen, G. (1978). Human laterality: cerebral dominance and handedness. *The Behavioral and Brain Sciences* 1, 295.
- Geffner, D. S. e Dorman, M. F. (1975). Hemispheric specialization for speech perception in four-year-old children from low and middle socio-economic classes. *Cortex* 12, 71 - 73.

- Geffner, D. S. e Hochberg, I. (1971). Ear laterality performance of children from low and middle socio-economic levels on a verbal dichotic listening task. *Cortex* 7, 193 - 203.
- Geschwind, N. (1984). The biology of cerebral dominance: Implications for cognition. *Cognition* 17, 193-208.
- Gibson, J. J. (1979). *An Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Gleitman, L. e Rozin, P. (1977). The structure and acquisition of reading (I): relations between orthographies and the structure of language. In A. S. Reber e D. L. Scarborough (Eds.) *Toward a Psychology of Reading*. New York: Wiley. (pp. 53-141)
- Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K. K. e Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology* 32, 401-424.
- Goodglass, H. e Geschwind, N. (1976). Language disorders (Aphasia). In E. C. Carterette e M. P. Friedman (Eds.) *Handbook of Perception VII*. New York: Academic Press.
- Hall, J. W., Ewing, A., Tinzmann, M. B. e Wilson, K. P. (1981). Phonetic coding in dyslexics and normal readers. *Bulletin of the Psychonomic Society* 17, 177-178.
- Hall, J. W., Wilson, K. P., Humphreys, M. S., Tinzmann, M. B. e Bower, P. M. (1983). Phonemic similarity effects in good vs. poor readers. *Memory & Cognition* 11, 520-527.
- Halwes, T. G. (1969). *Effects of dichotic fusion on the perception of speech*. Dissertação de doutoramento, Universidade de Minnesota.
- Hanson, V. L. (1977). Within-category discriminations in speech perception. *Perception & Psychophysics* 21, 423-430.
- Harnad, S. (Ed.) (1987). *Categorical Perception: The Groundwork of Cognition*. Cambridge, Mass.: Cambridge University Press.
- Hécaen, H. (1976). Acquired aphasia in children and the ontogenesis of hemispheric functional specialization. *Brain and Language* 3, 114 - 134.
- Heider, E. R. e Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology* 3, 337-354.
- Hellige, J. B. (Ed.) (1983) *Cerebral hemisphere Asymmetry: Method, Theory, and Application*. New York: Praeger.
- Hillenbrand, J. (1984). Speech perception by infants: categorization based on nasal consonant place of articulation. *Journal of the Acoustical Society of America* 75, 1613 - 1622.

- Hiscock, M. e Kinsbourne, M. (1977). Selective listening asymmetry in preschool children. *Developmental Psychology* 13, 217 - 224.
- Hiscock, M. e Kinsbourne, M. (1980). Asymmetries of selective listening and attention switching in children. *Developmental Psychology* 16, 70 - 80.
- Holender, D. (1987). Synchronic description of present-day writing systems: some implications for reading research. In I. K. O'Regan e A. Levy-Schoen (Eds.) *Eye Movements: from Physiology to Cognition*. Amsterdam: Elsevier North Holland. (pp. 397-420)
- Joanette, Y., Lecours, A. R., Lepage, Y. e Lamoureux, M. (1983). Language in right-handers with right-hemisphere lesions: a preliminary study including anatomical, genetic and social factors. *Brain and Language* 20, 217-248.
- Jusczyk, P. W., Smith, L. B. e Murphy, C. (1981). The perceptual classification of speech. *Perception & Psychophysics* 30, 10-23.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, N. J.: Prentice-Hall.
- Kaiser, J. F. (1986). Some observations on vocal tract operation from a fluid flow point of view. Conferência integrada no ciclo "Distinguished Lecturer Series" 1986-1987, Florida Atlantic University. Abril.
- Katz, R. B. (1986). Phonological deficiencies in children with reading disability: Evidence from an object-naming task. *Cognition* 22, 225-257.
- Keele, S. W. e Lyon, D. R. (1982). Individual differences in speech fusion: methodological and theoretical explorations. *Perception & Psychophysics* 32, 434-442.
- Kelso, J. A. S., Saltzman, E. e Tuller, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics* 14, 29-59.
- Kewley-Port, D. (1983) Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America* 73, 322-334.
- Kewley-Port, D., Pisoni, D. B. e Studdert-Kennedy, M. (1983). Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America* 73, 1779-1793.
- Kimura, D. (1963). Speech lateralization in young children as determined by an auditory test. *Journal of Comparative and Physiological Psychology* 56, 899 - 902.

- Kinsbourne, M. (1975). The mechanisms of hemispheric control of the lateral gradient of attention. In P. M. A. Rabbitt e S. Dornic (Eds.) *Attention and Performance V*. New York: Academic Press.
- Kinsbourne, M. e Hiscock, M. (1977). Does cerebral dominance develop? In S. I. Segalowitz e F. A. Gruber (Eds.) *Language Development and Neurological Theory*. New York: Academic Press. (pp. 171-191)
- Kolinsky, R. (1988) *La separabilité des propriétés dans la perception des formes*. Dissertação de doutoramento. Bruxelles: Université Libre de Bruxelles.
- Koopmans, M. (1986). Formal schooling and task familiarity: A reply to Morais et al. *Cognition* 24, 109-111.
- Krause, S. E. (1982). Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults. *Journal of the Acoustical Society of America* 71, 990-995.
- Kuhl, P. K. (1987). The special-mechanisms debate in speech research: Categorization tests on animals and infants. In S. Harnad (Ed.) *Categorical Perception: The Groundwork of Cognition*. Cambridge, Mass.: Cambridge University Press (pp. 355-386)
- Kuhl, P. K. e Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America* 63, 905-917.
- Kuhl, P. K. e Padden, D. M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & Psychophysics* 32, 542 - 550.
- Kuhl, P. K. e Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America* 73, 1003 - 1010.
- Lacerda, F. (1980). Estudo acústico-perceptivo de /f/, /s/ e /ʃ/ em português. *Relatório do Grupo de Fonética e Fonologia* 2, 1-29.
- Lacerda, F. (1982). Acoustic perceptual study of the Portuguese voiceless fricatives. *Journal of Phonetics* 10, 11-12.
- Ladefoged, P. (1982). *A Course in Phonetics*. New York: Harcourt Brace Jovanovitch.
- Lahiri, A., Gwirth, L. e Blumstein, S. E. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a crosslanguage study. *Journal of the Acoustical Society of America* 76. 391-404.
- Lane, H. L. (1965). Motor theory of speech perception: A critical review. *Psychological Review* 72, 275-309.

- Lasky, R. E., Syrdal-Lasky, A. e Klein, R. E. (1975). VOT discrimination by four to six-and-a-half-month-old infants from Spanish environments. *Journal of Experimental Child Psychology* 20, 215-225.
- Lauter, J. L. (1982). Dichotic identification of complex sounds: Absolute and relative ear advantages. *Journal of the Acoustical Society of America* 71, 701-707.
- Lecours, A. R. (1980). Social factors as determinants of human brain physiology. Does left brain specialization for verbal skills depend on literacy? Candidatura a subsídio da Fundação H. F. Guggenheim.
- Lecours, A. R., Mehler, J., Mehler, M. A., Caldeira, A., Castro, M. J., Dehaut, F., Delgado, R., Gurd, J., Karmann, D. F., Jakubowitz, R., Osorio, Z., Scliar-Cabral, L. e Junqueira, M. S. (1987-a). Illiteracy and brain damage-1. Aphasia testing in culturally contrasted populations (control subjects). *Neuropsychologia* 25, 231-245.
- Lecours, A. R., Mehler, J., Mehler, M. A., Aguiar, L. R., Silva, A. B., Caetano, M., Camarotti, H., Castro, M. J., Dehaut, F., Dumais, C., Gauthier, L., Gurd, J., Leitão, O., Macial, J., Machado, S., Memaragno, R., Oliveira, L. M. Paciornik, J., Sanvito, W., Silva, E. S., Silifrandi, M. e Torné, C. H. (1987-b). Illiteracy and brain Damage-2. Manifestations of unilateral neglect in testing auditory comprehension with iconographic materials. *Brain and Cognition* 6, 243-265.
- Lecours, A. R., Mehler, J., Mehler, M. A. e 31 colaboradores (1988). Illiteracy and brain damage-3: A contribution to the study of speech and language disorders in illiterates with unilateral brain damage (initial testing). *Neuropsychologia* 26, 575-589.
- Lemay, M. e Geschwind, N. (1975). Hemispheric differences in the brains of the great apes. *Brain, Behavior and Evolution* 11, 48-52.
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley.
- Lewrowicz, D. J. e Turkewitz, G. (1982). Influence of hemispheric specialization in sensory processing in infants: age and gender related effects. *Journal of Developmental Psychology* 18, 301-308.
- Liberman, A. M. (1970). Some characteristics of perception in the speech mode. In A.R.N.M.D. (Res. Publ.), *Perception and its Disorders*. USA: Association for Research in Nervous & Mental Disease.
- Liberman, A. M. (1982). On finding that speech is special. *American Psychologist* 37, 148-167.
- Liberman, A. M., Cooper, F. S., Harris, K. S. e MacNeilage, P. F. (1962). A motor theory of speech perception. *Proceedings of the Speech*

*Communication Seminar*. Stockholm: Royal Institute of Technology, D<sub>3</sub>.

