

INTRODUÇÃO

O atendimento fonoaudiológico de pacientes com deficiência auditiva inclui, em seus objetivos, a otimização da audição residual por meio de sistemas de amplificação e/ou estimulação elétrica, visando ao uso funcional da audição, principalmente no que se refere à percepção de fala. Autores como Ewing, Ewing (1971); Mencher (1976); Sanders (1977); Ling (1978,1989); Pollack (1985); Boothroyd (1982); Estabrooks (1994); Bevilacqua, Formigoni (1997); Mello, Novaes (2001), entre outros, argumentam que o aumento das possibilidades de receber o som de fala, por meio de recursos de amplificação, aumenta as chances de aquisição da linguagem falada, embora apenas esses recursos não sejam suficientes para eliminar as dificuldades em relação à linguagem de pessoas com deficiência auditiva.

O impacto das deficiências auditivas sobre a linguagem não pode ser considerado apenas do prisma da privação sensorial. As dificuldades de linguagem da criança surda em muito ultrapassam as condições fisiológicas de seu sistema auditivo, embora seja inegável que essas dificuldades tenham sua origem justamente na captação parcial e distorcida dos sinais acústicos. A fala concretiza-se no sinal acústico, que, sendo processado em níveis auditivo e cognitivo, possibilitará a compreensão do enunciado de um falante. Nesse sentido, as implicações das alterações auditivas sobre a percepção de fala e, conseqüentemente, sobre a sua produção são muitas, embora se observem variações individuais nessas habilidades, mesmo entre sujeitos com grau de deficiência auditiva semelhante.

A pesquisa a respeito dos efeitos das alterações auditivas sobre a percepção de fala no nível fonético indica que, em geral, essa diminui com o aumento do grau da perda, embora não possa ser prevista com exatidão pelo audiograma de tom puro (Boothroyd,1982). O indivíduo com alterações auditivas, principalmente aquelas de grau severo ou profundo, também fará uso de outros sentidos (visão e tato) para melhorar suas condições de perceber e produzir a fala, apoiando-se mais ou menos neles, de acordo com a natureza de sua alteração auditiva e as

condições acústicas do ambiente.

Apesar de existir, seguramente, uma forte relação entre os limiares audiométricos e as habilidades de percepção da fala, sabemos que a configuração do audiograma não é suficiente para prever as habilidades de percepção auditiva da fala (Sammeth et al, 1996), uma vez que reflete apenas uma medida de audibilidade dos sons nas várias frequências. A percepção auditiva da fala com uso de amplificação não pode ser prevista apenas com base no audiograma, pois as deficiências auditivas podem interferir também na resolução temporal, de frequência e de intensidade (Dorman, Hannley, 1985).

É reconhecida a importância da avaliação da percepção auditiva para a fala, embora os resultados dos testes disponíveis também não ofereçam uma indicação exata das habilidades de compreensão da língua falada pela pessoa com deficiência auditiva (Penrod, 1994).

O conhecimento sobre os processos de produção e percepção da fala, em pacientes com perdas auditivas, sem e com o auxílio de amplificação e, no decorrer do processo terapêutico, pode, sem dúvida, trazer subsídios para a clínica fonoaudiológica.

Crianças com perdas de audição congênitas ou adquiridas no início de suas vidas têm grande dificuldade em adquirir linguagem oral, porque percebem a fala do outro de forma muito distorcida e entrecortada, o que interfere no monitoramento do som de suas produções vocais. Nesses casos, o trabalho clínico fonoaudiológico, voltado para a aquisição da linguagem oral, enfatiza a orientação para o evento acústico como um potencial fator facilitador desse processo. Segundo Balieiro, Ficker (1997), no trabalho terapêutico, grande ênfase é dada ao uso da audição. A situação terapêutica permeada por um “jogo”, do qual “ouvir” faz parte, possibilita que a criança volte sua atenção para os sons do ambiente e da fala e, junto com o terapeuta, chegue a partilhar seus sentidos e significados. O que se espera é que o “jogo” em sua totalidade e não apenas o som em si seja enfatizado, sendo o aspecto relevante a relação dialógica estabelecida entre terapeuta e criança, na qual se introduz a dimensão sonora.

A informação auditiva (disponibilizada mediante recursos tecnológicos) irá contribuir para o processo de percepção da fala em diferente medida, de acordo com as habilidades auditivas de cada indivíduo e as habilidades de percepção visual e tátil das informações disponíveis para essas modalidades sensoriais, seja a partir do próprio falante, seja mediada por recursos tecnológicos, tais como vibradores táteis e *softwares* que tornam visíveis características acústicas dos sons de fala. Quando a deficiência de audição instala-se mais tardiamente em indivíduos que já adquiriram linguagem, o processo terapêutico enfocará, principalmente, as estratégias de comunicação, visando a melhorar as possibilidades de percepção de fala, por meio de pistas visuais, táteis e auditivas, bem como manter os padrões de produção de fala, que podem deteriorar-se sem a presença do *feedback* auditivo.

O conhecimento das características acústicas dos sons de fala é fundamental para a compreensão dos processos de produção e percepção de fala e, portanto, do impacto da perda auditiva sobre esses processos.

O interesse dos cientistas da fala na ampliação da percepção por parte de indivíduos com deficiência auditiva não é recente. Em 1945, Ralph Potter e seus colaboradores (citado por Baken, Daniloﬀ, 1991¹) publicaram um artigo sobre um novo instrumento que permitiria a visualização de padrões sonoros. Para os autores, a principal aplicação da espectrografia seria na comunicação de surdos, uma vez que poderia, um dia, possibilitar que sons fossem “ouvidos” visualmente. Assim, o surdo poderia aprender a falar ou melhorar sua produção de fala.

O uso da espectrografia com a finalidade de auxiliar a comunicação de surdos logo mostrou-se mais complexo do que imaginado, por não haver uma correspondência direta e evidente entre os aspectos articulatórios, acústicos e auditivos. Entretanto, ao possibilitar o desenvolvimento e aprofundamento dos estudos sobre a fala em condições normais, a espectrografia tornou-se um recurso de investigação extremamente relevante para a compreensão dos fenômenos de

¹ Potter RK et al. Visible patterns of sound (1945) apud Baken RS, Daniloﬀ RG. Readings in clinical spectrography of speech, 1991.

produção e percepção de fala. O conhecimento obtido por esse tipo de análise tem impulsionado o desenvolvimento da tecnologia visando à reabilitação do deficiente auditivo.

Poucos estudos sobre percepção e produção de fala de indivíduos com deficiência auditiva consideram os parâmetros fonético-acústicos do português brasileiro (doravante designado PB). Este trabalho visa a trazer uma contribuição nesse sentido, elegendo para análise os sons plosivos do PB.

O objetivo desta pesquisa é estudar a produção e percepção das consoantes plosivas do PB por um sujeito com deficiência auditiva. Para a sua consecução, será realizada a análise acústica de um *corpus* contendo as seis consoantes plosivas do PB e a avaliação da percepção auditiva para esses sons. Com a finalidade de obtermos referências para as medidas acústicas serão também analisadas as produções de um sujeito com audição normal, uma vez que não existem estudos que façam uma caracterização fonético-acústica desses sons no PB, que compreenda todos os parâmetros acústicos aqui considerados.

A escolha das consoantes plosivas deve-se à sua alta freqüência na língua portuguesa (Albano, 1995) e, portanto, à sua grande importância para a sua inteligibilidade. Cabe observar que se trata de consoantes encontradas em todas as línguas estudadas no mundo (Ladefoged, Madieson, 1996). Além disso, as consoantes plosivas são sons de fala com alto índice de alterações nas produções de sujeitos com deficiência auditiva, tanto no que se refere ao contraste de vozeamento, quanto ao ponto de articulação (Ochs et al, 1989; Tsui, Ciocca, 2000).

A análise acústica das produções das plosivas de um sujeito com deficiência auditiva e de um outro com audição normal apoiada na teoria acústica de produção de fala (Fant, 1960) possibilitará inferir alguns aspectos dos movimentos articulatórios dos sujeitos estudados.

Nossa intenção não é fazer considerações gerais sobre a produção das plosivas do PB, mas sim procurar compreender melhor como esse processo se revelou na fala desse sujeito com deficiência auditiva em particular, buscando também considerar como a produção desses sons de fala afeta sua percepção e é

afetada por ela. Cabe ressaltar que o objetivo deste trabalho não é fazer um estudo caracterizando a produção das plosivas do PB, tampouco caracterizar a produção desses sons por deficientes auditivos. Para tanto, deve ser considerado um maior número de sujeitos e um *corpus* que contemple outros contextos fonéticos.

Os resultados serão considerados à luz de um modelo gestual (Browman, Goldstein, 1986, 1990 e 1992) buscando-se uma maior compreensão da dinâmica de produção dos sons plosivos.

Por ser uma abordagem relativamente nova, o modelo gestual ainda não foi amplamente adotado para analisar questões clínicas relacionadas a alterações de fala (Kent, 1997a). Entretanto, a premissa fundamental da Fonologia Articulatória, que propõe que o sistema fonológico possa estar apoiado diretamente na articulação e não em traços distintivos, permite que as alterações encontradas na fala possam ser analisadas dinamicamente em termos da organização dos gestos articulatórios, o que não era possível numa análise que tinha como unidade, por exemplo, o fonema ou o traço distintivo. Nesses tipos de análise, a descrição das alterações restringe-se a considerações sobre substituições de fonemas ou traços, o que não traz muita contribuição para a compreensão da natureza das alterações da fala patológica.

Os dados da análise acústica serão comparados aos resultados de diversas tarefas de percepção auditiva das plosivas do PB, para podermos fazer algumas considerações sobre as relações entre a percepção e produção de fala nesse sujeito em particular.

No capítulo 1, que trata da revisão da bibliografia, iniciamos com algumas considerações sobre a relação entre a percepção e a produção de fala. Em seguida, apresentamos uma breve síntese da Fonologia Articulatória e da Teoria Acústica de produção de fala, enfocando alguns aspectos relevantes à produção das plosivas, tratando separadamente dos parâmetros acústicos relacionados ao contraste de vozeamento e de ponto de articulação. Por fim, são apresentados estudos que consideram a produção e percepção das plosivas por sujeitos com deficiência auditiva.

Em seguida, o capítulo 2, destina-se à descrição de um procedimento para a avaliação das habilidades de discriminação auditiva dos contrastes de sonoridade e ponto de articulação das plosivas do PB, o qual foi elaborado especificamente para a avaliação do sujeito com deficiência auditiva participante desta pesquisa.

O capítulo 3, que trata da metodologia, foi organizado a partir da descrição da elaboração e gravação do *corpus*, seleção dos sujeitos, realização de quatro diferentes tarefas de percepção auditiva das plosivas e dos procedimentos utilizados para a realização da análise acústica.

Os resultados das tarefas de percepção auditiva e das medidas acústicas estão apresentados no capítulo que se segue (capítulo 4). Inicialmente estão descritos os resultados das quatro tarefas de percepção auditiva das plosivas, duas realizadas por sujeitos com audição normal e duas pelo sujeito com deficiência auditiva. Em seguida, os dados da análise acústica da produção das plosivas do PB são expostos em tabelas e gráficos ilustrativos, acompanhados dos resultados do tratamento estatístico. A primeira medida acústica apresentada refere-se ao tempo de início do vozeamento (*Voice Onset Time – VOT*), seguida da frequência no início do primeiro formante da vogal subsequente à plosiva (F_1 no *onset* da vogal), da frequência fundamental no início da vogal subsequente (f_0 no *onset* da vogal). Posteriormente, são apresentadas as medidas de duração das consoantes plosivas e dos segmentos adjacentes, as medidas de energia da consoante em relação à vogal subsequente, os valores de frequência dos formantes das vogais adjacentes e os dados referentes à transição dos formantes em direção à vogal subsequente.

No capítulo 5 os resultados da análise acústica são discutidos à luz da Fonologia Articulatória e da Teoria Acústica de produção de fala, e são consideradas as relações entre esses dados e os resultados das tarefas de percepção auditiva para os sons plosivos. As considerações finais são apresentadas na sequência.

1 REVISÃO DA LITERATURA

1.1 Percepção e produção de fala

Quando nos deparamos com as conseqüências da deficiência de audição sobre o desenvolvimento da linguagem oral, não restam dúvidas de que a percepção auditiva dos sons da fala exerce papel fundamental sobre sua produção. Entretanto, a natureza dessa relação ainda não é totalmente compreendida. Várias teorias procuram explicar a percepção de fala e um dos maiores desafios que enfrentam é a questão da normalização do falante ou a busca da invariância, pois as características acústicas dos sons produzidos por diversos falantes — mulheres, crianças e homens — diferenciam-se em conseqüência das diferenças físicas das pregas vocais e do comprimento do trato vocal, bem como dos efeitos da coarticulação. Entretanto, os sons são reconhecidos por ouvintes como tais, apesar dessas variações. Segundo Hayward (2000), a questão da invariância, ou seja, de como os ouvintes relacionam sinais acústicos tão variáveis com unidades lingüísticas abstratas e invariantes, parece de mais difícil solução atualmente do que há cinquenta anos quando as pesquisas em busca das pistas acústicas iniciaram-se.