- Liberman, A. M., Cooper, F. S., Shankweiler, D. P. e Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review* 74, 431-461.
- Liberman, A. M., Delattre, P. e Cooper, F. S. (1952). The role of selected stimulus-variables in the perception of the unvoiced stop consonants. *American Journal of Psychology* 65, 497-516.
- Liberman, A. M., Delattre, P. C., Cooper, F. S. e Gerstman, L. J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied* 68, 1-13.
- Liberman, A. M., Harris, K. S., Eimas, P. D., Lisker, L. e Bastian, J. (1961-a). An effect of learning on speech perception: the discrimination of durations of silence with and without phonemic significance. *Language and Speech* 54, 175-195.
- Liberman, A. M., Harris, K. S., Hoffman, H. S. e Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 53, 358-368.
- Liberman, A. M., Harris, K. S., Kinney, J. A. e Lane, H. (1961-b). The discrimination of relative onset time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology* 61, 379-388.
- Liberman, A. M., Isenberg, D. e Rakerd, B. (1981). Duplex perception of cues for stop consonants: evidence for a phonetic mode. *Perception & Psychophysics* 30, 133-143.
- Liberman, A. M. e Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1-36.
- Liberman, A. M. e Pisoni, D. B. (1977). Evidence for a special speech-perceiving subsystem in the human. In T. H. Bullock (Ed.). *The Recognition of Complex Acoustic Signals*. Berlin: Dahlem Konferenzen. (pp. 59-76)
- Liberman, A. M. e Studdert-Kennedy, M. (1978). Phonetic perception. In R. Held, H. Leibowitz e H.-L. Teuber (Eds) *Handbook of Sensory Physiology Vol. 8: Perception*. New York: Springer-Verlag. (pp. 143-178)
- Liberman, I. Y. (1971). Basic research in speech and lateralization of language: some implications for reading disability. *Bulletin of the Orton Society* 21, 71-87.
- Liberman, I. Y., Shankweiler, D., Fisher, F. W. e Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology* 18, 201-212.

- Lieberman, P. (1986). Language, communication and rule-governed behaviour. Comunicação apresentada no curso OTAN "Evolutionary Biology of Intelligence", Poppi, Julho 1986.
- Lindblom, B. E. F. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773-1781.
- Lindsay, P. H. e Norman, D. A. (1977). *Human Information Processing: An Introduction to Psychology* (2<sup>a</sup> Ed.). New York: Academic Press.
- Lisker, L. (1975). Is it VOT or a first-formant transition detector? *Journal of the Acoustical Society of America* 57, 1547-1551.
- Lisker, L. (1978-a). In qualified defense of VOT. *Language and Speech* 21, 375-383.
- Lisker, L. (1978-b). Rapid vs. rabid: A catalogue of acoustic features that may cue the distinction. *Haskins Laboratories Status Report on Speech Research SR-54*, 127-132.
- Lisker, L. (1985). The pursuit of invariance in speech signals. *Journal of the Acoustical Society of America* 77, 1199-1202.
- Lisker, L. e Abramson, A. S. (1967). The voicing dimension: some experiments in comparative phonetics. In *Proceedings of the Sixth International Congress of Phonetic Sciences*. Prague: Academia Publishing House of the Czechoslovak Academy of Sciences.
- Lokker, R. e Morais, J. (1985). Ear differences in children at two years of age. *Neuropsychologia* 33, 127-129.
- Lundberg, I., Olofsson, A. e Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology* 21, 159-173.
- Luria, A. R. (1970). The functional organization of the brain. *Scientific American* 222, 66-78.
- Mann, V. A. (1984). Reading skill and language skill. *Developmental Review* 4, 1-15.
- Mann, V. A. (1985). A cross-linguistic perspective on the relation between temporary memory skills and early reading ability. *Remedial and Special Education* 6, 37-42.
- Mann, V. A. (1986). Phonological awareness: The role of reading experience. *Cognition* 24, 65-92.
- Mann, V. A. e Liberman, Y. I. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities* 17, 592-599.
- Mann, V. A. e Repp, B. H. (1980). Influence of vocalic context on the perception of [ ] - [s] distinction: I. Temporal factors. *Perception & Psychophysics* 28, 213-228.

- Mann, V. A. e Repp, B. H. (1981). Influence of preceding fricative on stop consonant perception. *Journal of the Acoustic Society of America* 69, 548-558.
- Marcel, A. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology* 15, 238-300.
- Marcel, A. (1986). Consciousness and processing: choosing and testing a null hypothesis. *The Behavioral and Brain Sciences* 9, 40-41.
- Marie, P. (1926). *Travaux et mémoires*. Paris: Masson (cit. in Lecours, 1980).
- Mark, L. S., Shankweiler, D. e Liberman, I. Y. (1977). Phonetic recoding and reading difficulty in beginning readers. *Memory & Cognition* 5, 623-629.
- Marques, J. H. F. (1969). *Estudo sobre a escala de inteligência de Wechsler para crianças: Sua adaptação e aferição para Portugal*. Lisboa: Ramos, Afonso e Moita.
- Marshall, J. C., Caplan, D. e Holmes, J. M. (1975). The measure of laterality. *Neuropsychologia* 13, 315-321.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition* 25, 71-102.
- Martins, M. R. D. (1971). Análise acústica das vogais orais tónicas em português. *Boletim de Filologia* 22, 303-314.
- Massaro, D. W. e Cohen, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 9, 753-771.
- Massaro, D. W. (1974). Perceptual units in speech recognition. *Journal of Experimental Psychology* 102, 199-208.
- Mateus, M. H. M. (1975). *Aspectos da Fonologia Portuguesa*. Lisboa: Instituto Nacional de Investigação Científica, Centro de Linguística da Universidade de Lisboa.
- Mattingly, I. G. (1972). Reading, the linguistic process and linguistic awareness. In I. F. Kavenagh e I. G. Mattingly (Eds.) *Language by Ear and by Eye*. Cambridge, Mass.: MIT Press.
- Mattingly, I. G. e Liberman, A. M. (1969). The speech code and the physiology of language. In K. N. Leibovic (Ed.). *Information Processing in the Nervous System*. New York: Springer-Verlag. (pp. 97-117)
- Mattingly, I. G. e Liberman, A. M. (1986). Specialized perceiving systems for speech and other biologically significant sounds. In G. M.

- Edelman, W. E. Gall e W. M. Cowan (Eds.), *Functions of the Auditory System*. New York: Wiley.
- Mattingly, I. G., Liberman, A. M., Syrdal, A. K. e Halwes, T. (1971). Discrimination in speech and nonspeech modes. *Cognitive Psychology* 2, 131-157.
- May, J. G. (1981). Acoustic factors that may contribute to categorical perception. *Language and Speech* 24, 273-284.
- McGurk, H. e McDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746-748.
- McKeever, W. F., Hoemann, H. W., Florian, V. A. e VanDeventer, A. D. (1976). Evidence of minimal cerebral asymmetries for the processing of English words and American sign language in the congenitally deaf. *Neuropsychologia* 14, 413 - 423.
- Mehler, J., Morton, J. e Jusczyk, P. W. (1984). On reducing language to biology. *Cognitive Neuropsychology* 1, 83-116.
- Miller, G. A. (1981). *Language and Speech*. San Francisco: W. H. Freeman and Company.
- Miller, G. A. e Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America* 27, 338-352.
- Miyawaki, K., Strange, W., Verbrugge, R. Liberman, A. M., Jenkins, Y. J. e Fujimura, O. (1975). An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, 18, 331-340.
- Molfese, D. L. e Molfese, V. J. (1979). Hemisphere and stimulus differences as reflected in the cortical responses of newborn infants to speech stimuli. *Developmental Psychology* 15, 505-511.
- Morais, J. (1981). Le test d'écoute dichotique en tant que prédicteur de la dominance cérébrale chez les normaux. *Acta Neurologica Belgica* 81, 144-152.
- Morais, J. (1982). The two sides of cognition. In J. Mehler, E. C. T. Walker e M. Garret (Eds.). *Perspectives on Mental Representation. Experimental and Theoretical Studies of Cognitive Processes and Capacities*. Hillsdale, N. J: Lawrence Erlbaum Associates, Inc. (pp. 277-309)
- Morais, J. (1985). Literacy and awareness of the units of speech: implications for research on the units of perception. *Linguistics* 23, 707-721.
- Morais, J. (1987). Phonetic awareness and reading acquisition. *Psychological Research* 49, 147-152.

- Morais, J., Alegria, J. e Content, A. (1987-a). The relationship between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de Psychologie Cognitive* 7, 415-438.
- Morais, J., Alegria, J. e Content, A. (1987-b). Segmental awareness: respectable, useful, and almost always necessary. *Cahiers de Psychologie Cognitive* 7, 530-556.
- Morais, J., Bertelson, P., Cary, L. e Alegria, J. (1986). Literacy training and speech segmentation. *Cognition* 24, 45-64.
- Morais, J., Cary, L., Alegria, J. e Bertelson, P. (1979). Does awareness of speech as a sequence of phonemes arise spontaneously? *Cognition* 7, 323-331.
- Morais, J., Cluytens, M., Alegria, J. e Content, A. (1986). Speech-mediated retention in dyslexics. *Perceptual and Motor Skills* 62, 119-126.
- Moutier, F. (1908). *L'aphasie de Broca*. Paris: Steinheil. (cit. in Lecours, 1980).
- Neisser, U. (1967) *Cognitive Psychology*. New York: Appleton-Century Crofts.
- Nittrouer, S. e Studdert-Kennedy, M. (1986). The role of coarticulatory effects in the perception of fricatives by children and adults. *Haskins Laboratories Status Report on Speech Research* SR-88, 73-93.
- Norman, D. e Bobrow, D. (1975). On data limited and resource limited processing. *Journal of Cognitive Psychology* 7, 44-60.
- Norris, D. e Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Perception & Psychophysics* 43, 541-550.
- Notebohm, F. (1970). Ontogeny of bird song. *Science* 167, 950-956.
- Nusbaum, H. C. , Walley, A. C., Carrell, T. D. e Ressler, W. H. (1982). Controlled perceptual strategies in phonemic restoration. *Research on Speech Perception Progress Report* 8, 83-103. Bloomington, Ind.: Department of Psychology, Indiana University.
- Obrzut, J. E., Hynd, G. W. e Boliek-Uphoof, C. (no prelo). Lateral asymmetries in learning disabled children: A review. In S. J. Ceci (Ed). *Handbook of Cognitive, Social and Neuropsychological Aspects of Learning Disabilities*. Hillsdale, N. J.: Lawrence Erlbaum Associates, Inc.
- Orchik, D. J. e Burgess, J. (1977). Synthetic sentence identification as a function of the age of the listener. *Journal of the American Audiological Society* 3, 42-46.
- Pardal, E. d' A. (1977) *Aspects de la Phonologie (Generative) du Portugais*. Lisboa: Instituto Nacional de Investigação Científica, Centro de Linguística da Universidade de Lisboa.

- Pastore, R. E., Ahron W. A., Baffuto, K. J., Friedman, C., Puleo, J. S. e Fink, E. A. (1977). Common-factor model of categorical perception. *Journal of Experimental Psychology: Human Perception and Performance* 3, 686-696.
- Perkell, J. B. e Klatt, D. H. (Eds) (1986). *Invariance and Variability in Speech Processes*. Hillsdale, N. Y.: Lawrence Erlbaum Associates.
- Peterson, G. E. e Shoup, J. E. (1966). A physiological theory of phonetics. *Journal of Speech and Hearing Research* 9, 5-67.
- Pipe, M. E. (1985). Attenuation of dichotic-listening ear advantages by stimulus bias. *Neuropsychologia* 23, 437-440.
- Pinto, A. (1986). *Amplitude de Memória Imediata*. Faculdade de Psicologia e de Ciências de Educação, Universidade do Porto.
- Pirozzolo, F. J., Rayner, K., e Hynd, G. W. (1983). The measurement of hemispheric asymmetries in children with developmental reading disabilities. In J. B. Hellige (Ed.) *Cerebral Hemisphere Asymmetry: Method, Theory and Application*. New York: Praeger.
- Pisoni, D. B. (1973). Perceptual processing time for consonants and vowels. *Journal of the Acoustical Society of America* 53, 369 (Sumário).
- Pisoni, D. B. (1975). Auditory short-term memory and vowel perception. *Memory & Cognition* 3, 7-18.
- Pisoni, D. B. (1985). Speech perception: Some new directions in research and theory. *Journal of the Acoustical Society of America* 78, 381-388.
- Pisoni, D. B., Aslin, R. N. Perey, A. J. e Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8, 297-314.
- Pisoni, D. B. e Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition* 25, 21-52.
- Pizzamiglio, L. e Cecchini, M. (1971). Development of the hemispheric dominance in children from 5 to 10 years of age and their relation with the development of cognitive processes. *Brain Research* 31, 361 - 378.
- Poeck, K. (Ed) (1982). *Klinische Neuropsychologie*. Stuttgart: Thieme.
- Poltrock, S. E. e Hunt, E. (1977). Individual differences in phonological fusion and separation errors. *Journal of Experimental Psychology: Human Perception and Performance* 3, 62-74.
- Porter, R. J. e Hughes, L. F. (1983). Dichotic listening to CVS: Method, interpretation, and application. In J. B. Hellige (Ed.) *Cerebral Hemisphere Asymmetry: Method, Theory and Application*. New York: Praeger.

- Pylyshyn, Z. (1980). Cognitive representation and the process-architectural distinction. *The Behavioral and Brain Sciences* 3, 154-169.
- Rabbitt, P. M. A. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology* 20, 241-248.
- Rand, T. C. (1974). Dichotic release from masking for speech. *Journal of the Acoustical Society of America* 55, 678-680.
- Raphael, L. J. (1972). Preceding vowel duration as cue to the perception of voicing of American English consonants in word-final position. *Journal of the Acoustical Society of America* 51, 1296-1303.
- Read, C. e Ruyter, L. (1985). Reading and spelling skills in adults of low literacy. *Remedial and Special Education* 6, 43-52.
- Read, C., Zhang, Y., Nie, H. e Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic spelling. *Cognition* 24, 31-44.
- Remez, R. E., Rubin, P. E., Pisoni D. B. e Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science* 212, 947-950.
- Repp, B. H. (1975). Dichotic masking of consonants by vowels. *Journal of the Acoustical Society of America* 57, 724-735.
- Repp, B. H. (1977-a). Dichotic competition of speech sounds: the role of acoustic stimulus structure. *Journal of Experimental Psychology: Human Perception and Performance* 3, 37-50.
- Repp, B. (1977-b). Measuring laterality effects in dichotic listening. *Journal of the Acoustical Society of America* 62, 720-737.
- Repp, B. H. (1980). Stimulus dominance in fused dichotic syllables: Trouble for the category goodness hypothesis. *Journal of the Acoustical Society of America* 67, 288-305.
- Repp, B. H. (1981). On levels of description in speech research. *Journal of Acoustical Society of America* 69, 1462-1464.
- Repp, B. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin* 92, 81-110.
- Repp, B. H. (1984). Categorical perception: issues, methods, findings. In N. J. Lass (Ed.). *Speech and Language: Advances in Basic Research and Practice* Vol. 10, 243-335.
- Repp, B. H. e Bentin, S. (1984) Parameters of spectral/temporal fusion in speech perception *Perception & Psychophysics* 36, 523-530.

- Repp, B. e Liberman, A. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.) *Categorical Perception: The Groundwork of Cognition*. Cambridge, Mass.: Cambridge University Press.
- Repp, B. H. Milburn, C. e Ashkenas, J. (1983). Duplex perception: Confirmation of fusion. *Perception & Psychophysics* 33, 333-337.
- Rogers, L. J. e Anson, J. M. (1979). Lateralization of function in the chicken fore-brain. *Pharmacology, Biochemistry and Behavior* 10, 679-686.
- Rosen, S. M. e Howell, P. (1981). Plucks and bows are not categorically perceived. *Perception & Psychophysics* 30, 156-168.
- Rosen, S. e Howell, P. (1987). Auditory, articulatory and learning explorations of categorical perception in speech. In S. Harnad. *Categorical Perception: the Groundwork of Cognition*. (pp. 113-160)
- Rosner, J. e Simon. D. P. (1971). The auditory analysis test: An initial report. *Journal of Reading Disabilities* 4, 384-392.
- Rozin, P. e Gleitman, L. R. (1977). The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principale. In A. S. Reber e D. L. Scarborough (Eds.) *Toward a Psychology of Reading*. Hillsdale. Lawrence Erlbaum.
- Samuel, A. G. (1977). The effect of discrimination training on speech perception: noncategorical perception. *Perception & Psychophysics* 22, 321-330.
- Samuel, A. G. e Ressler, W. H. (1986). Attention within auditory word perception: Insights from the phonemic restoration ilusion. *Journal of Experimental Psychology: Human Perception and Performance* 12, 70-79.
- Sawusch, Y. R. e Nusbaum, H. C. (1983). Auditory and phonetic processes in place perception for stops. *Perception & Psychophysics* 34, 560-568.
- Scholes, R. J. (Manuscrito em preparação). Using letters to make phonemic judgements: I. Segment delection.
- Schöner, G. e Castro, S. L. (submetido) Ciência da fala e sinérgica.
- Schöner, G. e Kelso, S. (1988). A synergetic theory of environmentally-specified and learned patterns of movement coordination. I. Relative phase dynamics. *Biological Cybernetics* 58, 71-80.
- Schwartz, S. e Kirsner, K. (1984). Can group differences in hemispheric asymmetry be inferred from behavioral laterality indices? *Brain & Cognition* 3, 57-70.
- Searle, C. L., Jacobson, J. Z. e Rayment, S. G. (1979). Discrimination based on human audition. *Journal of the Acoustical Society of America* 65, 799-809.