Whalen (1999) enfatiza que, embora exista uma incontestável relação entre a produção e a percepção de fala, a maioria dos estudos encontrados privilegia apenas um desses aspectos. Segundo o autor, as teorias sobre percepção de fala seguem três principais tendências: de base auditiva, de base motora e outra que não considera necessária a relação entre percepção e produção de fala, muito utilizada em trabalhos de reconhecimento automático de fala, não será aqui considerada.

A relação entre percepção e produção de fala é abordada na proposta da teoria quântica desenvolvida por Stevens (1972; 1997), que contempla as relações entre a configuração articulatória e os parâmetros acústicos e a maneira como esses parâmetros são percebidos auditivamente, estabelecendo uma estreita relação entre bandas críticas “invariantes”, dentro das quais variações de

configurações articulatórias não determinam variações de parâmetros acústicos. Tais variações só ocorrem nas fronteiras entre essas bandas, quando uma pequena variação articulatória resultaria em variação do parâmetro acústico, num salto quântico. A teoria quântica explicaria a capacidade de identificar fonemas, apesar das variações inter e intra falantes, uma vez que essas variações estariam ocorrendo dentro dos limites dessas bandas. O sistema auditivo estaria altamente afinado com essas bandas críticas, organizado em bandas equivalentes, de células ciliadas com a mesma especificidade (Stevens, 1978; Syrdal, Gopal, 1986).

Outros autores advogam que ocorre uma adaptação do falante em relação às demandas auditivas, ou seja, a partir de certos requisitos auditivos o falante ajusta seus articuladores para atingir esses efeitos. Segundo a teoria da variabilidade adaptativa de Lindblom (1990) a fala não precisa ser invariante. Sua função fundamental é a comunicação eficiente e, na produção de fala, sempre aplica-se o princípio da economia (princípio comum a outras funções biológicas). Assim, a forma acústica da fala, ao contrário de invariante, sempre será a resultante da interação entre a demanda de comunicação eficiente (determinada pelo ouvinte e condições do ambiente) e a economia de fala (inerente ao falante). Variabilidade adaptativa significa que o falante é capaz de adaptar sua produção às demandas da situação, produzindo fala de hiper a hipo articulada, num contínuo de adaptações entre estes extremos. Na “fala clara”, ou hiperarticulada, o falante não só aumenta a intensidade de sua fala, mas realiza reorganizações articulatórias. Nessa teoria, a percepção auditiva é uma função relativamente simples e a tarefa do mecanismo de produção é assegurar que os alvos acústicos sejam alcançados.

Há também autores que partem do pressuposto de que o ouvinte percebe ações significativas do trato vocal (gestos), uma vez que essas são sinalizadas por suas conseqüências acústicas (e, às vezes, visuais e táteis). A teoria motora de percepção de fala (Lieberman, Mattingly, 1996²) faz uma ligação explícita e direta entre a produção e a percepção. Sua versão revisada afirma que o objeto da percepção de fala são as intenções de gestos fonéticos dos falantes representadas no cérebro como comandos motores invariantes que eliciam os movimentos

² Trabalho originalmente publicado em 1985 – Lieberman AM, Mattingly IG. The motor theory of speech perception revised. *Cognition*, 21, 1985 p. 1-36.

articulatórios através de configurações lingüisticamente significativas. Nesse caso, não existe mapeamento entre percepção e produção porque as unidades envolvidas são as mesmas.

A não-diferenciação entre as unidades envolvidas na percepção e na produção de fala também é defendida por Fowler (1991) que propôs a teoria do realismo direto, uma evolução da teoria motora. Segundo o realismo direto, a percepção é mais bem estudada a partir do exterior do sujeito que percebe, ou seja, deve-se estudar a natureza do estímulo que age sobre o organismo e a maneira como esse reage. Argumenta-se que os órgãos sensoriais recuperam as informações sobre o objeto distal, a partir de informações contidas no estímulo proximal. Assim, o que se percebe é o próprio objeto e não a imagem do objeto. O realismo direto propõe um sistema no qual os gestos de fala são especificados diretamente pelo sinal de fala, que é vista como um entre outros eventos do mundo que são especificados pelo sinal acústico que originam. As estruturas fonológicas têm de ser percebidas pelo ouvinte como ações lingüisticamente significantes do trato vocal. Para tanto o ouvinte se apóia na estrutura do sinal acústico e por vezes visual, não como estrutura a ser percebida em si mesma, mas como informação de sua fonte causal, ou seja, das ações do trato vocal.

Para Fowler (1991), o que se percebe são os movimentos dos articuladores, e a informação sobre esses movimentos é recuperada pelo ouvinte a partir do sinal acústico. A autora, em outro trabalho, (Fowler, 1996), argumenta em favor da sua teoria (que propõe que os primitivos da percepção de fala são os gestos articulatórios), respondendo a críticas de Ohala (1996) e Sussman, (1989) ao realismo direto e à teoria motora. Esses autores argumentam que a motivação dessas teorias é a “ainda não encontrada invariância no sinal acústico”. Nesse mesmo sentido, O’Shaughnessy (1996) afirma que não há evidências fortes de que os ouvintes extraem características articulatórias do sinal acústico e que o fato de pesquisadores não terem ainda encontrado invariâncias no sinal acústico provavelmente deve-se à precariedade dos recursos tecnológicos de análise acústica da fala se comparados à capacidade do cérebro em termos de organização neural. A favor do realismo direto, Fowler (1996) cita algumas evidências experimentais de que os gestos são percebidos. Entre elas destaca o

efeito McGurk, que se refere à integração visual e auditiva na percepção de estímulos auditivos e visuais conflitantes, resultando na percepção de outro som: quando a sílaba /ba/ é apresentada auditivamente em sincronia com a apresentação visual da produção da sílaba /ga/, os ouvintes relatam perceber o som como /da/. Esse fenômeno é tomado como evidência de que a percepção de fala se dá a partir do gesto articulatório.

Albano (2001) questiona a hipótese de Ohala (1995) de que o traço é mapeado diretamente na acústica e na articulação, sendo a informação processada em um só estágio, de maneira rápida e econômica, mesmo considerando relações acústico-articulatórias não lineares. Albano, que propõe um modelo de Fonologia Acústico-Articulatória do PB ressalta:

“A viabilidade de um mapeamento neutro do traço é posta em questão pelo fato de a coarticulação tornar o sinal de fala mais opaco em termos acústicos do que em termos articulatórios. Mapear o ponto de articulação, por exemplo, em eventos acústicos tais como transições entre consoante e vogal, estouros ou ruídos só é possível assumindo-se uma certa sobreposição temporal das unidades fônicas cuja origem é inegavelmente articulatória.”(Albano, 2001, p. 41).

O modelo proposto avança em relação à Fonologia Gestual de Browman, Goldstein (1986, 1990, 1992), porque traz para o modelo a importância do acústico como elemento que impõe limites ao gesto articulatório. Sobre esse aspecto Browman, Goldstein (1986, 1990 e 1992) apenas citam a teoria quântica de Stevens (1972 e 1997) sem, entretanto, discutir como ela poderia ser integrada ao modelo que desenvolvem. Albano, retomando uma expressão sua (Albano, 1990) diz que:

“[...]’tocar a fala de ouvido’ é ajustar o período de gestos fônicos que somos naturalmente aptos a fazer desde meados do primeiro ano de vida à orquestração da linguagem e das ações que nos cercam. [...] a evolução nos fez os únicos primatas capazes de ouvir gestos que podem, de maneiras várias, apontar para outros gestos, inaudíveis mas preñes de significação.” (Albano, 2001, p. 179).

1.2 Fonologia Articulatória

Apresentamos uma breve síntese da proposta da Fonologia Articulatória (doravante FAR), apontando algumas vantagens do modelo para a análise das alterações de fala decorrentes de uma patologia como a deficiência auditiva.

Tradicionalmente, os pesquisadores que se dedicam ao estudo da fala, consideram a existência de duas estruturas: uma física e outra cognitiva, e a abordagem da relação entre as duas pressupõe a necessidade de uma “tradução” de um domínio (cognitivo) para outro (físico). A Fonética é a ciência que tem se ocupado das questões físicas da fala (as características articulatórias, acústicas e perceptivas dos sons das línguas), e a Fonologia das questões cognitivas (o estudo das funções dos sons da fala nas línguas do mundo).

A Fonologia Articulatória parte de um novo princípio, que considera os domínios fonético e fonológico como dois níveis de descrição de um único sistema complexo (Browman, Goldstein, 1986, 1990 e 1992): um nível mais elevado comprometido com a descrição do planejamento da emissão e um nível mais baixo, voltado para a descrição de sua execução³.

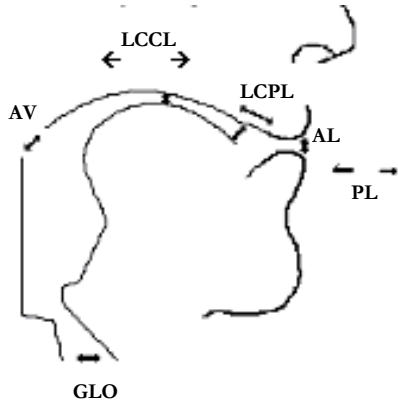
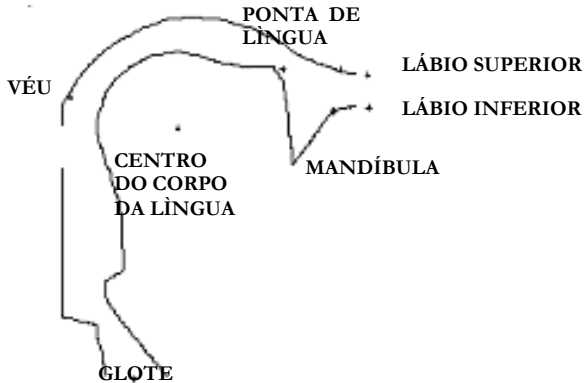
Segundo Tatham (1996), a FAR foi proposta por Browman, Goldstein (1986), em parte, como uma tentativa de unificar as descrições fonética e fonológica da produção de fala. Browman, Goldstein (1986) criticam a visão da Fonologia tradicional que adota a hipótese segmental propondo uma seqüência de feixes de traços que constituem os segmentos e determinam uma matriz de traços binários organizados em células que não se sobrepõem. Essa descrição, que enfatiza o aspecto estático da articulação, não permite capturar a variedade de ocorrências fonológicas e incorporar, na descrição das línguas, aspectos temporais que são específicos de cada uma. Para esses autores, era necessário desenvolver uma nova concepção que permitisse lidar com aspectos da dinâmica da fala (como, por exemplo, a sobreposição de gestos articulatórios) e que permitisse a emergência das relações temporais entre as estruturas articulatórias.

³ Para uma excelente revisão da FAR consultar D'Angelis (1998).

A FAR se distingue das outras fonologias por postular que a representação fonológica é feita em termos da organização articulatória em oposição aos traços ou outras unidades fonológicas comumente assumidas. Embora outras teorias fonológicas também façam referência a propriedades articulatórias, elas não consideram a articulação como a base do sistema fonológico. O ponto mais importante dessa abordagem é que, para a FAR, a unidade básica de contraste fonológico é o gesto, definido como uma caracterização abstrata do evento articulatório, cada um com um tempo (ou duração) intrínseco. Tal abordagem dinâmica da fala introduz a possibilidade de dar conta tanto do aspecto gradiente como do aspecto categórico da fala. Os enunciados são modelados como padrões organizados de gestos (ou constelações de gestos) nos quais as unidades gestuais podem sobrepor-se no tempo. As estruturas fonológicas definidas dessa maneira resultam num conjunto “natural” de classes articulatórias (Kent, 1997a).

Segundo Browman, Goldstein (1992), gestos são eventos que se desdobram durante a produção de fala e suas conseqüências podem ser observadas nos movimentos dos órgãos fonoarticulatórios. Esses eventos consistem na formação e soltura de constrições no trato vocal. Para ajudar a explicitar o modelamento desses eventos, os gestos são definidos em termos de tarefas dinâmicas (Saltzman, Kelso, 1987; Saltzman, Munhall, 1989). No caso da fala, a tarefa envolve a formação de várias constrições relevantes para que uma língua particular seja falada, as quais são expressas por equações que caracterizam o movimento. Um aspecto importante da tarefa dinâmica é que a mobilidade da variável do trato que é caracterizada dinamicamente, e não o movimento de um articulador individual.