- Segalowitz, S. J. e Chapman, J. S. (1980). Cerebral arymmetry for speech in neonates: A behavioral measure. *Brain and Language* 9, 281-288.
- Seidenberg, M. S. e Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory* 5, 546-554.
- Sexton, M. A. e Geffen G. (1981). Phonological fusion in dichotic monitoring. *Journal of Experimental Psychology: Human Perception and Performance* 7, 422-429.
- Shallice, T. (1984). More functionally isolable subsystems but fewer "modules"? *Cognition* 17, 243-252.
- Shankweiler, D. e Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition* 24, 139-168.
- Shankweiler, D., Liberman, I. Y., Mark, L. S. Fowler, C. A. e Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory* 5, 531-545.
- Shankweiler, D. e Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to the left and right ears. *Quarterly Journal of Experimental Psychology* 19, 59-63.
- Sherman, G. F., Garranati, J. A., Rosen, G. D., Yutzer, D. A. e Denenberg, U. H. (1980). Brain and behavior asymmetries in spatial preference in rats. *Brain Research* 192, 61-67.
- Simon, C. e Fourcin, A. J. (1978). Cross-language study of speech-pattern learning. *Journal of the Acoustical Society of America* 63, 925-935.
- Smith, B. L. (1978). Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics* 6, 37-67.
- Sorin, C. e Thouin-Daniels, S. C. (1983). Effects of auditory fatigue on speech intelligibility and lexical decision in noise. *Journal of the Acoustical Society of America* 74, 456-466.
- Spiegel, M. F. e Watson, C. S. (1981). Factors in the discrimination of tonal patterns. III. Frequency discrimination with components of well-learned patterns. *Journal of the Acoustical Society of America* 69, 223-230.
- Sprott, D. A. e Bryden, M. P. (1983). Some problems with "Some problems with Bryden and Sprott's statistical determination of degree of lateralization". *Neuropsychologia* 21, 299 - 300.
- Stanovich, K. E. (1983). Individual differences in the cognitive processes of reading: I. Word decoding. *Journal of Learning Disabilities* 15, 485-492.

- Starck, R., Genesee, F., Lambert, W. E. e Seitz, M. (1977). Multiple language experience and the development of cerebral dominance. In S. J. Segalowitz e F. A. Gruber (eds). *Language Development and Neurological Theory*. New York: Academic Press.
- Stevens, K. N. (1978). The speech signal. In Y. F. Kavannagh e W. Strange (Eds.) *Speech and Language in the Laboratory, School and Clinic*. Cambridge Mass.: Massachussets Institute of Technology.
- Stevens, K. N. (1981). Constraints imposed by the auditory system on the properties used to classify speech sounds: data from phonology, acoustics and psychoacoustics. In T. Myers, Y. Laver e J. Anderson (Eds.) *The cognitive representation of speech*. North-Holland. (pp. 61-74)
- Stevens, K. N. e Blumstein, S. (1978) Invariance cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America* 64, 1358-1368.
- Stevens, K. N. e Blumstein, S. E. (1981) The search for invariant correlates of phonetic features. In P. E. Eimas e Y. Miller (Eds.) *Perspectives on the Study of Speech*. Hillsdale, N. Y.: Erlbaum. (pp. 1-38)
- Stevens, K. N., Liberman, A. M., Ohman, S. E. G. e Studdert-Kennedy, M. (1969). Cross-language study of vowel perception. *Language and Speech* 12, 1-23.
- Strange, W. e Halwes, T. (1971). Confidence ratings in speech perception research: evaluation of an efficient technique for discrimination testing. *Perception & Psychophysics* 9, 182-187.
- Streeter, L. A. (1975). Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259, 38-41.
- Studdert-Kennedy, M. (1974). The perception of speech. In T. A. Sebeok *Current Trends in Linguistics*. The Hague: Mouton.
- Studdert-Kennedy, M., Liberman, A. M., Harris, K. S. e Cooper, F. S. (1970). Motor theory of speech perception: A reply to Lane's critical review. *Psychological Review* 77, 234-249.
- Studdert-Kennedy, M. e Shankweiler, D. (1970). Hemispheric specialization for speech. *Journal of the Acoustical Society of America* 48, 579-594.
- Summerfield, Q. (1979). Use of visual information for phonetic perception. *Journal of Phonetic* 36, 314-331.
- Summerfield, Q. (1982). Differences between spectral dependencies in auditory and phonetic temporal processing: Relevance to the perception of voicing in initial stops. *Journal of the Acoustical Society of America* 72, 51-61.

- Swets, J. A. (Ed.). (1964). *Signal detection and recognition by human observers: Contemporary Readings*. New York: Wiley.
- Summerfield, Q. e Mc Grath, M. (1984). Detection and resolution of audiovisual incompatibility in the perception of vowels. *Quarterly Journal of Experimental Psychology* 36A, 51-74.
- Tallal, P. (1980). Auditory temporal perception, phonics and reading disabilities in children. *Brain and Language*. 9, 182-198.
- Tanner, W. e Swets, J. (1954). A decision making theory of visual detection. *Psychological Review* 61, 401-409.
- Tartter, V. C. (1982). Vowel and consonant manipulations and the dual-coding model of auditory storage: A re-evaluation. *Journal of Phonetics* 10, 217-224.
- Tartter, V. C. e Blumstein, S. E. (1981). The effects of pitch and spectral differences on phonetic fusion in dichotic listening. *Journal of Phonetics* 9, 251-259.
- Taylor, I. (1981). Writing systems and reading. In G. E. Mackinnon e T. G. Waller. *Reading Research* (Vol. 2) New York. Academic Press. (pp. 1-51).
- Teng, E. L. (1981). Dichotic ear difference is a poor index for the functional asymmetry between the cerebral hemispheres. *Neuropsychologia* 19, 235-240.
- Treisman, A. (no prelo). Properties, parts and objects. In K. Boff, L. Kaufman e J. Thomas (Eds.) *Handbook of Perception and Human Performance*. London: Wiley.
- Treisman, A. e Patterson, R. (1984). Emergent features, attention and object perception. *Journal of Experimental Psychology: Human Perception and Performance* 10, 12-32.
- Treisman, A. e Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology* 14, 107-141.
- Treisman, A. e Souther, J. (1986). Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General* 114, 285-310.
- Trevarthen, C. (1983). Development of the cerebral mechanisms for language. In U. Kirk (Ed.) *Neuropsychology of Language, Reading, and Spelling*. New York: Academic Press.
- Tsunoda, T. (1975). Functional differences between right and left cerebral hemispheres detected by the key-tapping method. *Brain and Language* 2, 152 - 170.

- Tuller, B. e Kelso, J. A. S. (1984). The relative timing of articulatory gestures: Evidence for relational invariants. *Journal of the Acoustical Society of America* 71, 1534-1543.
- Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review* 80, 1-52.
- Tzavaras, A., Kaprinis, G. e Gatzoyas, A. (1981). Literacy and hemispheric specialization for language: digit dichotic listening in illiterates. *Neuropsychologia* 19, 565-570.
- Von Mundy, V. G. (1957). Zur Frage der paarig veranlagten Sprachzentren. *Der Nervenarzt* 28, 212 - 216 (cit. in Lecours, 1980).
- Walley, A. C. e Carrell, T. D. (1983). Onset spectra and formant transitions in the adults and child's perception of place of articulation in stop consonants. *Journal of the Acoustical Society of America* 73, 1011-1022.
- Ward, T. B. (1983). Response tempo and separable-integral for an integral-to-separable processing sequence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*. 37, 471-482.
- Warren, R. M., Obusek, C. J. e Farmer R. M. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science* 164, 586-587.
- Weber, E. L. (1904). Das Schreiben als Ursache der einseitigen Lage des Sprachzentrums. *Zentralblatt für Physiologie* 18, 341.
- Werker, J. F. e Tees, R. C. (1984). Cross-language speech-perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.
- Werker, J. F. e Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology* 41, 48-61.
- Wexler, B. E. e Halwes, T. (1981). Increasing the power of dichotic methods: the fused rhymed words test. *Neuropsychologia* 21, 59-66.
- Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics* 35, 49-64.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman e D. R. Davies (Eds.) *Varieties of Attention*. Orlando: Academic Press.
- William, D. R., Verbrugge, R. R. e Studdert-Kennedy, M. (1983). Judging sine wave stimuli as speech and nonspeech. *Journal of the Acoustical Society of America* 74 S66.

- Williams, L. (1977). The perception of stop consonant voicing consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics* 21, 289-297.
- Witelson, S. F. e Pallie, W. (1973) Left hemisphere specialization for language in the newborn: Neuroanatomical evidence of asymmetry. *Brain* 96, 641-647.
- Woods, B. T. (1980). Observations on the neurological basis for initial language acquisition. In D. Caplan (Ed.). *Biological Studies of Mental Processes*. Cambridge, Mass: MIT Press.
- Yoneshige, J. e Elliot, L. L. (1981). Pure tone sensitivity and ear canal pressure at threshold in children and adults. *Journal of the Acoustic Society of America* 70, 1272-1276.
- Young, A. W. (1981). Asymmetry of cerebral hemispheric function during development. In J. W. T. Dickerson e H. Mc Gurk (Eds.). *Brain and Behavior: A Developmental Perspective*. Glasgow: Blackie.
- Zhurova, L. Y. (1973). The development of analysis of words into their sounds by preschool children. In C. A. Ferguson e D. I. Slobin (Eds.) *Studies of Child Language Development*. New York: Holt, Rinehart e Winston. (pp. 141-165)
- Zlatin, M. A. e Koenigsknecht, R. A. (1975). Development of the voicing contrast: perceptions of stop consonants. *Journal of Speech and Hearing Research* 18, 541-553.

**ANEXO 1. AFI: ALFABETO FONÉTICO INTERNACIONAL: símbolos relevantes para o português.**

**VOGAIS ORAIS**

Anterior ou palatal		
Abertura mínima	"livro"	[i]
Anterior (palatal)		
Abertura média	"medo"	[e]
Anterior (palatal)		
Abertura máxima	"pedra"	[ɛ]
Central		
Abertura máxima	"caso"	[a]
Central		
Mais fechada	"cão"	[ɑ]
Posterior ou velar		
Abertura máxima	"bola"	[ɔ]
Posterior (velar)		
Abertura média	"boca"	[o]
Posterior (velar)		
Abertura mínima	"luta"	[u]
Central		
Abertura mínima	"velar"	[ɐ]

**VOGAIS NASAIS**

Anterior		
Abertura mínima	"pinto"	[ĩ]
Anterior		
Abertura média	"pente"	[ẽ]
Central	"canto"	[ã]
Posterior		
Abertura média	"longe"	[õ]
Posterior		
Abertura mínima	"mundo"	[ũ]

**CONSOANTES OCLUSIVAS**

Bilabial surda	"pá"	[p]
Bilabial sonora	"bata"	[b]
Bilabial sonora nasal	"mala"	[m]
Dental-alveolar surda	"toda"	[t]
Dental-alveolar sonora	"dado"	[d]
Dental-alveolar sonora nasal	"nada"	[n]
Velar surda	"galo"	[k]
Velar sonora	"galo"	[g]
Palatal sonora nasal	"banho"	[ɲ]

**FRICATIVAS**

Bilabial sonora	"tábua"	[β]
Labio-dental surda	"fado"	[f]
Labio-dental sonora	"yala"	[v]
Dental-alveolar sonora	"fada"	[s]
Alveolar surda	"sala"	[ʃ]
Alveolar sonora	"zero"	[z]
Palato-alveolar surda	"chama"	[ʃ]
Palato-alveolar sonora	"gelo"	[ʒ]
Velar	"água"	[χ]

## APROXIMANTES

Lateral		
Dental-alveolar sonora	"lado"	[l]
Lateral		
Palatal sonora	"galho"	[ʎ]
Central (semivogal)		
Palatal sonora	"pai"	[j]
Central (semivogal)		
Labio-velar sonora	"quando"	[w]

## VIBRANTES

Batimento		
Apico-alveolar sonoro	"caro"	[r]
Trilo		
Uvular sonoro	"carro"	[R]

ANEXO L. 1: Teste de fluência de leitura. O teste era apresentado aos sujeitos com esta disposição colunar, e em espaço duplo, sendo-lhes pedido que lessem "tudo seguido, a uma velocidade rápida mas confortável".

ARCO	CACETE	EXEMPLO
ELA	MÃE	CAIXOTE
GRILO	CRISTAL	ABSURDO
SACO	PÔE	ARMAZÉM
TRIGO	ERGUER	ADMIRADO
TOLO	SOMBRA	RELUZENTE
PRATO	CÃES	AMENDOIM
MEU	LÍNGUA	MINÚSCULO
CAUDA	MÃOS	COMPANHEIRO
GEMA	QUANTIA	QUILÓMETRO
LEI	LIMÕES	ORQUESTRA
OUVIDO	BARULHO	INSTRUMENTO
NOZ	CALÇADA	INTERROGAÇÃO
MESMO	SILÊNCIO	PROXIMIDADE
MAÇÃ	PÁSSARO	CONSTITUCIONAL
NEBLINA	GIGANTE	

## Anexo ED.1: Teste de audição dicótica 1: Pares de palavras dissimilares

Pista 2.	Pista 1.	Pista 2.	Pista 1.
papa	côto	puro	bela
tosse	gato	cunho	dote
goto	cacho	bomba	pilha
cota	paço	pode	bulha
pato	cura	bote	dano
<u>capa</u>	<u>bota</u>	<u>pano</u>	<u>tolo</u>
pouco	barro	ponta	barra
tara	côco	garra	conto
gola	tacho	dique	bôda
dente	cara	bufo	pinto
baço	pica	tique	galo
<u>ponto</u>	<u>carro</u>	<u>baça</u>	<u>tonta</u>
dura	tela	calo	toda
bilha	gago	tufu	cano
calha	bica	pança	tinto
bota	pula	burra	penete
pata	cola	touro	gama
<u>gula</u>	<u>posse</u>	<u>gana</u>	<u>peso</u>
duro	palha	doca	canga
pala	bode	cana	teso
bago	punho	turra	caça
pomba	devo	ganga	dado
tela	pulha	dança	toca
pote	tela	gado	couro

## Anexo ED.2: Teste de audição dicótica 2: Pares de palavras similares

Pista 2.	Pista 1.	Pista 2.	Pista 1.
carro	barro	bolo	tôlo
pouco	côco	tonta	ponta
cola	gola	bebo	devo
dente	penete	pomba	bomba
tara	cara	bela	tela
<u>bata</u>	<u>pata</u>	<u>pulha</u>	<u>bulha</u>
gôto	côto	toda	bôda
capa	papa	dique	tique
pato	gato	tinto	pinto
cota	bota	ponto	conto
tacho	cacho	bufo	tufu
<u>baço</u>	<u>paço</u>	<u>galo</u>	<u>calo</u>
dura	cura	turra	burra
posse	tosse	cama	gama
bica	pica	teso	peso
palha	calha	pilha	bilha
gula	pula	couro	touro
<u>pala</u>	<u>bala</u>	<u>gana</u>	<u>cana</u>
puro	duro	pano	dano
gago	bago	barra	garra
baça	caça	ganga	canga
cunho	punho	dança	pança
pode	bode	gado	dado
dote	pote	doca	toca

ANEXO ED.3/1: Análise de variância ao número de respostas correctas, com grupo e ordem como variáveis intersujeito, e ouvido e tarefa como variáveis intrasujeito (Experiência I, 1ª Passagem)\*

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F
Grupo (Grp.)	2	252.32	11.54
Ordem (Ord.)	1	9.13	.42
Intra-Cela	40	21.87	-
Ouvido (Ouv.)	1	237.39	23.41
Tarefa (Tar.)	1	22.96	2.16
Ouv. x Tar.	1	.91	.18
Grp. x Tar.	2	15.50	1.46
Grp. x Ouv.	2	1.26	1.46
Ord. x Tar.	1	2.87	.27
Ord. x Grp.	2	24.73	1.13
Ord. x Ouv.	1	151.57	14.95
Grp. x Tar. x Ouv.	2	7.67	1.47
Grp. x Ouv. x Ord.	2	20.07	1.98
Ord. x Grp. x Tar.	2	18.33	3.73
Ord. x Ouv. x Tar.	1	7.44	1.42
Ord.x Grp.x Tar. x Ouv.	2	12.17	2.33

ANEXO ED.3/2: Análise de variância ao número de respostas correctas, com grupo e ordem como variáveis intersujeito, e ouvido e tarefa como variáveis intrasujeito (Experiência I, 2ª Passagem)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F
Grupo (Grp.)	2	90.47	5.08
Ordem (Ord.)	1	5.62	.32
Intra-Cela	34	17.80	-
Ouvido (Ouv.)	1	129.60	11.37
Tarefa (Tar.)	1	18.22	2.18
Ouv. x Tar.	1	1.22	.39
Grp. x Tar.	2	55.12	6.60
Grp. x Ouv.	2	12.43	1.09
Ord. x Tar.	1	6.40	.77
Ord. x Grp.	2	24.04	1.35
Ord. x Ouv.	1	172.22	15.10
Grp. x Tar. x Ouv.	2	8.22	2.64
Grp. x Ouv. x Ord.	2	15.81	1.39
Ord. x Ouv. x Tar.	2	8.22	6.64
Ord.x Grp.x Tar. x Ouv.	2	1.11	.36

\* O grau de pormenor com que serão apresentados os resultados das análises de variância variam consoante o programa utilizado. Recorremos aos programas de uso corrente na Laboratório de Psicologia Experimental da Universidade Livre de Bruxelas, nestas análises das Experiências I e II.

ANEXO ED.3/3: Análise de variância ao número de respostas correctas, com grupo e ordem como variáveis intersujeito, e ouvido e tipo de par (dissimilar vs similar) como variáveis intrasujeito.(Experiência II)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F
Grupo (Grp.)	2	285.13	37.2
Ordem (Ord.)	1	1.48	.19
Intra-Cela	68	7.66	-
Ouvido (Ouv.)	1	356.84	105.97
Par	1	854.76	276.52
Ouv. x Par	1	44.67	16.42
Grp. x Par	2	9.33	3.02
Grp. x Ouv.	2	.76	.23
Grp. x Par	2	9.33	3.02
Ord. x Grp.	2	8.59	1.12
Ord. x Ouv.	1	.76	.23
Grp. x Ouv. x Par	2	16.93	6.23
Grp. x Par x Ordem	2	10.68	3.46
Ord.x Grp.x Par. x Ouv.	2	.84	.31

ANEXO ED.3/4: Análise de variância aos valores dos score *f*, com grupo e ordem como variáveis intersujeito, tipo de par como variável intrasujeito.(Experiência II)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F
Grupo (Grp.)	2	1745.61	2.15
Ordem (Ord.)	1	.61	.0
Intra-Cela	68	11.51	-
Par	1	700.91	1.11
Grp. x Par	2	740.08	1.17
Ord. x Par	1	1942.78	3.08
Ord. x Grp.	2	526.81	.65
Ord. x Grp. x Par	2	13.71	.02

ANEXO ED.4: Teste de audição dicótica utilizado na Experiência II. T<sub>1</sub> e T<sub>2</sub> referem-se a ensaios de treino.