Uma variável do trato caracteriza uma dimensão da constrição do trato vocal e os articuladores que contribuem para a formação e soltura da constrição são organizados numa estrutura coordenada. Por exemplo, a variável do trato - abertura de lábios - é afetada pela ação de três articuladores: lábio superior, lábio inferior e mandíbula. As variáveis do trato atuais e seus componentes articulatórios estão descritos na figura 1 (Browman, Goldstein (1990, p.344).

Variáveis do trato	Articuladores envolvidos
PL protrusão de lábios AL abertura de lábios LCPL local de constrição da ponta da língua GCPL grau de constrição da ponta da língua LCCL local de constrição do corpo da língua GCCL grau de constrição do corpo da língua AV abertura vélica GLO abertura glotal	lábio superior e inferior, mandíbula lábio superior e inferior, mandíbula ponta da língua, corpo da língua, mandíbula ponta da língua, corpo da língua, mandíbula corpo da língua, mandíbula corpo da língua, mandíbula véu palatino glote
	

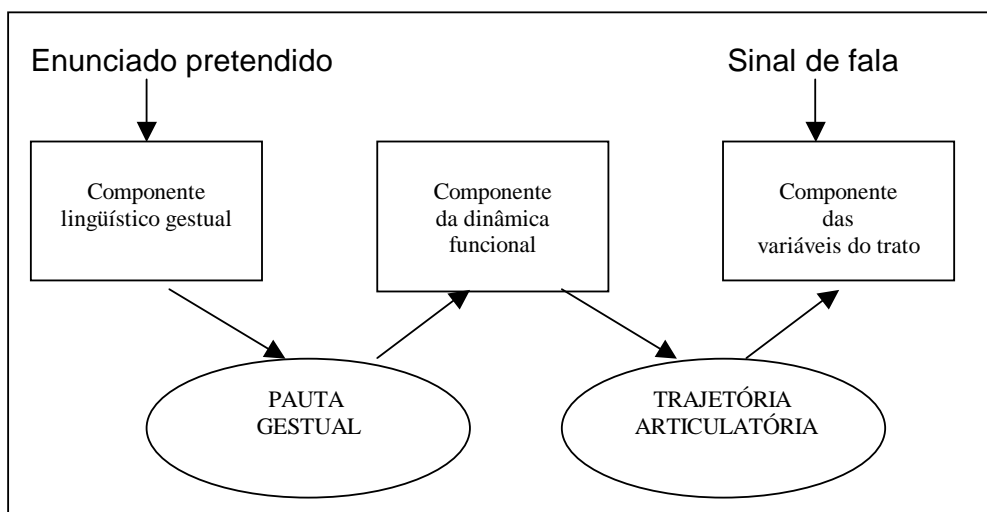
Fonte: Browman, Goldstein (1990, p. 344).

Figura 1 - Variáveis do trato e seus componentes articulatorios

O controle de uma variável é especificado pelo conjunto de articuladores usados para obter uma constrição, e os valores dos parâmetros na equação dinâmica descrevem seu movimento: alvo, rigidez e amortecimento. Esses parâmetros fornecem um tipo de estrutura interna para um regime de controle que subjaz ao evento espacial e temporal em todas as suas instâncias. Um gesto, em FAR, é especificado por um conjunto de variáveis do trato relacionadas. Por exemplo, nas constrições do trato oral, o local e o grau da constrição são duas dimensões da mesma constrição e, portanto, são consideradas variáveis do trato relacionadas. As variáveis do trato relacionadas são representadas pelas mesmas letras no final (L – lábios, PL – ponta da língua, CL – corpo da língua). Isto significa

que cada gesto é uma constrição local, definida com relação a um dos cinco conjuntos de variáveis do trato.

A figura 2 sintetiza o sistema computacional desenvolvido no *Haskins Laboratories* (Browman, Goldstein, 1990).



Fonte: Browman, Goldstein (1990, p. 342).

Figura 2 - Sistema computacional desenvolvido no *Haskins Laboratories*

Para a produção da emissão pretendida, no componente lingüístico gestual ocorre a composição da pauta gestual (escolha dos gestos que compõem o enunciado ou a palavra e a organização deles em ordem canônica). No componente da dinâmica funcional (*task dynamics*), ocorre a escolha das variáveis do trato para implementar os gestos da pauta; a geração das trajetórias articulatórias associadas a cada variável do trato se dá a partir da equação que descreve o movimento de oscilação de uma massa presa a uma mola. As trajetórias descritas mudam continuamente no tempo de acordo com os valores atribuídos aos parâmetros da equação e esses valores alimentam o modelo do trato vocal ou sintetizador articulatório, gerando o *output* acústico ou sinal de fala.

Kent (1997a) faz uma síntese do modelo funcional de produção de fala. A

teoria de sistemas dinâmicos conceitualiza a fala ou qualquer outro comportamento motor em termos da interação entre sistemas biomecânicos e variáveis do ambiente. Essa interação define um modelo não linear no qual ocorrem várias regiões operacionais estáveis, com características relativamente simples de controle. Uma das formas de expressar essa simplicidade é dizer que o sistema possui um grau de liberdade relativamente pequeno. A questão do grau de liberdade é o maior obstáculo para dar conta do controle do movimento. Mesmo o controle de movimentos relativamente simples pode ser um desafio, se todas as ações dos músculos e articulações individuais envolvidos tiverem de ser especificadas individualmente para serem realizadas. Entretanto, se vários músculos e articulações forem considerados como uma unidade, o seu controle passa a ser um problema menor. Em outras palavras, um sistema altamente complexo, quando considerado em seus componentes individuais passa a ser um sistema mais simples se os componentes estão funcionalmente ligados em unidades.

O problema de graus de liberdade será resolvido se considerarmos um sistema composto de agrupamentos funcionais (também chamados de *linkages*, sinergia ou estruturas coordenadas entre outras denominações). Na teoria de sistemas dinâmicos, a maior parte do problema do controle fica resolvida à medida que as estruturas que podem ser apropriadamente coordenadas são identificadas. Tais estruturas coordenadas têm uma tarefa específica, são sensíveis e adaptativas ao contexto e transformam-se num meio econômico e poderoso de dar conta do controle motor. Outra vantagem das estruturas coordenadas é que podem regular os parâmetros relacionados a mudanças na velocidade ou deslocamento total. Esse ajuste depende da distinção entre parâmetros essenciais e não essenciais de um movimento. Um parâmetro essencial define o aspecto qualitativo de um movimento e os não essenciais referem-se ao aspecto quantitativo. Um parâmetro essencial de uma plosiva bilabial é o fechamento dos lábios e um parâmetro não essencial seriam as possíveis diferenças no deslocamento do lábio inferior em direção ao fechamento, em consequência do contexto fonético, acento ou taxa de elocução. A interação entre esses parâmetros (essenciais e não essenciais) determina os ajustes do movimento. Os parâmetros essenciais estão

relacionados aos aspectos fonético-fonológicos que constituem os gestos articulatórios abstratos que induzem o movimento dos articuladores e os não essenciais às variações advindas dos efeitos do acento, taxa de elocução e outras variáveis escalares que modificam o movimento sob as demandas de restrições fonéticas.

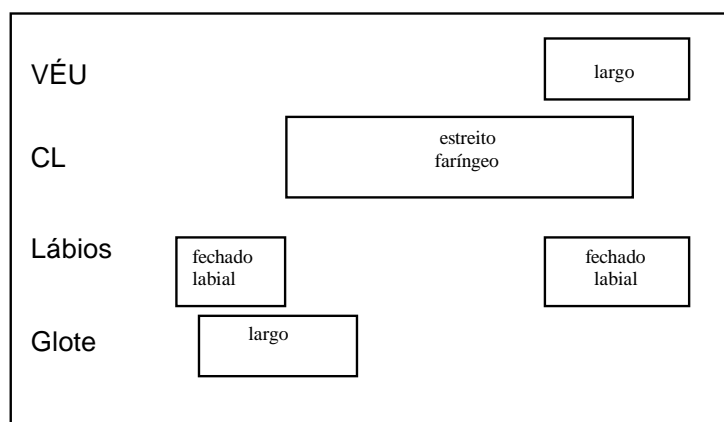
A relação entre os componentes de uma estrutura coordenada e a consequência motora dessa relação sinérgica é determinada pelas equações de restrição, as quais determinam a tarefa, ou seja, como os membros de um grupo podem interagir dentro dos limites de uma ação particular em dada circunstância ambiental. As equações permitem que a variação nas ações dos componentes seja baseada em princípios.

Browman, Goldstein (1992) definem as pautas gestuais como combinações de gestos que se sobrepõem. Segundo eles, caracterizar um gesto como um evento físico significa assumir que ele ocorre no tempo e no espaço. Considerando que o gesto tem um tempo de duração intrínseco, eles podem se sobrepor uns aos outros e, uma vez que gestos são eventos físicos, eles são afetados por processos físicos que ocorrem durante o ato de fala.

Os gestos que são empregados numa dada emissão são organizados ou coordenados em estruturas mais amplas, maiores. A organização formada por esses gestos particulares constitui a estrutura fonológica de tal emissão (ou, pelo menos, parte dessa estrutura). Nem todas as emissões de uma língua têm uma organização individual, já que há diversos princípios que definem como as classes de gestos são organizadas ou “faseadas”. Esses princípios captam o aspecto sintagmático da estrutura fonológica da língua, enquanto o inventário de gestos que podem participar dessa organização captura o aspecto paradigmático.

No componente lingüístico gestual do modelo computacional descrito, uma primeira aproximação desses princípios de “faseamento” é usada para coordenar os gestos, uns com os outros (Browman, Goldstein, 1990). Esse “faseamento” de gestos resulta numa estrutura chamada pauta gestual. As figuras 3 e 4 mostram as pautas gestuais da palavra *palm* e da palavra *pawn* em duas formas de

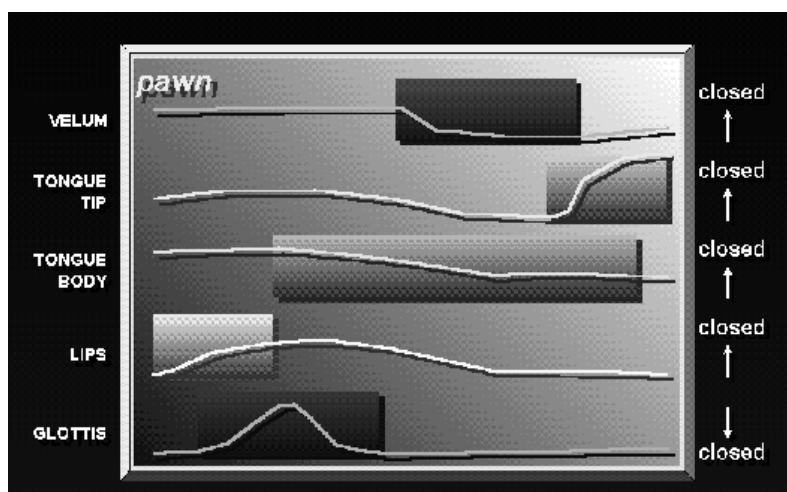
representação possíveis.



Fonte: Browman, Goldstein (1989 p. 212)

Figura 3 - Pauta gestual da palavra *palm*

Outra representação possível, na qual os blocos se referem aos gestos e o traçado ao movimento dos articuladores.



Fonte: Tatham (1996, p. 2)

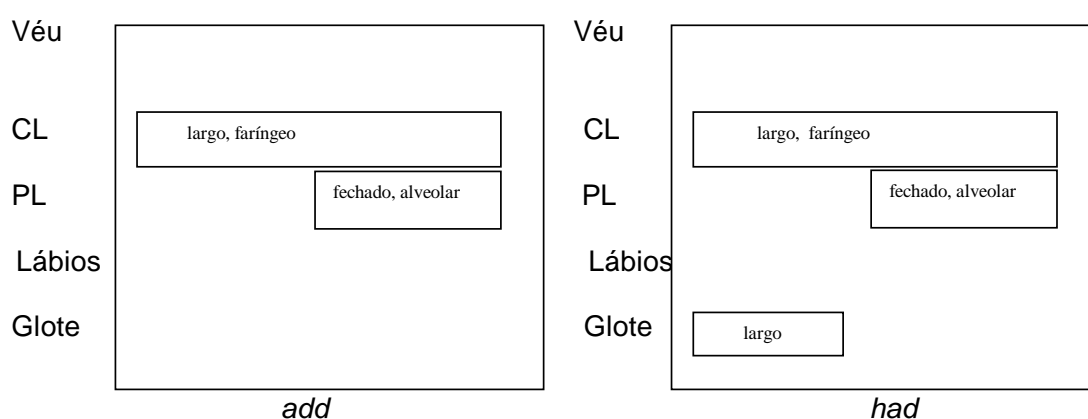
Figura 4 - Pauta gestual da palavra *pawn*

Essas representações mostram a duração dos gestos individuais assim como a sobreposição entre os gestos, que são organizados em “camadas articulatórias”.

A extensão horizontal dos blocos indica o intervalo discreto de tempo durante o qual um conjunto particular de valores dos parâmetros dinâmicos é ativado. A sobreposição significa que, num instante específico, uma série de diferentes gestos ou conjuntos de valores está afetando o trato vocal simultaneamente.

A riqueza da estrutura fonológica nessa proposta está nos padrões como os gestos estão coordenados no tempo, uns em relação aos outros, e o termo *constelação* ou *pauta* é utilizado para fazer referência a essa coordenação. Emissões que contêm os mesmos gestos podem contrastar umas com as outras em consequência da maneira como eles estão organizados, ou seja, os mesmos gestos podem formar diferentes constelações ou pautas.

Como o gesto é considerado a unidade de análise fonológica, dois itens lexicais irão contrastar se diferirem quanto à composição da pauta gestual. Essa diferença pode envolver a presença ou ausência de um dado gesto ou diferenças entre a organização de mesmos gestos. Os itens lexicais contrastam gestualmente, em primeiro lugar, se um dado gesto está presente ou ausente, como, por exemplo, em *add* x *had* (figura 5).



Fonte: de Browman, Godlstein (1992, p. 158)

Figura 5 - Pauta gestual esquemática de *add* e *had*

Numa análise tradicional fonológica, a diferença entre “*had*” em relação a “*add*” (exemplificado na figura 5) seria atribuída à presença de um segmento *a* mais. Já em *bad* x *pad*, o contraste estaria apenas em um traço, ou seja, vozeamento. Em termos gestuais, esses contrastes são carregados pela presença ou ausência de um gesto, ou seja, a presença do gesto glotal. Outro tipo de contraste provém de diferenças na pauta gestual quanto aos conjuntos de articuladores e variáveis do trato envolvidos na composição do gesto articulatorio, como, por exemplo, fechamento de lábios *versus* fechamento de ponta de língua (*bad* x *dad*).

Os gestos também podem diferir parametricamente, ou seja, nos valores dos parâmetros dinâmicos que definem a estrutura temporal e espacial do evento articulatorio, tais como valores-alvo para o grau de constrição da ponta da língua que poderia definir o fechamento total *versus* um valor crítico que levaria à geração de turbulência (diferindo as obstruintes plosivas das fricativas).

Além dos valores-alvos para o grau de constrição, outros parâmetros dinâmicos servem para distinguir os gestos (Browman, Goldstein, 1989,1990): o local da constrição; a rigidez (*stiffness*) o que provavelmente contrasta vogais de líquidas; o amortecimento, o que provavelmente distingue *flaps* de plosivas nas línguas que as contrastam.

Uma vez que tais diferenças não são inerentemente categóricas, Browman, Goldstein (1992) sugerem que diferentes gamas de valores dos parâmetros para um conjunto particular de articuladores são selecionadas, por cada língua, com base na relação quântica acústico-articulatoria proposta por Stevens (1989), e/ou com base nos princípios da dispersão adaptativa propostos por Lindblom (1990). O fato de que gestos são eventos com extensão temporal pode também eliminar a necessidade de certas restrições fonológicas por adjacência, as quais, por sua vez, podem ser compreendidas como consequência direta da sobreposição de gestos. Do ponto de vista da sobreposição de gestos, muitos casos de “assimilação”, ou aparente espalhamento de traços em função da coarticulação, são consequências diretas da co-ocorrência de vários gestos. Não há necessidade de espalhar um traço, uma vez que os gestos já têm uma extensão temporal inerente.

Quanto ao estilo geral de coordenação (ou “faseamento”) entre gestos Browman, Goldstein (1992) referem que esta coordenação pode variar de língua para língua.

Em geral, os proponentes da FAR referem-se a uma análise gestual como uma análise do *input* e a análises mais tradicionais como análise do *output*, onde *input* e *output* referem-se respectivamente, à descrição da organização gestual e o resultante formato/acústico do trato vocal como um todo. Na análise segmental tradicional, são feitas descrições dos efeitos combinados dos gestos, sendo tipicamente descrições acústicas, ou do *output* na terminologia da FAR. Mesmo descrições de traços freqüentemente se referem a atributos de segmentos e, portanto, são freqüentemente descrições de *output*. Essa é a origem das diferenças entre as descrições da abordagem gestual por um lado e, por outro, das descrições baseadas em segmentos ou traços.

Uma série de exemplos de análise gestual de casos que tradicionalmente foram analisados em termos segmentais e/ou de traços como diferentes tipos de variação alofônica pode ser utilizada para demonstrar que uma análise gestual captura uma gama maior de comportamentos e pode fazê-lo a usando um princípio geral, ao invés de regras especiais de mudança de categorias. A estrutura subjacente do *input* captura o contraste de uma forma muito simples. Citamos como exemplo o tratamento das variantes de /t/, [t] e [ɰ] no inglês americano. Conforme aponta Albano (2001, p.59) a variante [ɰ] resulta do encurtamento do gesto consonantal de ponta da língua e da aproximação dos gestos vocálicos, como se pode verificar na pauta gestual representada na figura 6.

nos seus parâmetros dinâmicos. O gesto consonantal caracteriza-se por um grau maior de constrição e uma constante de tempo mais curta (maior rigidez) do que ocorre com os gestos vocálicos. A organização silábica (unidades do tamanho de uma sílaba) é definida pelo faseamento de gestos (orais) consonantais e vocálicos uns em relação aos outros. A relação básica é que o *offset* das consoantes iniciais é coordenado com o *onset* do gesto vocálico e o *onset* das consoantes finais com o alvo do gesto vocálico. Isso resulta em organizações nas quais ocorre uma sobreposição temporal substancial entre os movimentos associados com os gestos vocálicos e consonantais na posição de coda silábica.

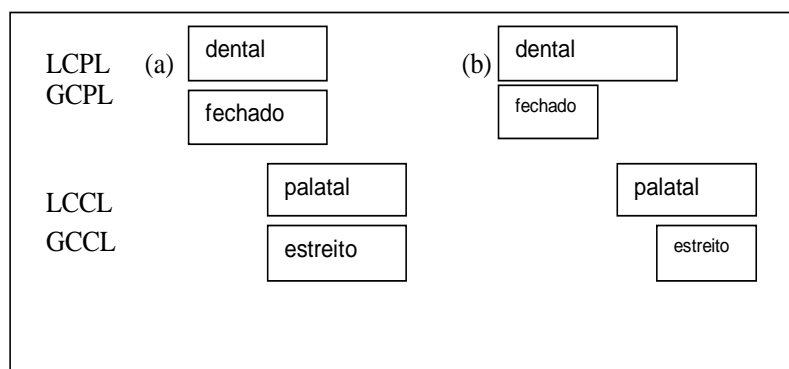
Quando o mesmo gesto consonantal é co-produzido com diferentes gestos vocálicos sobrepostos (como em [ada] x [idi]), o movimento do articulador produzido pelo modelo da tarefa dinâmica irá diferir, refletindo as demandas do gesto vocálico sobre os articuladores, que têm em comum a mesma consoante. A natureza da variação será diferente dependendo de os gestos que se sobrepõem estarem relacionados à mesma ou a diferentes variáveis do trato. No caso das variáveis serem distintas (como PL – ponta da língua para [d] e CL – corpo da língua para as vogais), o gesto consonantal irá atingir o seu alvo independentemente de qual vogal esteja se sobrepondo, embora a participação dos articuladores usados para atingir este alvo – mandíbula, corpo da língua e ponta da língua – seja diferenciada em função da vogal. Assim, o formato geral do trato vocal produzido durante o fechamento da ponta da língua será diferente para a produção de [ada] e [idi]. As trajetórias articulatórias diferentes irão produzir diferentes formantes acústicos de transição para as duas plosões, embora, aparentemente, o percepto da consoante não se modifique.

No caso em que as consoantes e vogais dividem a mesma variável do trato (CL – corpo da língua) - como a consoante [g] em [aga] ou [igi], os gestos consonantais e vocálicos não podem atingir simultaneamente seus alvos, uma vez que estão tentando mover exatamente a mesma estrutura para posições diferentes. Como resultado, o local (e não o grau) da constrição para a consoante vai variar em consequência da vogal que se sobrepõe. Nesse caso, entretanto, a diferença é perceptível, ao menos para os foneticistas, que representam a plosiva velar, algumas vezes, como distintos alofones “anteriores” e “posteriores”.

Os exemplos de sobreposição consoante/vogal ilustram dois pontos importantes sobre a estrutura gestual:

- mostram de que maneira os gestos, como unidades fonológicas específicas invariantes, podem originar trajetórias acústicas e articulatórias dependentes do contexto, sem que seja necessária a postulação de nenhuma regra de implementação para converter unidades específicas (fonológicas) invariantes em parâmetros (físicos) variáveis. As variações decorrem diretamente da definição das unidades como sistemas parametrizados por uma tarefa dinâmica, sua organização fonológica (padrão de sobreposição) e princípios gerais de como as unidades sobreponentes se fundem. A mesma estrutura gestual simultaneamente caracteriza propriedades fonológicas da emissão (unidades contrastivas e organização sintagmática) e propriedades físicas;
- sugerem como a mesma estrutura sintagmática (padrão de sobreposição) pode levar a diferentes tipos de variação (alofônica ou apenas articulatório-acústica) em função dos gestos que estão envolvidos, ou seja, se os gestos usam os mesmos ou diferentes articuladores.

Em relação ao tratamento da variação alofônica, Albano (2001,p.81) apresenta proposta que difere da FAR, introduzindo o que chamou de “hipótese da dissociabilidade da duração das trajetórias das variáveis do trato responsáveis pelo ponto e modo de articulação”. Como exemplo, podemos citar sua proposta para explicar as versões chiantes e não chiantes no que se refere à palatalização do [t] e [d] diante de [i] no PB, mostrada na figura 7:



Fonte: Albano (2001, p. 82).

Figura 7 - Pautas gestuais para as versões (a) não-chiante e (b) chiante da palatalização de [t, d] diante de [i], conforme a revisão da FAR proposta por Albano (2001)

Ambas as realizações compreendem as mesmas variáveis do trato, a saber: LCPL (dental) e GCPL (fechado); LCCL (palatal) e GCCL (estreito). A diferença entre eles ocorre em função do encurtamento da oclusão e do adiamento da constrição da vogal na versão chiante.

Farnetani, Recasens (2000) afirmam que a FAR representa, até o momento, a melhor tentativa de superar a distância entre os aspectos cognitivos e físicos da linguagem. Os gestos como unidades com um tempo intrínseco são passíveis de serem sobrepostos no tempo e, portanto, podem explicar as variações da coarticulação. A idéia geral da teoria da co-produção é que os gestos mesclam sua influência num articulador comum.

Por ser uma abordagem relativamente nova, o modelo gestual ainda não foi amplamente adotado para analisar questões clínicas relacionadas a alterações de fala. Entretanto, a premissa fundamental da FAR, que propõe que o sistema fonológico pode estar apoiado diretamente na articulação e não em traços, permite que as produções de fala alteradas possam ser analisadas dinamicamente em relação à organização dos gestos articulatorios, o que não era possível fazer numa análise que tinha como unidade o traço distintivo (Weismer et al, 1995). Antes, o

máximo que se fazia era a descrição das alterações de fala como uma substituição de fonemas ou traços, e isso pouco contribuía para a compreensão da natureza das alterações da fala patológica e para o fornecimento de subsídios para a implementação de estratégias terapêuticas no trabalho de fala.

Tomando a FAR como referencial teórico para a análise das alterações de fala, Kent (1997a) discute como o aspecto dinâmico da organização dos gestos pode explicar as alterações de fala de sujeitos com patologias de fala decorrentes alterações neurológicas. Destaca os seguintes aspectos encontrados na fala de sujeitos com apraxia:

- A lentificação do movimento articulatorio - para tais falantes, que lentificam o movimento articulatorio, as especificações temporais e espaciais de tais gestos seriam inerentemente desviantes da fala normal;
- O escalonamento desviante - pode ser definido como o ajuste da magnitude dos deslocamentos articulatorios em função do tempo e, em geral, apresenta-se alterado em pacientes com alterações neurológicas;
- Alterações no faseamento de gestos - podem provocar escorregamentos /deslizamentos de gestos em relação a outros na mesma pauta e, conseqüentemente, dar origem a alterações no percepto fonético. As alterações de faseamento podem responder pelas alterações de coordenação de gestos que resultam em distorções ou substituições do som de fala;
- Maior variabilidade na produção de fala por sujeitos com patologias neurológicas – decorrentes de falhas na organização temporal dos movimentos articulatorios poderiam ser explicadas por uma dificuldade no faseamento de gestos no modelo dinâmico de produção de fala;
- Segmentalização da fala - na fala patológica em função da diminuição da coarticulação a fala parece ser produzida de forma segmentada, não se apresentando concatenada. A medida da frequência de F2 no início da transição serve como um parâmetro acústico do grau de segmentalização da fala;

- Adaptação restrita a variações escalonares - variações no escalonamento impostas por padrões de acento, taxa de elocução em geral estão prejudicadas na fala patológica. Assim, o aumento da magnitude dos movimentos articulatórios esperado em uma sílaba acentuada pode ocorrer de forma alterada em pacientes com alterações de fala;
- Alteração da pauta gestual - para explicar alterações na fala de pacientes apráxicos, relacionadas às dificuldades da organização temporal na realização dos movimentos articulatórios, os autores sugerem uma deteriorização da pauta gestual acompanhando a disfunção de movimentos específicos.