T <sub>1</sub> -DENTE	CARA	T <sub>2</sub> -puro	duro
- PATO	CURA	- cota	bota
- BURRA	PONTO	- TUFO	CAMA
- gado	dado	- CALO	TODA
- couro	touro	- PODE	BULHA
- pato	gato	- doca	toca
- CAPA	BOTA	- baça	caça
- COTA	PAÇO	- GANA	PESO
- gago	bago	-BOLO	DANO
- teso	peso	- ponto	conto
- PATA	COLA	- TOSSE	GATO
- DANÇA	TOCA	- baço	paço
- pulha	bulha	- cola	gola
- PURO	BELA	- tonta	ponta
- GARRA	CONTO	- BILHA	GAGO
- bufo	tufo	- capa	papa
- pano	dano	- turra	burra
- goto	côto	- GADO	COURO
- DURO	PALHA	- PONTA	BARRA
- CUNHO	DOTE	- pala	bala
- PALA	BODE	- GOTO	CACHO
- TIQUE	GALO	- PANO	TOLO
- pode	bode	- gana	cana
- tacho	cache	- pilha	bilha
- bela	tela	- PAPA	CÔTO
- BAÇA	TONTA	- GOLA	TACHO
- PENTE	CARRO	- BAGO	PUNHO
- bolo	tolo	- toda	boda
- DURA	BALA	- cama	gama
- ganga	canga	- posse	tosse
- tinto	pinto	- TARA	CÔCO
- dura	cura	- galo	calo
- PANÇA	TINTO	- BUFO	PINTO
- carro	barro	- TURRA	CAÇA
- palha	calha	- barra	garra
- BOMBA	PILHA	- pomba	bomba
- DOCA	CANGA	- TOURO	GAMA

**ANEXO EO.1:** Análise de variância das respostas correctas, com os factores intersujeito Grupo (Univ., Lic., Prim., Préprim.) e Sexo (M, F), e Intrasujeito Ouvido (E, D).\* (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	3	909.94	37.89	0
Sexo	1	161.33	6.71	.0111
Grupo x Sexo	3	11.77	0.49	.6898
Intra-Cela	88	24.01	-	-
( $R^2 = .581$ )				
Ouvido (Ouv.)	1	247.52	38.11	0.002 E-008
Ouvido x Grp.	3	3.29	0.50	0.6778
Ouvido x Sexo	1	9.18	1.41	0.2374
Ouv. x Grp. x Sexo	3	7.63	1.17	0.3238
Intra-Cela (Ouv.)	88	6.49	-	-
( $R^2 = .336$ )				

**ANEXO EO.2:** Análise de variância das proporções de erro na primeira consoante ( $C_1$ ), depois de submetidas à transformação arcoseno, em que  $x' = 2(\text{arcoseno } \sqrt{x})$ ; grupo e sexo são factores intersujeito, ouvido é factor intrasujeito. (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	3	4.25	15.85	0.530 E-008
Sexo	1	0.00	0.00	0.9264
Grupo x Sexo	3	0.20	0.74	0.5272
Intra-Cela	88	0.26	-	-
( $R^2 = .361$ )				
Ouvido	1	0.06	0.38	0.5384
Ouvido x Grp.	3	0.18	1.09	0.3544
Ouvido x Sexo	1	0.00	0.02	0.8889
Ouv. x Grp. x Sexo	3	0.30	1.75	0.1617
Intra-Cela (Ouv.)	88	0.17	-	-
( $R^2 = .092$ )				

\* Salvo se expressamente indicado, o modelo utilizado nesta Anova e nas que seguirão considera todos os efeitos principais e interacções possíveis. Os programas utilizados foram os do Statistical Package for the Social Sciences, SPSS, ou do Statistical Analysis Software, SAS.

Uma análise aos mesmos resultados expressos em proporção e submetidos à transformação arcoseno, em que  $x=2(\text{arcoseno } \sqrt{x})$ , revelou o mesmo padrão de efeitos significativos.

**ANEXO EQ.3:** Análise de variância das proporções de erros globais, *i.e.*, a soma dos erros  $C_1V_1$  e  $C_1V_1C_2(V_2)$ , submetidas à transformação arcoseno, em que  $x' = 2(\text{arcoseno } \sqrt{x})$ ; grupo e sexo são factores intersujeito, e ouvido é factor intrasujeito. (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	3	4.24	23.28	0
Sexo	1	0.70	3.88	0.0518
Grupo x Sexo	3	0.14	0.77	0.5138
Intra-Cela	88	0.18	-	-
				( $R^2 = .464$ )
Ouvido	1	0.09	0.45	0.5000
Ouvido x Grp.	3	0.10	0.49	0.6846
Ouvido x Sexo	1	0.00	0.03	0.8622
Ouv. x Grp. x Sexo	3	0.08	0.40	0.7521
Intra-Cela (Ouv.)	88	0.20	-	-
				( $R^2 = .035$ )

**ANEXO EQ.4:** Análise de variância das proporções de erro em pelo menos os três primeiros segmentos,  $C_1V_1C_2(V_2)$ , previamente transformados (arcoseno), com as variáveis intersujeito grupo e sexo e intrasujeito ouvido. (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	3	4.46	38.47	0
Sexo	1	0.70	6.08	0.0155
Grupo x Sexo	3	0.28	2.43	0.0701
Intra-Cela	88	24.01	-	-
				( $R^2 = 0.549$ )
Ouvido	1	0.37	3.87	0.0521
Ouvido x Grp.	3	0.13	1.37	0.2544
Ouvido x Sexo	1	0.29	3.02	0.0852
Ouv. x Grp. x Sexo	3	0.02	0.26	0.8537
Intra-Cela (Ouv.)	88	-	-	-
				( $R^2 = 0.118$ )

**ANEXO EQ.5:** Análise de variância das proporções de erro  $C_1$  (transformadas arcoseno), com grupo como variável intersujeito. (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo	2	.79	12.36	.0001
Intra-Cela	35	.065	-	-

ANEXO EO.6: Análise de variância das proporções de erros em função do tipo de contraste (transformadas arco-seno), com grupo como variável intersujeito e tipo de contraste como variável intrasujeito. (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	4	9.04	21.51	0.0001
Intra-Cela	115	0.42	-	-
Contraste (Cont.)	2	0.49	6.45	0.0019
Grp. x Cont.	8	0.13	1.72	0.09
Intra-Cela	230	0.07	-	-

ANEXO EO.7: Análise de variância à proporção de não-palavras (transformadas arco-seno), com grupo e sexo como variáveis intersujeito. (Dados ontogenéticos)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	3	0.95	1.91	0.114
Sexo	1	0.00	0.02	0.875
Grupo x Sexo	3	0.53	1.07	0.376
Intra-Cela	88	0.12	-	-

ANEXO EO.Blend: Controlo *a posteriori* de efeitos lexicais nos *blendings*. Lista de *blendings* e anómalas possíveis, conforme sejam palavras ou não-palavras (por questão de clareza assinala-se apenas as palavras, P).

a. Pares com atenção a sonoro (n=7)

Estímulos	<i>Blendings</i>	Anómalos
1. GATO-tosse	/dato/ /kato/ p	/bato/ /pato/ p
2. BOTA-capa	/gota/ p /pota/	/dota/ /tota/
3. BICA-calha	/gica/ p /pica/	/dica/ p /tica/
4. DOTE-cunho	/gote/ p /tote/	/bote/ p /pote/ p
5. GAMA-touro	/dama/ p /cama/ p	/bama/ /pama/
6. DEVO-pomba	/bevo/ (p) /tevo/	/gevo/ /kevo/
7. GALO-tique	/dalo/ p /calo/	/balo/ /palo/
Total	(Sonora) P=2(3) (Surda) P=4	P=2 P=2

## b. Pares com atenção a surdo (n = 7)

Estímulos	Blendings	Anómalos
1. POSSE-gula	/bossa/ p /coça/ p	/dossa/ /tossa/ (p)
2. PESO-gama	/beso/ /keso/	/deso/ /teso/ p
3. TACHO-gola	/dacho/ /cacho/ p	/bacho/ /pacho/ p
4. CARA-dente	/gara/ /tara/ p	/bara/ /para/ p
5. PALHA-dura	/balha/ /talha/ p	/galha/ /calha/ p
6. TONTA-bala	/dōta/ /pōta/ p	/gōta/ /cōta/ p
7. CANGA-doca	/gāga/ p /tāga/ p	/bāga/ /pāga/
Total	(sonora) P=2 (surda) P=6	P=0 P=5(6)

ANEXO IP.1: Lista das 40 palavras usadas na experiência de inteligibilidade, na ordem em que aparecem na condição de ruído binaural.

1 - CURA	21 - CACHO
2 - PONTO	22 - TELA
3 - DADO	23 - TONTA
4 - TOURO	24 - CARRO
5 - GATO	25 - TOLO
6 - BOTA	26 - CANGA
7 - PAÇO	27 - BALA
8 - BAGO	28 - PINTO
9 - PESO	29 - BARRO
10 - COLA	30 - TINTO
11 - TOCA	31 - PILHA
12 - BULHA	32 - TODA
13 - BELA	33 - GOLA
14 - CONTO	34 - PONTA
15 - TUFO	35 - PAPA
16 - DANO	36 - BODA
17 - PALHA	37 - TOSSE
18 - DOTE	38 - CALO
19 - GALO	39 - GARRA
20 - BODE	40 - POMBA

**ANEXO IP.2/1:** Análise de variância das respostas correctas com grupo e ordem como factores intersujeito e condição como factor intrasujeito. (Inteligibilidade de palavras)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo(Grp.)	2	267.79	14.58	0.0001
Ordem	1	399.31	21.75	0.0001
Grp. x Ordem	2	20.70	1.13	0.3283
Intra-Cela	90	18.36	-	-
Condição (Cond.)	2	16777.88	1795.45	0.0
Cond. x Grp.	4	42.81	4.58	0.0015
Cond. x Ordem	2	573.62	61.38	0.001
Cond. x Grp.P x Ordem	4	7.36	0.79	0.5342
Intra-Cela	180	9.34	-	-

**ANEXO IP.2/2:** Correlações entre idade e nível de desempenho (percentagem de respostas correctas) por condição e ordem, no grupo das iletradas (idade média 52.3, d.p. = 6.3).