Kent (1997a) e Weismer et al (1995) discutem as alterações encontradas na fala desses sujeitos a partir da realização de análise acústica de suas produções. Os dados da análise acústica da produção de fala interpretados à luz da Teoria Acústica de produção de Fala permitem, no modelo Fonte e Filtro, fazer inferências sobre as configurações do trato vocal na produção dos diversos segmentos de fala. Abordaremos a seguir alguns aspectos da Teoria Acústica de Produção de Fala, particularmente no que se refere à produção das plosivas.

1.3 Teoria Acústica da Produção de Fala – Fonte e Filtro

A Teoria Acústica da Produção de Fala (Fant, 1960), também chamada Teoria Fonte-Filtro linear (fundamentada em um modelo matemático linear), estabelece que na produção de fala a frequência produzida na glote (fonte) é modificada no trato vocal, que funciona como um filtro, aumentando ou diminuindo a amplitude de determinadas frequências⁴.

A vibração das pregas vocais produz um som que possui um espectro de amplitude no qual a energia está distribuída em frequências discretas, múltiplos inteiros da frequência fundamental chamados de harmônicos. Nesse modelo,

⁴ Uma excelente revisão sobre a teoria acústica de produção de fala pode ser encontrada em Camargo (2002).

considera-se que a variação da frequência fundamental e da amplitude de vibração das pregas vocais independe do filtro. A relativa independência entre os efeitos da fonte e do filtro permite a identificação dos sons de fala, mesmo com variações do tom laríngeo.

As ressonâncias naturais do trato vocal são chamadas de formantes, identificados por números inteiros (frequência, largura de banda e amplitude), os quais correspondem à função de transferência do trato vocal dentro do modelo de Fonte/Filtro. Estão potencialmente associados a um pico de energia no espectro do sinal de fala e não geram energia, apenas modificam a energia gerada na fonte.

A partir da modelagem de um tubo que simula o trato vocal, estuda-se a produção das ressonâncias, que variam de acordo, sobretudo, com a posição e o grau de constrição no trato bem como com os acoplamentos de tubos secundários. O modelo utilizado para explicar a vogal neutra é do tubo fechado em uma extremidade (glote) e aberto na outra (lábios). Como já foi dito, o trato vocal é o ressoador e funciona como um filtro, que modifica a amplitude dos componentes espectrais da fonte glotal, (função de transferência), determinando os formantes (picos de ressonância). A fonte é representada pelo tom laríngeo e/ou pelo ruído produzido em algum ponto do trato por uma constrição parcial no caso da articulação de algumas consoantes.

As características acústicas de cada som de fala são determinadas exclusivamente pelas constrições e bifurcações do trato vocal e pela frequência fundamental, que é definida pelo número de vibrações das pregas vocais por unidade de tempo.

A Teoria Fonte-Filtro possibilita prever as frequências dos formantes. Para a modelagem da vogal neutra central [ɪ], ou *schwa*, o tubo é considerado uniforme na sua dimensão transversal. Para o tubo uniforme, aberto em uma das extremidades, o cálculo das ressonâncias é feito através da fórmula - $f_n = (2n-1) \cdot c/4L$, onde n é o número do formante; c é a velocidade do som no ar e L é o comprimento do tubo.

Observando a fórmula, verificamos que modificações no comprimento do

tubo, determinam alterações acústicas. Um tubo com o comprimento maior apresenta ressonâncias em frequências mais baixas do que um de comprimento menor. O comprimento do trato vocal (tubo) varia em função da idade (crescimento) e do sexo, sendo mais curto em crianças e mulheres que, portanto, apresentam formantes mais altos em relação aos dos homens.

O padrão dos formantes é a principal pista para a percepção das vogais, sendo que a sua distinção se dá principalmente em função de F1, F2, F3. Pode-se dizer que F1 e F2 variam de acordo com o posicionamento da língua, ou seja, F1 varia principalmente com a altura da língua e F2 com o seu avanço. Quanto mais alta é a vogal, mais baixo é o primeiro formante e quanto mais anterior, mais alto é o segundo formante, e vice-versa.

Do ponto de vista da acústica, a descrição das vogais é mais simples do que a das consoantes, pois as informações relevantes estão contidas nelas mesmas. No caso das consoantes, é a combinação de eventos acústicos que ocorrem no momento em que as estruturas do trato vocal se movem em direção ou se afastam da configuração da constrição que fornece as principais pistas para a percepção de uma consoante em particular. Além disso, a descrição das vogais é feita principalmente em termos da frequência dos formantes e para a descrição das consoantes várias medidas de dimensões acústicas diversas são necessárias, tais como intensidade e tempo, além da frequência (Hayward, 2000).

Para a análise das consoantes, também é possível aplicar-se o modelo do tubo. Do ponto de vista acústico, não existe uma linha divisória muito bem marcada entre vogais e consoantes. As nasais e laterais, por exemplo, à semelhança das vogais, não apresentam fontes de ruído. Já as obstruintes são caracterizadas pela presença de ruído. No caso das plosivas, o ruído é transiente e resulta da liberação repentina da obstrução. O ruído das fricativas é contínuo e é resultado da corrente de ar que passa entre dois articuladores, gerando turbulência em virtude da passagem estreita.

A produção de qualquer consoante envolve o contato ou pelo menos a aproximação entre dois articuladores em algum ponto do trato vocal, entre os

lábios e a laringe. A constrição é mantida por um certo período de tempo, gerando conseqüências específicas. Ela não ocorre isoladamente: os articuladores precisam se mover em direção à posição-alvo. Quando uma vogal precede ou sucede uma consoante, os movimentos relacionados à produção da consoante provocarão reflexos nos movimentos dos formantes das vogais em relação à sua porção estável. Esses movimentos são conhecidos como transição de formantes. Além disso, deve-se levar em conta a atividade das pregas vocais e suas conseqüências acústicas durante a produção da consoante. Se ela é não-vozeada durante o intervalo de constrição, não será observada nenhuma vibração e não se visualizará no espectrograma a barra de sonoridade. A não-ocorrência da barra de sonoridade praticamente coincide com o tempo da constrição, com diferença de um pequeno intervalo de tempo entre o início da consoante até o vozeamento (*Voice Onset Time* ou *VOT*). Se a consoante for aspirada, esse intervalo de tempo será mais longo. (Hayward, 2000).

1.4 Consoantes plosivas

No PB, existem seis consoantes plosivas, duas bilabiais, duas dentais/alveolares⁵ e duas velares, sendo que, para cada ponto de articulação, uma é vozeada e a outra não vozeada - /p/; /b/; /t/; /d/; /k/; /g/.

As características articulatórias essenciais para uma consoante plosiva são o fechamento completo momentâneo do trato vocal e a conseqüente interrupção do fluxo de ar seguido da soltura, com o movimento dos articuladores em direção à configuração articulatória do som subsequente, caracterizando a transição dos formantes. O correlato acústico do fechamento é o silêncio, embora, nas plosivas vozeadas, uma fraca energia em baixas freqüências possa ser detectada, decorrente da vibração das pregas vocais. Para o inglês, a duração do silêncio pode variar de 50 a 100 ms, podendo sofrer influência das condições prosódicas (Kent, Read, 1992b). O silêncio é uma pista necessária para a percepção de uma plosiva. A transição do formante é a variação dinâmica do padrão de freqüência em

⁵ O ponto dental/alveolar será referido neste trabalho apenas como alveolar

função do tempo. Até que ponto a transição dos formantes sinaliza o ponto de articulação é uma questão de grande interesse na pesquisa em fonética acústica. Outro ponto a ser estudado é a possibilidade de inferir características da coarticulação pela análise dos padrões de transição de formantes.

Os sons plosivos são produzidos por um complexo de movimentos do trato vocal. Com a cavidade nasal fechada, um rápido fechamento e posterior abertura é realizado em algum ponto da cavidade oral. Atrás desse ponto, a pressão aumenta e é subitamente aliviada quando o fechamento é desfeito. Durante o período de fechamento, as cordas vocais podem ou não estar vibrando, determinando a produção de uma consoante plosiva vozeada ou não-vozeada (Halle et al, 1991⁶).

Fant (1991⁷) descreve as plosivas do sueco em posição inicial de sílabas tônicas. Segundo o autor, uma plosiva não-vozeada completamente desenvolvida em posição tônica em posição inicial apresenta cinco segmentos sucessivos:

- oclusão - durante a qual pode ocorrer o silêncio ou vozeamento;
- transiente - resposta do trato vocal à soltura da pressão. A duração do ruído transiente é da ordem de 2 a 30 ms e geralmente menos que 10ms;
- segmento fricativo - caracterizado pelo ruído produzido na cavidade anterior à constrição. As anti-resonâncias interagem de forma a cancelar os formantes cavidade posterior, enquanto os de cavidade anterior prevalecem.
- aspiração - é caracterizado por um ruído semelhante ao /h/ originado de alguma fonte próxima à glote excitando todos os formantes, exceto o mais baixo. Os formantes F2, F3 e F4 são os constituintes mais típicos. O segmento aspirado pode, em parte, co-ocorrer com o segmento fricativo e sobressai-se na medida em que o grau de abertura aumenta. Formantes subglotais podem surgir devido ao acoplamento entre cavidades sub e supra-glóticas;

⁶ Publicado originalmente em 1957- Halle M, Hughes GW, Radley JP. Acoustics properties of stop consonants. J Acoustic Soc Am 1957; 27: 107-116.

⁷ Publicado originalmente em 1969 - Fant G. Stops in CV-syllables: speech analyses. Spectrographic characteristics of normal speech. Labs. Quartely Progress Reports 1969; 41.

- porção inicial do som vozeado subsequente - ou transição dos formantes, na medida em que este é influenciado pela coarticulação com a plosiva.

O ruído transiente, o segmento fricativo e a aspiração são chamados de *burst* estouro ou ruído plosivo. Adotaremos, neste trabalho, a denominação *burst*.

Stevens (1998) descreve os eventos acústicos e articulatórios relacionados à produção das plosivas. Para ele, a produção de uma consoante plosiva em posição intervocálica consiste numa seqüência de uma série de eventos acústicos e articulatórios. Na fase inicial, ocorre o fechamento em algum ponto do trato vocal por meio da manipulação de uma estrutura articulatória específica. Em seguida, o trato vocal permanece fechado durante um intervalo de tempo que é seguido pela soltura. Durante o período de fechamento e nas suas adjacências o estado das pregas vocal pode ser ajustado para gerar características particulares no som – vozeado e não-vozeado.

As conseqüências acústicas dos movimentos das estruturas articulatórias são inibir a vibração das pregas vocais durante o período de oclusão do trato vocal e produzir em algumas regiões de freqüência um decréscimo abrupto na amplitude na implosão e um aumento abrupto na soltura, gerando modificações espectrais distintivas na implosão e na plosão, dependendo da estrutura que é usada para realizar a constrição. Para que a constrição seja efetiva, são freqüentemente necessários ajustes também do corpo da língua e da mandíbula, embora essas estruturas tenham papel secundário e apresentem deslocamentos suaves.

Ocorre uma descontinuidade na velocidade da mudança da área da secção transversal na constrição, que atinge a taxa de mudança máxima logo antes da implosão e após a soltura. Se a constrição é feita por lábios ou ponta de língua, a aceleração (velocidade de mudança) na área da secção transversal fica entre 40 e 100 cm² por segundo. Se o fechamento é feito pelo dorso da língua a velocidade de mudança é um pouco menor, em torno de 25 cm² por segundo. Até alguns milissegundos antes da implosão, os movimentos supraglotaais não exercem muita influência no padrão de vibração glotal e não causam modificação importante na pressão intraoral, o que passa a ocorrer a partir do instante do fechamento.

Imediatamente após o fechamento, a pressão cresce atrás da constrição e provoca o deslocamento das paredes do trato vocal e da glote. No instante do fechamento, a pressão intraoral é igual a zero. À medida que a pressão intraoral cresce, provoca-se uma força abduutora nas pregas vocais e, como consequência, há um deslocamento posterior. No momento em que a pressão intraoral aproxima-se da pressão na região da glote, o fluxo de ar decresce. O aumento da pressão intraoral é mais rápido para as constrições mais posteriores no trato, como no caso das velares e uvulares.

No momento da soltura, ocorrem um aumento abrupto no fluxo de ar por meio da constrição e da glote e um rápido decréscimo da pressão intraoral. Nos primeiros milissegundos da soltura, há um fluxo inicial transiente que aumenta de acordo com a velocidade de aumento da área da secção transversal. Depois da soltura, ocorre um rápido aumento na pressão no nível da glote, que passa a vibrar quando ultrapassa 3 cm H₂O, o que corresponde a um intervalo de cerca de 2 ms quando a soltura é rápida (lábios e ponta de língua) e 15 ms quando é mais lenta (dorso da língua). Durante o tempo imediatamente anterior e posterior ao fechamento, as modificações na área da secção transversal da constrição provocam mudanças nas frequências naturais do trato vocal. O primeiro formante abaixa à medida que a área da constrição diminui, sendo, portanto, o seu deslocamento mais rápido para as constrições bilabiais e alveolares e mais lento para as velares. Imediatamente antes do fechamento, o primeiro formante abaixa e apresenta uma tendência de alargamento da largura de banda. Nas proximidades da oclusão, a amplitude do primeiro formante cai cerca de 10 dB em relação à vogal anterior, indicando uma diminuição na amplitude dos pulsos glotais. No caso das velares, a diminuição de F1 é mais lenta e a diminuição da amplitude é ainda maior, provavelmente em virtude da menor área disponível para expansão no trato vocal (atrás da constrição) se comparada à oclusão bilabial e alveolar.

Para as bilabiais, a maior parte da transição de F1 não é visível no espectro, uma vez que a transição está praticamente completa antes que as pregas vocais comecem a vibrar. O posterior aumento de F1 provavelmente deve-se ao fato de que a constrição labial continua a se expandir à medida em que a mandíbula

abaixa para a realização de uma vogal como /ɪ/. A transição de F1 para as bilabiais deve estar completa em cerca de 20 ms, sendo um pouco mais longa para alveolares e velares. No caso das velares, a amplitude no início da transição de F1 está bem menor em relação à vogal.

Quando a constrição é feita nos lábios, todas as ressonâncias do trato vocal estão relacionadas a ressonâncias da cavidade posterior à constrição. Todas as frequências tendem a abaixar antes da oclusão, mas os formantes de frequências mais agudas tendem a cair mais lentamente do que os de frequência mais baixa. Depois da soltura, a trajetória dos formantes espelha a trajetória antes da oclusão e tendem a subir rapidamente. A trajetória dos formantes que segue à soltura de uma plosiva bilabial depende em parte da vogal subsequente, uma vez que o corpo da língua pode já estar posicionado no momento da soltura. No caso da vogal subsequente ser /ɪ/, a configuração articulatória pode ser modelada a partir de um ressoador de dois tubos.

Se considerarmos que os dois tubos têm comprimentos iguais, a frequência natural de cada secção individual é a mesma. Com uma área seccional de $0,7 \text{ cm}^2$ (tubo estreito) e 7 cm^2 e um comprimento total de 17 cm, as frequências naturais mais baixas serão 800 Hz e 1180 Hz. Após a soltura da constrição (de 0 para 7 cm^2 , considerando a taxa de variação a 100 cm^2 por segundo), haverá uma rápida ascensão de F1, embora o movimento observável de F1 seja relativamente pequeno em virtude da rápida modificação ocorrer muito próxima da soltura, não podendo ser visualizado. Observa-se uma transição relativamente discreta de F2, visto que as ressonâncias da cavidade posterior são pouco afetadas pelas modificações da cavidade anterior. Assim, uma discreta transição positiva de F2 é típica de uma constrição labial associada a uma vogal posterior.

Um padrão um pouco diferente surge quando a constrição é feita com a ponta da língua. Nesse caso, a constrição é feita num ponto cerca de 1,5 cm a 2,5 cm atrás da abertura dos lábios. Para produzir uma plosiva alveolar, é necessário que o corpo da língua fique em posição anteriorizada de forma que a ponta possa atingir os alvéolos. Essa anteriorização pode ser representada por um estreitamento do tubo na região anterior e um alargamento na região posterior. No

momento após a soltura da constrição da ponta da língua em contato com os alvéolos, ocorrerá um aumento na área de constrição e um movimento da língua um pouco para frente, no caso da vogal subsequente ser anterior, podendo ser observada uma discreta subida de F1 e F2 e uma transição ascendente ainda menor para F3. No caso de uma vogal posterior, o corpo da língua se move para trás, formando um estreitamento da região faríngea e isso resulta num abaixamento de F2, embora o padrão ascendente permaneça para F1. Para as alveolares, F3 apresenta uma ligeira ascensão. Essas características da transição dos formantes para as alveolares contrastam com as bilabiais.

As velares apresentam características bastante diferentes por diversas razões: a localização da constrição é posterior em relação às bilabiais e alveolares; a taxa de modificação da secção é menor e a extensão da oclusão é maior. À medida que o corpo da língua movimenta-se em direção ao palato mole ou à parte posterior do palato duro, as frequências mais baixas da cavidade anterior estarão associadas ou a F2 ou a F3. Além disso, uma das frequências da cavidade posterior estará relativamente próxima da relacionada à cavidade anterior, resultando numa aproximação de F2 e F3. No instante da soltura, F1 está em torno de 200 Hz, enquanto F2 e F3 estão próximos. Após a soltura, F1 sobe e F2 e F3 separam-se, F2 descendo e F3 subindo. A transição observada será mais lenta dado que a taxa de modificação da área seccional após a soltura é mais lenta.

Quando uma plosiva ocorre em posição adjacente a uma vogal, geralmente estão presentes as três pistas: silêncio, ruído transiente e transição, ou transição, silêncio e ruído transiente, das quais apenas o silêncio é uma pista obrigatória. O silêncio e a transição ou o silêncio e o *burst* são suficientes para a percepção da plosiva.

O *burst*, que leva esse nome dado sua brevidade (de 5 a 40 ms), assemelha-se ao segmento de ruído das fricativas. Entretanto, sua duração é bem mais curta, porque, na abertura dos articuladores da posição de oclusão para a vogal subsequente, a turbulência cessa rapidamente. O *burst* é facilmente localizado no espectrograma pois corresponde a uma breve espícula vertical. O espectro do ruído da plosiva é explicado pelo efeito dos formantes da cavidade anterior e está

relacionado com o ponto de articulação. Halle et al (1991), apontam que, no caso das labiais, existe uma predominância de frequências baixas (de 500 a 1500 Hz), altas para as dentais (acima de 4000 Hz) e médias para as velares (entre 1500 a 4000 Hz), ocorrendo uma variação para velares e bilabiais de acordo com a vogal que sucede a consoante. Para as velares, a frequência aumenta com vogais anteriores e abaixa com posteriores; para as bilabiais a frequência abaixa com vogais posteriores. Segundo esses autores, o *burst*, isoladamente, pode fornecer pistas para o reconhecimento do ponto de articulação da plosiva.

Ao contrário do *burst* muito facilmente de identificado no espectrograma, o início do intervalo de silêncio não é muito simples de ser localizado porque o bloqueio do fluxo do ar não é tão súbito quanto a soltura, uma vez que os articuladores vão se juntando gradativamente. Além disso, apesar do véu palatino estar levantado e os articuladores estarem juntos, o selamento pode não ser total. Dessa forma, alguma energia acústica pode estar ainda visível no início do intervalo de silêncio (Hayward, 2000).

O ruído transiente e os formantes de transição são pistas acústicas complementares, e a sua integração, provavelmente, confere identidade ao fonema. Os formantes das vogais adjacentes são alterados pelo ponto de articulação da plosiva e, portanto, na vogal também estarão contidas pistas acústicas da consoante.

As consoantes plosivas podem ser vozeadas ou não-vozeadas. Entretanto, não é para todas as línguas que a presença do vozeamento é determinante para a percepção de uma plosiva vozeada: no inglês, por exemplo, ele não é crucial para essa distinção. A diferença mais importante entre essas duas classes de plosivas é o fato de que, na produção das não-vozeadas, a pressão intra-oral atrás da obstrução é maior, resultando na maior intensidade do *burst*, e na ocorrência de aspiração (Halle et al, 1991). Ambas as classes de plosivas também apresentam diferenças nos espectros do *burst* e na duração da vogal anterior. Devido ao fato de, no inglês, a diferença entre vozeadas e não-vozeadas não ser determinada pela presença ou ausência do vozeamento, os autores propõem a nomenclatura “*tenso*” para /p/ /t/ /k/ e “*não tenso*” para /b/ /d/ /g/.

Os termos vozeados e não-vozeados podem ser usados no sentido que lhes dá a Fonética, significando que há ou não vibração das pregas vocais durante o intervalo de fechamento. Contudo tais termos também podem ser usados para diferenciar contrastes fonêmicos, como no caso das vozeadas do inglês /b/, /d/, /g/, em cuja produção não ocorre vozeamento durante a oclusão, mas se opõem a /p/, /t/, /k/, que são aspiradas. A ocorrência da aspiração deve-se ao ruído conseqüente da passagem de ar pelas pregas vocais parcialmente fechadas.

No caso de plosivas aspiradas, o intervalo sem vibração de pregas vocais estende-se mais do que no caso das não-aspiradas, quando o vozeamento inicia-se logo depois da plosão, com o início da vogal. Na produção das aspiradas, o intervalo sem vozeamento estende-se além da constrição, sobrepondo-se à transição dos formantes para a vogal subsequente. As plosivas vozeadas no sueco não apresentam aspiração. As não-vozeadas e vozeadas distinguem-se pela duração do tempo de início do vozeamento (*voice onset time – VOT*) de cerca de 125 ms para as primeiras contra de 10 a 25 ms para as últimas. No caso das plosivas totalmente vozeadas (como no PB), o intervalo de constrição não é completamente silencioso, porque as pregas vocais continuam a vibrar (Behlau, 1986; Madureira, et al, 2002). Uma barra de baixa frequência, conhecida como barra de sonoridade, pode ser visualizada no espectrograma (Stevens, 1998).

As consoantes plosivas não ocorrem isoladamente e, portanto, sua análise está necessariamente submetida aos fenômenos da coarticulação. O contexto fonético, ou seja, os sons adjacentes, influenciam os padrões acústicos dos sons da fala. A fala é fundamentalmente coarticulada, isto é, compreende a superposição de gestos articulatórios, implicando ajustes que podem ser antecipatórios ou perseverativos. O foco também altera os parâmetros acústicos de duração, intensidade e frequência fundamental dos sons. Já a taxa de elocução modifica as medidas das características acústicas, principalmente em termos da duração e frequência de formantes. Existe também a variabilidade em conseqüência da idade e do sexo do falante.

Segundo Farnetani, Recasens (2000), a variabilidade da fala se deve em

grande parte ao fenômeno universal da coarticulação. Os autores se referem à coarticulação como a “penetrante, sistemática e recíproca influência entre os segmentos de fala contíguos e mesmo não contíguos” (Farnetani, Recasens, p. 31). Eles analisam a questão da coarticulação à luz das teorias e modelos de produção de fala atuais e apresentam uma revisão da literatura. De acordo com algumas teorias, as variações da coarticulação ocorrem no nível do planejamento da fala e, assim, modificam o próprio planejamento. Outras atribuem os fenômenos da coarticulação aos processos de produção de fala, de forma que a invariância, em níveis mais abstratos, estaria preservada.

Öhman (1966), analisando sílabas VCV do sueco, conclui que a transição do segundo formante não depende apenas da identidade da consoante, mas também da identidade da outra vogal. Esse efeito da transição transconsonantal é chamado de coarticulação vogal para vogal. Os dados foram comparados com dados do inglês americano e do russo, sendo que, para esse último, os efeitos da segunda vogal não foram significativos, o que sugere que esse é um fenômeno lingüístico e não meramente decorrente das restrições impostas por aspectos físicos do sistema articulatorio.

1.4.1 Contraste vozeado – não vozeado

Nas plosivas em posição inicial, o contraste vozeado – não-vozeado está associado a vários parâmetros acústicos: tempo de início do vozeamento (*voice onset time* – *VOT*), frequência fundamental (f_0) no início da vogal subsequente, contorno de f_0 no início da vogal subsequente, *onset* e transição do primeiro formante (F_1), intensidade do *burst*, duração da consoante e duração das vogais adjacentes.

O *voicing onset time* (*VOT*) corresponde ao intervalo entre a soltura da obstrução na articulação das consoantes plosivas e o início do vozeamento (Lisker, Abramsom, 1964). O *VOT* nas plosivas não-vozeadas é positivo e nas vozeadas pode ser zero ou negativo. Sua medida varia em virtude do ponto de articulação, já que, quanto maior a distância da constrição em relação à fonte glotal, melhores as condições para o início da vibração das pregas vocais. Assim sendo, encontra-se

uma graduação nos valores do *VOT*, que são menores para as bilabiais, seguidos dos alveolares e velares.