	Monaural		R. Contral.		R. Binaural	
	O <sub>1</sub>	O <sub>2</sub>	O <sub>1</sub>	O <sub>2</sub>	O <sub>1</sub>	O <sub>2</sub>
<i>r</i>	-.12	-.24	-.08	-.42	-.16	-.04
<i>p</i>	.65	.29	.74	.06	.53	.85

**ANEXO IP.2/3:** Análise de variância das percentagens de decréscimo em ruído binaural (itens correctamente identificados em binaural/itens correctamente identificados em monaural), submetidas à transformação arcoseno com grupo e ordem como factores intersujeito. (Inteligibilidade de palavras)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	0.14	8.27	0.0005
Ordem	1	1.05	56.59	0.0001
Grp. x Ordem	2	0.36	0.31	0.737
Intra-Cela	90	0.01	-	-

$$R^2 = .45$$

**ANEXO IP.2/4:** Análise de variância das percentagens de decréscimo em ruído contralateral ruído contralateral/itens correctamente identificados em (itens correctamente identificados em monaural), submetidas à transformação arcoseno, com grupo e ordem como factores intersujeito. (Inteligibilidade de palavras)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	2.26	2.26	0.10
Ordem	1	50.61	0.67	0.43
Grp. x Ordem	2	2.72	2.72	0.07
Intra-Cela	90	-	-	-

$$R^2 = .10$$

ANEXO IP.2/5: Análise de variância às respostas correctas na condição de ruído binaural em subgrupos emparelhados segundo exactidão em monaural/controlo, com grupo e ordem como factores intersujeito. (Inteligibilidade de palavras)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	124.10	5.75	0.004
Ordem	1	1132.11	55.98	0.000
Grp. x Ordem	2	10.17	0.50	0.607
Intra-Cela	51	20.22	-	-

$R^2 = .57$

ANEXO IP.2/6: Análise de variância às respostas correctas na condição de ruído contralateral em subgrupos emparelhados segundo exactidão em monaural/controlo, com grupo e ordem como factores intersujeito. (Inteligibilidade de palavras)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	7.56	1.13	0.331
Ordem	1	8.21	1.23	0.273
Grp. x Ordem	2	19.98	2.98	0.059
Intra-Cela	51	6.70	-	-

$R^2 = .17$

ANEXO IP.2/7: Análise de variância ao número de dígitos correctamente repetidos (avaliação da Memória a Curto Termo) com grupo como variável intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	55.02	41.26	.0001
Intra-Cela	77	1.33	-	-
Tarefa	1	91.81	223.31	.0001
Tarefa x Grp.	2	.14	.35	.7091
Intra-Cela	77	.41	-	-

ANEXO IP.2/8: Correlações entre Memória a Curto Termo (score de repetição de dígitos em ordem directa) e desempenho na condição de ruído binaural (percentagem de respostas correctas), por grupo e ordem (Ordem 1: condições controlo, ruído contralateral e ruído binaural; ordem 2: sequência contrária à anterior). (Inteligibilidade de palavras)

	Letradas		Semiletradas		Iletradas	
	O <sub>1</sub>	O <sub>2</sub>	O <sub>1</sub>	O <sub>2</sub>	O <sub>1</sub>	O <sub>2</sub>
Directa						
r	-.19	-.05	.37	.27	.04	.48
p						.03
Inversa						
r	.02	-.10	.18	.20	.04	.44
p						.04

ANEXO IP.2/9: Análise de variância às proporções de erro (transformadas arcoseno) com grupo como factor intersujeito e tipo de erro (C1(V2) vs global mais silábico) como variável intrasujeito, considerando todos os sujeitos que fizeram a experiência na ordem I.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	.039	.60	.5495
Intra-Cela	41	.064	-	-
				$r^2 = .02$
Tipo Erro (T.E.)	1	5.12	44.45	24.87 E-008
Grp. x T.E.	2	.807	6.99	.0024
				$r^2 = .58$

ANEXO IP.2/10: Análise de variância à proporção de erros de tipo C<sub>1</sub>(V<sub>2</sub>) (transformadas arcoseno) com grupo como variável intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo	2	.351	3.93	.0274
Intra-Cela	41	.089	-	-

ANEXO IP.2/11: Análise de variância à proporção de erros de tipo silábico mais global (transformadas arcoseno) com grupo como variável intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo	2	.49	5.49	.0077
Intra-Cela	41	.09	-	-

ANEXO IP.2/12: Análise de variância às proporções de erro (transformadas arcoseno) com grupo como factor intersujeito e tipo de erro (C1(V2) vs global mais silábico) como variável intersujeito, considerando apenas os sujeitos com nível de desempenho equivalente (cerca de 42%), na ordem I.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	.03	.89	.4206
Intra-Cela	25	.03	-	-
				$r^2 = .06$
Tipo Erro	1	3.97	34.71	33.97 E-007
Grp. x T.E.	2	.40	3.55	.0438
Intra-Cela	25	.11	-	-
				$r^2 = .62$

ANEXO IP.2/13: Análise de variância à proporção de erros de tipo silábicos mais global (transformadas arco-seno) com grupo como variável intersujeitos com o mesmo nível de desempenho.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo	2	.32	4.14	.0278
Intra-Grupo	25	.07	-	-

ANEXO IP.2/14: Análise de variância à proporção de erros de tipo C<sub>1</sub>(V<sub>2</sub>) (transformadas arco-seno) com grupo como variável intersujeitos com o mesmo nível de desempenho.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo	2	.11	1.59	.2253
Intra-Grupo	25	.07	-	-

ANEXO IP.3: Análise por item dos estímulos cuja C<sub>2</sub> não é oclusiva

a. FRICATIVAS: [z,s,f,z] em "peso, paço, tosse, tufo, cacho".

Os estímulos são em geral correctamente identificados. No conjunto dos 5 estímulos e dos 3 grupos, houve erros em apenas 22.7% dos casos. Estes erros raramente envolviam a identificação incorrecta da segunda consoante: num total de 50, apenas 15 (30%) eram desse tipo. Desses 15, quase metade (7), consistia em dar como resposta outras fricativas.

b. LATERAIS [l, λ]

No conjunto das 11 palavras com [l] ou [λ] como segunda consoante, distinguem-se nitidamente dois subgrupos, um constituído por palavras facilmente identificáveis e outro muito pouco inteligível.

No primeiro, composto por "bela", "tela", "cola", "calo", "pilha", e "palha"; a percentagem de respostas erradas relativamente ao total de apresentações foi cerca de 24% (13% nas letradas, 26% nas semiletradas e 32% nas iletradas). Neste subgrupo de palavras, apenas para o estímulo "calo" se observam erros de segunda consoante, 5 vezes ouvido como "caldo" e uma vez como "cacho", e para os estímulos "pilha" e "palha", em que o [λ] foi substituído por [p].

O outro subgrupo de 5 estímulos, "bala", "galo", "gola", "tolo", e "bulha", teve uma elevada percentagem de respostas erradas: no conjunto das observações, cerca de 84% (81% nas letradas, 86% nas semiletradas e iletradas). Exceptuando o caso de "gola", em que os erros foram quase sempre "cola", nos restantes as palavras simplesmente não eram entendidas: 40% dos erros são de tipo "ininteligível". Os erros na segunda consoante apareceram quase exclusivamente em resposta aos itens "tolo" e "bulha". No primeiro caso, em nenhuma das identificações erradas apareceu a líquida [l]; no total das 22 respostas obtidas, 10 tinham [t] como C<sub>2</sub>, e 7 tinham [b]. Todavia, não nos parece muito prometedora uma análise mais detalhada a este nível devido ao eventual papel de factores lexicais. De facto, todos estes erros eram palavras com sentido, como "ponto", "tonto", tendo-se mesmo dado a nasalização das vogais para conseguir a aproximação semântica. São de referir as respostas dadas por duas iletradas que preservam o carácter líquido da segunda consoante: "coelho" e "tulho".

Quanto às identificações incorrectas do item "bulha", não emergiu nenhum padrão claro; num total de 12 erros, 1/4 consiste em palavras significativas com [R] como segunda consoante, 33% em palavras significativas com preservação de vogal e troca de consoantes, sendo a segunda [d]. As restantes 4 respostas consistiam em trocas de consoantes (a segunda duas vezes para [l], uma vez para [b] e outra para [t]) com manutenção de vogal posterior (em vez de [u], resposta com [o]); eram todas palavras significativas.