Behlau (1986) estudou a medida do *VOT* das plosivas do PB, encontrando a mesma relação entre os valores descritos na literatura. Por meio da manipulação desse parâmetro, a autora concluiu que essa é uma dimensão satisfatória para a discriminação da sonoridade das plosivas do PB. As médias encontradas para falantes adultos são (em milissegundos): [p] 10,37; [b] -35,91; [t] 14,85; [d] -33,27; [k] 43,73; [g] 0,17.

A duração da consoante plosiva também está associada ao contraste vozeado/não-vozeado. Barbosa (1996) encontra valores maiores de duração para as plosivas não-vozeadas em relação a seus pares vozeados, seguindo a tendência universal. No caso da oclusiva vozeada, ocorre o aumento da pressão supra-glótica pelo vedamento do trato vocal, o que dificulta a manutenção da vibração das pregas vocais. As medidas de duração média em milissegundos das consoantes plosivas encontrados por esse autor para um sujeito do sexo masculino (desvio-padrão entre parênteses) são as seguintes:

p - 120 (20) ms	b - 86 (17) ms
t - 113 (20) ms	d - 71 (17) ms
k - 121 (21) ms	g - 67 (16) ms

Veloso (1995) estudou a duração da oclusão nas plosivas vozeadas e não vozeadas do português europeu e também encontrou valores significativamente maiores para as durações das plosivas não-vozeadas se comparadas com seus pares vozeados:

p - 132 (13) ms	b - 80 (17) ms
t - 133 (23) ms	d - 70 (15) ms
k - 123 (10) ms	g - 75 (14) ms

Outro parâmetro acústico ligado ao constraste de vozeamento é a duração

das vogais adjacentes. Segundo Peterson, Lehiste (1991⁸), as vogais que seguem plosivas vozeadas são cerca de 30ms mais longas do que as que seguem plosivas não vozeadas e essa diferença corresponde a cerca de metade da duração do *burst*. Fant (1991) estudando a seqüência [CaC]:Ca] (C=p, b, t, d, k, g), encontra dados semelhantes, ou seja, o prolongamento da vogal depois da plosiva vozeada. Ele observou o alongamento das plosivas não vozeadas e das vogais que seguem as vozeadas e a diferença entre as consoantes era maior do que entre as vogais, resultando na duração mais longa das sílabas com plosivas não-vozeadas. Para a primeira sílaba, a pré-tônica, a vogal após a vozeada era 40 ms mais longa, enquanto na sílaba tônica a vogal foi 60 ms mais longa no contexto vozeado. Na vogal da última sílaba, pós-tônica, o autor não observou diferença de duração entre o contexto vozeado e não-vozeado. Entretanto, segundo Peterson, Lehiste (1991), a duração da vogal anterior à plosiva sofre mais influência do vozeamento do que a posterior, já que nesta ocorre um alongamento quando antecede consoantes vozeadas. Para a língua inglesa, os autores encontram uma média de 297 ms na duração das vogais que precedem consoantes vozeadas contra 197 ms no contexto de não vozeadas. Todavia, tal variação na duração da vogal (*vowel length effect* – *VLE*) não ocorre em todas as línguas.

Em relação ao PB, encontramos o trabalho de Brito (2000). Entre outras medidas acústicas, a autora analisou a duração da vogal /a/ precedendo plosivas e fricativas não-vozeadas e vozeadas, na posição intervocálica na sílaba tônica. Para um grupo composto por crianças de 6 a 10 anos, com fala considerada normal, a média da duração das vogais que precedem as plosivas vozeadas foi 22% maior que a das que precedem as não vozeadas - 119 ms para as vozeadas (variação de 76 a 158 ms) e 97 ms para as não vozeadas (variação de 72 a 150 ms). Pela análise estatística, ela sugere que essas crianças podem, embora de forma insuficiente, estar usando a duração da vogal como indicador de diferença entre os fonemas obstruintes vozeados e não-vozeados⁹.

⁸ Publicado originalmente em 1960 – Peterson GE, Lehiste I. Duration of syllabic nuclei in english. J. Acoust. Soc. Am. 1960; 32: 693-703.

⁹ Em nosso trabalho, essa pista não poderá ser investigada, pois a plosiva encontra-se em posição inicial de palavra. Mesmo sendo a palavra inserida na frase-veículo, existe uma

A frequência de F1 no início da vogal subsequente também é apontada como um parâmetro acústico vinculado ao vozeamento. Kluender (1991), em um experimento no qual investiga a resposta a estímulos sonoros sintetizados por codornas japonesas, mostra que a frequência do início do primeiro formante (*onset* de F1) diferencia perceptivamente as plosivas vozeadas das não-vozeadas. O primeiro formante das vogais que seguem as plosivas vozeadas apresenta *onset* de F1 mais baixo, enquanto as vogais que seguem as plosivas não-vozeadas mostram o *onset* do F1 mais alto.

Shimizu (1996) estudou os contrastes de vozeamento das plosivas em seis línguas asiáticas, encontrando tendência de aumento da frequência de *onset* do F1, a partir das plosivas vozeadas (as mais baixas), até as não-vozeadas aspiradas (as mais altas). Nas plosivas não-vozeadas, o intervalo entre o ruído transiente e o início da periodicidade resulta em uma transição mais abrupta de F1 e a frequência no início do primeiro formante é mais alta. Esse efeito varia com a vogal subsequente e é mais evidente para as vogais baixas porque essas tendem a ter F1 mais altos. Tais diferenças podem ser explicadas pelas variações na produção dessas duas classes de plosivas. No caso das vozeadas, há o vozeamento durante a oclusão e o F1 apresenta uma frequência muito baixa. No momento da soltura oral, F1 sobe rapidamente a partir de uma região de frequência bem baixa. No caso das não-vozeadas, pela demora no início do vozeamento, não se observa a transição de F1, que já inicia em frequências mais altas.

Benki (2001), por meio de um experimento no qual o VOT e a transição de F1 (frequência do início do F1 e duração da transição) eram manipulados, fornece mais evidências a favor da hipótese de que a transição de F1 mais rápida e com frequência mais alta em seu início influencia a percepção de uma plosiva não-vozeada.

Outro parâmetro acústico que varia em decorrência do vozeamento é a

fronteira sintático-semântica entre a vogal anterior e a plosiva, o que pode interferir em sua duração.

freqüência fundamental (f_0). O contraste de vozeamento das plosivas está associado a modificações no f_0 no início da vogal subsequente. As plosivas não vozeadas estão associadas a uma faixa de freqüência mais alta e as vozeadas, a freqüências mais baixas. Shimizu (1996) aponta a hipótese aerodinâmica como uma explicação para essas diferenças: durante a plosiva vozeada, a pressão oral aumenta gradualmente, favorecendo a diminuição da pressão entre as pregas vocais, o que, por sua vez, causa uma diminuição na freqüência de f_0 . Durante a soltura oral, a pressão volta ao normal, resultando na diminuição seguida do aumento da freqüência de f_0 após a soltura. Já na produção das plosivas não-vozeadas, o fluxo de ar passando nas pregas vocais é supostamente bastante grande na soltura, reforçando o efeito de Bernoulli¹⁰, o que resulta num fechamento mais rápido das pregas vocais, aumentando a freqüência de suas vibrações logo após a soltura. Depois de algum tempo, o fluxo de ar volta ao normal e f_0 também. Dessa forma, no início da vogal, logo após a soltura da plosiva não-vozeada, observa-se uma elevação de f_0 com posterior decréscimo.

Holt, et al (2001) afirmam que o fato de as vogais que sucedem as plosivas vozeadas tenderem a apresentar f_0 mais baixo do que seus pares não-vozeados é tão estável entre as línguas que chegou a ser atribuído à restrição fisiológica da produção de fala. Entretanto, eles ressaltam que estudos em várias línguas demonstram que a correlação entre VOT e f_0 só ocorre nas consoantes que são usadas distintivamente nas línguas. Um f_0 mais baixo também favorece a percepção de uma consoante como vozeada. Os autores desenvolvem um experimento com codornas que aprenderam a responder a estímulos manipulados de três maneiras: a covariância de f_0 e VOT seguia o padrão natural, ou seja, menor VOT , menor f_0 ; o padrão era invertido, sendo menor VOT , maior f_0 e finalmente, um padrão aleatório no qual f_0 e VOT não apresentavam covariação. O objetivo foi investigar duas hipóteses relacionadas a essa covariância: a primeira hipótese, a da saliência auditiva, preconiza que essa condição cria a facilitação da percepção do vozeamento, uma vez que torna as freqüências baixas mais evidentes; a segunda atribui esse padrão perceptual ao fato de ocorrer um

¹⁰ A lei de Bernoulli (chamada de efeito de Bernoulli) estabelece que o aumento na velocidade do ar ocasiona a redução da pressão estática.

aprendizado. Os resultados favorecem a hipótese de que ocorre um aprendizado, sugerindo que a frequência fundamental não é determinante sobre a percepção de vozeamento. Os autores, considerando os resultados de seu estudo, ressaltam que a experiência influencia a percepção e afirmam que o sinal de fala, sendo estruturado a partir de regularidades impostas por limites físicos sobre os processos articulatórios e por restrições lingüísticas, molda os hábitos dos falantes.

1.4.2 Contraste por ponto de articulação

Os dois principais correlatos acústicos do ponto de articulação parecem ser o espectro do *burst* e a transição dos formantes para a vogal subsequente. Experimentos com percepção mostram que esses dois parâmetros de fato funcionam como pistas para o ponto de articulação. Entretanto, para os dois casos, verifica-se falta de invariância (Hayward, 2000). Para as plosivas do inglês /b/, /d/./g/, o *burst* do /b/ tem menos energia do que o /d/ e o /g/. O *burst* do /d/ tem mais energia nas frequências altas. O *burst* do /g/ apresenta uma concentração de energia nas frequências médias e, por isso, as velares são chamadas de compactas. Já as bilabiais e alveolares são designadas difusas, pois a energia do *burst* não está concentrada.

No trabalho clássico de Liberman et al (1996¹¹), uma série de sílabas do tipo CV (consoante+vogal) foram sintetizadas a partir de doze *bursts*, variando nas frequências de 360 a 4320 Hz combinados com vogais de dois formantes. Os ruídos transientes compostos por frequências altas (acima de 2800 Hz) foram identificados como /t/. Já os *bursts* formados por frequências mais baixas foram identificados ora como /p/, ora como /k/, dependendo da vogal subsequente. Como regra geral, pode-se dizer que, se a frequência do *burst* está próxima do F2 da vogal subsequente, o som ouvido tende a ser /k/. Esse estudo foi, em parte, encorajador, uma vez que mostrou que é possível identificar o fonema a partir da

¹¹ Originalmente publicado em 1952.- Liberman AM, Delattre PC, Cooper FS. The role of selected stimulus variables in the perception of unvoiced stop consonants. American Journal of Psychology 1952; 65: 497-516.

informação do *burst*, embora a relação não seja tão direta quanto era esperado.

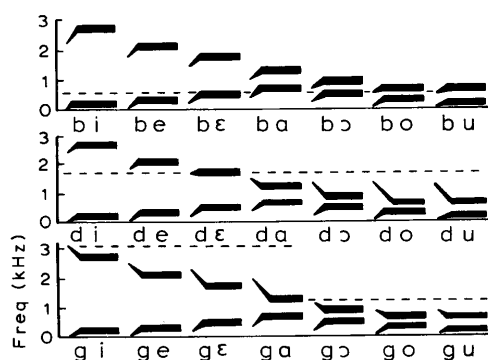
A transição dos formantes é outra pista para a percepção do ponto de articulação. Se o ruído transiente for retirado e forem mantidas as transições, é possível perceber os três pontos de articulação das plosivas - labial, alveolar e velar. Também se for retirada a porção aperiódica da fricativa /s/ em /say/ e forem mantidas as transições dos formantes, a palavra será percebida como /day/. Os experimentos com falantes do inglês têm se concentrado nas vozeadas /b/, /d/ e /g/, já que a aspiração das não-vozeadas sobrepõem-se à transição.

As transições são geralmente descritas considerando uma sílaba do tipo CV. Assim se, por exemplo, for dito que para as labiais a transição de F2 é ascendente, subentende-se que essa referência seja relacionada à transição da consoante em direção à vogal. No caso da consoante suceder a vogal, ocorre o espelhamento do movimento da transição do formante, ou seja, o movimento do formante da vogal para a consoante será descendente, embora variações na velocidade da transição possam existir (Stevens, 1998).