### c. VIBRANTE E BATIMENTO

Tínhamos três estímulos em que a segunda consoante era [R], "barro", "garra" e "carro". Em todos eles a percentagem de erros foi elevada, cerca de 92% considerando o total das observações (3 itens x 44 sujeitos); este valor era semelhante para os três grupos. O tipo de erro mais frequente consistia na troca da vibrante por fricativa, considerando aqui apenas as fricativas propriamente ditas [z,s,ʃ,f,v]. No total dos três estímulos e considerando todas as respostas erradas incluindo as "ininteligíveis", 47% dos erros caíam naquela categoria. A fricativa mais usada era o [z], ocorrendo em 46 vezes no total de 57 erros deste tipo. Este resultado não pode ser inteiramente atribuído a um constrangimento semântico, pois das cinco pseudo-palavras observadas, duas envolviam a transformação em fricativa, e as três restantes não. O número de pseudo-palavras era aproximadamente o mesmo neste tipo de erros e nos outros.

Os outros erros distribuíam-se pela categoria dos "ininteligíveis" (13%), e na posição de segunda consoante, preservação da vibrante (16%), transformação em oclusivas\* (7%) e outros. As transformações em [r] e [l] foram pouco frequentes, respectivamente apenas três e duas vezes. De notar a ocorrência de uma africada, [ʃk], em "casco" como resposta a "carro". Ocorreram também dois encontros consonânticos, em [karlu] de resposta a "garra", [baldru] de resposta a "barro", ambas palavras sem sentido referidas por iletradas.

Dois itens tinham o batimento [r] como segunda consoante, "touro" e "cura". A percentagem de erros para estes dois estímulos era elevada (cerca de 76% considerando todos os casos). O padrão de erros observados parece ser específico a cada estímulo. O erro mais comum na identificação de "touro" consiste na troca do [r] pelo [l]; esta transformação dá conta de 49% de todos os erros observados. Na identificação de "cura", na maioria dos erros é preservada a segunda consoante; quando isto não acontece, apenas em oito casos num total de 30, as substituições consistem na audição de [ɣ] (4 vezes) e [λ] (4 vezes). Estas transformações são acompanhadas pela manutenção da vogal anterior e levam o sujeito a ouvir em todos os casos excepto um palavras com sentido: "fuga" (2 letradas) "bulha" (uma semiletrada e uma iletrada) "pulha" (1 iletrada) e "gulha" (1 semiletrada).

### d. NASAL [n]

Apenas num item, "dano", aparece uma nasal. Trata-se também de um estímulo de difícil reconhecimento (36 respostas erradas num total de 44), em que a maioria erros são de tipo "ininteligível" (53%). Os restantes preservam em geral apenas as vogais, não sendo evidente nenhum padrão típico de respostas quanto às consoantes.

\* Consideramos aqui somente as oclusivas surdas e a nasal [m] pois as sonoras intervocálicas são realizadas como fricativas.

ANEXO C1/1: Análise de variância ao número de respostas sonoras para os *continua* velares, com grupo (estudantes vs letradas) como variável intersujeito, e estímulo e modo de construção (adição ou corte de pré-vozeamento) como variáveis intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	1	.10	.06	.7999
Intra-cela	27	1.59	-	-
Estímulo (Est.)	5		456.94	0
Est. x Grp.	5		1.14	.3669
Construção (Const.)	1	17.48	26.95	.0002
Const. x Grp.	1	1.17	1.80	.1904
Const. x Est.	5		4.22	.0071
Const. x Est. x Grp.	5		.3	.9081

ANEXO C.1/2: Análise de variância ao número de respostas sonoras para os *continua* labiais, com grupo (estudantes vs letradas) como variável intersujeito, e estímulo e modo de construção (adição ou corte de pré-vozeamento) como variáveis intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	1	1.40	.429	.5176
Intra-cela	128	3.26	-	-
Estímulo (Est.)	5		277.76	0
Est.x Grp.	5		.62	.6824
Construção (Cons.)	1	6.13	20.42	.0001
Const. x Grp.	1	1.36	4.54	.0418
Const. x Est.	5		3.11	.0262
Const. x Est. x Grp.	5		.86	.5203

ANEXO C.1/3: Análise de variância ao número de respostas sonoras para os *continua* labiais, com grupo (letradas, semiletradas e iletradas) como variável intersujeito, e estímulo e modo de construção (adição ou corte de pré-vozeamento) como variáveis intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	18.68	4.86	.0127
Intra-Cela	41	3.84		
Estímulo (Est.)	5		296.27	0
Est. x Grp.	10		2.04	.0403
Construção (Const.)	1	.03	.04	.8257
Const. x Grp.	2	.83	1.35	.2699
Const. x Est.	5		2.08	.0892
Const. x Est. x Grp.	10		1.08	.3860

**ANEXO C.1/4:** Análise de variância ao número de respostas sonoras para os *continua* velares, com grupo (letradas, semiletradas e iletradas) como variável intersujeito, e estímulo e modo de construção (adição ou corte de pré-vozeamento) como variáveis intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	1.54	.93	.4014
Intra-Cela	46	1.65	-	-
Estímulo (Est.) x Grp.	10	*	1.83	.0664
Construção (const.)	1	5.14	12.76	.0008
Const. x Grp.	2	3.03	7.52	.0014
Const. x Est.	5	*	5.55	.005
Cons. x Est. x Grp.	10	*	3.16	.0017

**ANEXO C.1/5:** Análise de variância ao número de respostas sonoras, para o *continuum* velar construído por adição de pré-vozeamento, com grupo (letradas, semiletradas e iletradas) como variável intersujeito e estímulo como variável intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	3.74	4.64	0.01459
Intra-Cela	46	0.81	-	-
			$r^2=168$	
Estímulo (Est.)	5	242.69	542.77	0.0
Est. x Grp.	10	1.56	2.99	0.0014

**ANEXO C.1/6:** Análise de variância ao número de respostas sonoras, para o *continuum* velar construído por corte de pré-vozeamento, com grupo (letradas, semiletradas e iletradas) como variável intersujeito e estímulo como variável intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	0.56	0.46	0.6357
Intra-Cela	46	1.22	-	-
Estímulo (Est.)	5	194.10	369.90	0.0
Est. x Grp.	10	1.56	2.99	0.001

**ANEXO C.1/7:** Análise de variância à diferença máxima entre dois estímulos consecutivos no *continuum* labial por adição de vozeamento, com grupo como variável intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Intergrupo	2	.002	.002	.9976
Intragrupo	31	1.02	-	-

ANEXO C.1/8: Análise de variância à diferença máxima entre dois estímulos consecutivos no *continuum* labial por corte de vozeamento, com grupo como variável intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Intergrupo	2	1.06	1.26	.3013
Intragrupo	22	.83	-	-

ANEXO C.1/9: Análise de variância à diferença máxima entre dois estímulos consecutivos no *continuum* velar por corte de vozeamento, com grupo como variável intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Intergrupo	2	.813	1.29	.2864
Intragrupo	38	.629	-	-

ANEXO C.1/10: Análise de variância à diferença máxima entre dois estímulos consecutivos no *continuum* velar por adição de vozeamento, com grupo como variável intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Intergrupo	2	2.53	.43	.0412
Intragrupo	44	.73	-	-

ANEXO FF.1. Conjunto de tríades usadas para constituir o material experimental.

- |     |         |        |       |
|-----|---------|--------|-------|
| 1.  | Clara,  | cara,  | lara  |
| 2.  | Plena,  | pena,  | lena  |
| 3.  | Pelar,  | par,   | lar   |
| 4.  | Feliz,  | fiz,   | lis   |
| 5.  | Prato,  | pato,  | rato  |
| 6.  | Pregar, | pegar, | regar |
| 7.  | Grama,  | gama,  | rama  |
| 8.  | Frita,  | fita,  | rita  |
| 9.  | Regato, | rato,  | gato  |
| 10. | Rebuço, | ruço,  | buço  |
| 11. | Traça,  | taça,  | raça  |
| 12. | Grita,  | guita, | rita  |
| 13. | Melaço, | maço,  | laço  |
| 14. | Fardo,  | fado,  | faro  |
| 15. | Falda,  | fada,  | fala  |
| 16. | Casta,  | cata,  | caixa |
| 17. | Rosto,  | roto,  | roxo. |

**ANEXO FF.1:** Análise de variância ao número de fusões para os seis pares de estímulos, "pena/lena", "pato/rato", "gama/rama", "cara/lara", "par/lar", "fiz/liz", com grupo como factor intersujeito e atenção como factor intrasujeito. (Fusões fonológicas)

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	44.14	8.81	0.003
Intra-Cela	81	5.00	-	-
			$r^2 = .17$	
Atenção (At.)	1	.07	.078	.7811
At. x Grp.	2	.39	.425	.6552
Intra-Cela	81	.93	-	-
			$r^2 = .01$	

**ANEXO FF.2:** Análise de variância ao número total de fusões para os pares anteriores, com ouvido como factor intrasujeito e grupo como factor intersujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	54.92	9.01	.0002
Intra-Cela	81	5.87	-	-
Ouvido (Ouv.)	1	.48	.58	.4474
At. x Grp.	2	2.0	2.41	.0955
Intra-Cela	81	.82	-	-

**ANEXO FF.3:** Análise ao número de fusões com grupo como factor intersujeito e tipo de par ("cluster" com líquida ou com vibrante ou com ortografia não-congruente) como variável intrasujeito.

Fonte Variação	Graus Liberdade	Quadrado Médio	Valor F	Probabilidade F
Grupo (Grp.)	2	70.48	9.23	.0002
Intra-Cela	81	7.63	-	-
Tipo Par	2	-	80.00	0
Grupo x Tipo Par	4	-	11.43	3.45 E-008