Para os pontos de articulação orais, o F1 é sempre ascendente na transição CV, o que está de acordo com a teoria que diz que, durante a oclusão da cavidade oral, o primeiro formante é zero (Formante de Helmholtz). Entretanto, F2 e F3 variam conforme o ponto de articulação e também de acordo com a vogal subsequente. De forma geral, para as bilabiais, as transições de F2 e F3 são planas ou ascendentes, sendo que a subida de F2 é mais acentuada quando a vogal seguinte é anterior e, portanto, o F2 tende a ser mais alto. Quando a vogal seguinte é posterior e conseqüentemente o F2 é mais baixo, a transição fica mais nivelada. Para as velares, a transição de F2 geralmente começa um pouco acima do segundo formante, num movimento descendente em direção à vogal e o F3 inicia no mesmo ponto, subindo, o que, na figura do espectrograma, parece evidenciar uma “forma de asa”, que varia de acordo com a vogal.

As alveolares apresentam transições de F2 ascendentes ou niveladas com vogais anteriores e descendentes no contexto de vogais posteriores. A transição de F3 será geralmente nivelada ou descendente.

Os primeiros estudos sobre percepção e transição de formantes, utilizando síntese de fala focalizaram o comportamento de F2. O comportamento de F2, para a alveolar, descrito acima, pode ser explicado pelo conceito de *locus* (Dellatre et al 1955). O *locus* é um ponto na escala de frequência onde a transição inicia ou “aponta para” (no caso das plosivas, a direção para onde a transição aponta é um conceito necessário uma vez que o início da transição, correspondente ao movimento da língua em direção à vogal subsequente, não é visível no espectrograma). Assim, o *locus* de F2 (que reflete os movimentos da língua) de uma consoante, independentemente da vogal que vem a seguir, localiza-se em um mesmo ponto e o que varia é o modo como a transição desloca-se de acordo com a vogal. No caso das alveolares, o *locus* de F2 está por volta de 1800 Hz. Por essa razão, a transição sempre inicia em torno desse valor e cai para vogais posteriores que apresentam F2 mais baixos que esse valor. Para as bilabiais, o *locus* está em torno de 720 Hz e, por isso, a transição de F2 é sempre ascendente, mais acentuada em vogais anteriores e menos acentuada em sílabas com vogais posteriores. No caso das velares, o *locus* está próximo de 3000 Hz para vogais anteriores e 1300 Hz para vogais posteriores, devido à tendência de o ponto de constricção da língua para a obstrução variar de acordo com o ponto da vogal posterior. A figura 9, retirada de Kent, Read (1992), é uma adaptação da figura original de Dellatre et al (1955) e ilustra o conceito de *locus*:



Obs. As linhas pontilhadas estimam o *locus* de F2 para cada ponto de articulação.
 Fonte: Kent, Read (1992, p. 117), figura adaptada de Dellatre et al (1955).

Figura 8 - Espectrogramas estilizados (padrões de F1 e F2) de sílabas do tipo CV - plosivas /b/, /d/ e /g/ em combinação com as 7 vogais do inglês

A transição de F2 para as alveolares e velares pode ser muito semelhante e por essa razão, a transição de F3, que se diferencia mais para os dois pontos de articulação, parece exercer um papel fundamental em sua distinção (Hayward, 2000).

Os trabalhos de Stevens, Blumstein (1978) e Blumstein, Stevens (1979) sugerem que as características espectrais estáticas dos primeiros 20 ms após a soltura são mais importantes para a percepção do ponto de articulação do que o *locus* e transição dos formantes tratados isoladamente. Os autores consideram que o ponto de articulação de uma plosiva em posição inicial pode ser identificado, independentemente do contexto vocálico, pelas características espectrais observáveis na soltura da consoante, que incluem informações não apenas da consoante, mas também da porção inicial da transição para a vogal. Eles postulam moldes estáticos que corresponderiam aos pontos de articulação, sendo que um espectro de configuração descendente, com um pico por volta de 1 kHz ou menos, está associado às plosivas bilabiais; um espectro compacto com o pico em torno de 1500 Hz a 4000 Hz está associado a plosivas velares e um espectro ascendente com pico nas frequências mais altas que 3 kHz está associado às plosivas alveolares.

Posteriormente, Kewley-Port (1982), Kewley-Port (1983), Kewley-Port et al (1983) sugerem que as características dinâmicas do espectro após a plosão identificam melhor o ponto de articulação do que os moldes derivados de espectros estatísticos propostos por Stevens, Blumstein, (1978). Kewley-Port (1982) enfatiza que as frequências de F2 e F3 no início da vogal subsequente também são pistas importantes para a identificação do ponto de articulação das plosivas e propõe um novo *locus* considerando as frequências de F2 e F3 no início da transição: para as bilabiais, F2 varia de 1100 Hz a 1500 Hz e F3, de 2200 Hz a 2400 Hz; para as alveolares, F2 gira em torno de 1800 Hz e F3, de 2500 Hz a 2700 Hz; para as velares, F2 varia de 1500 Hz a 2500 Hz e F3, de 2200 Hz a 3000 Hz.

Complementando os achados dos estudos anteriores, Ohde, Stevens (1983) concluíram que a amplitude relativa do *burst* também pode ser uma pista para a distinção do ponto de articulação entre os sons plosivos bilabiais e alveolares.

Mais tarde, Sussman et al (1991) investigaram a adequação da equação de *locus*¹² para identificar o ponto de articulação das plosivas. O objetivo da pesquisa foi explorar a ainda pouco conhecida métrica da equação de *locus* como um índice quantitativo de propriedades acústicas invariantes para o ponto de articulação no intervalo entre o início e o meio da transição de F2. Com a análise das produções de dez falantes do inglês americano, os resultados demonstraram que a equação de *locus* pode ser um poderoso descritor de ponto de articulação nos dez contextos de vogal investigados.

Engstrand et al (2000) discutem a eficácia de medidas para a identificação do ponto de articulação de consoante plosivas usando a equação de *locus* e a métrica baseada na relação entre F2 e F3 a partir da análise de sílabas do tipo CVC, onde a primeira consoante corresponde a uma das plosivas do sueco /p,t,k,b,d,g/; a segunda, à plosiva vozeada homorgânica e V, a uma das cinco vogais do sueco, ou seja: /pVb/, /bVb/, /tVd/, /dVd/, /kVg/ e /gVg/. A sílaba foi incluída em frase veículo e gravada por três falantes nativos do sueco em quatro diferentes condições de acento. Os resultados apontam para uma sobreposição entre os pontos de articulação em diferentes condições de acento quando identificados a partir da equação de *locus*. Por outro lado, a separação por ponto foi mais bem realizada a partir da métrica de F2 *versus* F3 no início da transição. Entretanto, mesmo assim, verificou-se uma sobreposição entre as velares e labiais, o que pode ser parcialmente resolvido pela informação do VOT (embora o VOT esteja principalmente associado ao contraste de vozeamento, também fornece pistas para o ponto de articulação já que varia em decorrência do local de constrição no trato vocal (Lisker, Abramsom, 1964)).

Smits et al (1996), por sua vez, realizaram um estudo com a finalidade de avaliar a participação das informações contidas no *burst* e na transição dos formantes para a identificação do ponto de articulação das consoantes plosivas em posição pré-vocálica. Manipularam sílabas com as plosivas do holandês /p/, /b/, /t/,

¹² A equação de *locus* foi originalmente concebida por Lindblom (1963) *apud* Sussman et al (1991) e é derivada de uma linha de regressão a partir de pontos formados pela relação entre F2 no início da transição – eixo y – e os valores de F2 no centro da vogal – eixo x.

/d/ e /k/, combinadas com as vogais /a, i, y e u/ a serem apresentadas para identificação para 20 estudantes universitários: sílabas com apenas o *burst*, sem o *busrt*, ou com o *burst* trocado. Os resultados mostram que a relevância perceptiva do *burst* e da transição varia de acordo com o ponto de articulação (bilabial, dental ou velar) e com o contexto da vogal. Além disso, os dados mostraram que o *burst* é mais efetivo nas plosivas não-vozeadas (as plosivas vozeadas do holandês apresentam pré-vozeamento e as não-vozeadas não são aspiradas). O *burst* da velar /k/ é mais efetivo do que o das bilabiais e dentais. Quando apresentado isoladamente, foi quase perfeitamente identificado e também resultou na melhor identificação nos estímulos com o *burst* trocado. Por outro lado, a sílaba iniciada com a velar foi pouco identificada quando o *burst* faltava, ou seja, quando só a transição foi apresentada; e, ainda, a melhor identificação a partir da transição de formantes foi para as bilabiais e em seguida para as alveolares. Tudo isso sugere que o *burst* do /k/ é uma pista robusta para a identificação desse ponto de articulação, enquanto a transição é uma pista mais robusta para o ponto bilabial e alveolar. Quanto ao contexto da vogal, os dados indicam que as pistas do *burst* foram mais eficientes com vogais anteriores e que a transição foi uma pista mais robusta com vogais posteriores.

Considerando todas as contribuições dos estudos citados, conclui-se que a informação sobre o ponto de articulação das plosivas é proveniente de várias propriedades acústicas que, perceptivamente integradas, fornecem pistas mais robustas do que relações tão específicas, como, por exemplo, a equação de *locus*. Além disso, o peso de cada pista acústica para a percepção do ponto de articulação parece variar para os diversos pontos de articulação e diferentes contextos de vogal, não sendo possível estabelecer uma regra genérica a respeito da participação isolada do *burst*, *locus* ou transição de F2 e F3.

1.4.3 Produção das plosivas por sujeitos com deficiência auditiva

Monsen (1976) argumenta sobre a importância de se estudar as alterações de fala de sujeitos com deficiência auditiva, não apenas a partir de avaliações

perceptivas que julgam as produções alteradas, mas também da análise acústica dessas produções. Estudou a produção de 37 sujeitos com deficiência auditiva e 6 sujeitos com audição normal no que se refere à implementação do contraste de vozeamento e realizou a medida de *VOT* para as seis plosivas do inglês. Os resultados mostram que, de maneira geral, os sujeitos com maior perda auditiva tendem a diferenciar menos os fonemas vozeados do que os não-vozeados, embora alguns sujeitos, mesmo com perdas auditivas profundas, possam implementar diferenças no *VOT* de forma similar a que ocorre com os sujeitos com audição normal. Os sujeitos que produziram apropriadamente a distinção entre as plosivas vozeadas e não-vozeadas (no caso do inglês, não-aspiradas e aspiradas) foram aqueles que apresentaram melhor inteligibilidade de fala.

Segundo o autor, a relação entre melhor inteligibilidade e implementação adequada do contraste de vozeamento nas plosivas não se deve apenas à melhor produção das plosivas. Se esses indivíduos são capazes de produzir adequadamente a diferenciação desses fonemas que exigem uma precisa coordenação de gestos glóticos e supraglóticos, provavelmente são capazes de produzir adequadamente outros aspectos da fala. Quando houve falha na diferenciação das vozeadas e não-vozeadas, constatou-se a tendência da realização de uma única variante para as duas plosivas de mesmo ponto, sendo que para as bilabiais e alveolares a tendência encontrada foi a produção do som não-aspirado e para as velares do som aspirado. Quanto à variação das medidas de *VOT* em relação ao ponto de articulação, os resultados comprovam a mesma tendência descrita na literatura, com valores maiores para as velares, seguida das alveolares e das bilabiais, mesmo para os sujeitos que não apresentaram diferenças entre os vozeados e não-vozeados.

A literatura revela que as alterações relacionadas à organização temporal da fala são muito freqüentes nas produções de sujeitos com deficiência de audição, tais como taxa de elocução lentificada, prolongamento excessivo de segmentos, inserção de pausas quando sintaticamente não caberia uma fronteira, dificuldades em marcar temporalmente sílabas tônicas entre outras (Osberger, Levitt, 1979). Segundo McGarr, Löfqvist (1988), durante a produção de consoantes não vozeadas do inglês, uma série de eventos articulatórios ocorre simultaneamente no

trato vocal e os eventos laríngeos e supralaríngeos devem estar coordenados. McGarr, Löfqvist (1982) estudaram as produções de obstruintes vozeadas e não vozeadas (/p/ /b/ /t/ /d/ e /s/ /z/) do inglês por três sujeitos com deficiência auditiva profunda e encontraram similaridades e diferenças em relação à produção de sujeitos com audição normal. Esses três sujeitos diferiam quanto à inteligibilidade da fala: a fala do sujeito 1 foi classificada como inteligível; a do sujeito 2 como difícil de entender e a do terceiro como pouco inteligível. A atividade da laringe foi monitorada pela transiluminação¹³ por meio de fibroscópio flexível inserido pelo nariz. Os padrões temporais da articulação oral foram gravados a partir de eletrodos posicionados no lábio inferior e superior que identificavam o início e fim das constrições de lábios e língua contra o palato.

McGarr, Löfqvist (1982) descrevem o padrão encontrado para sujeitos com audição normal, sendo que, na produção de uma consoante plosiva não-vozeada, pode-se observar o gesto de abdução/adução, o que não ocorre nas plosivas vozeadas. O padrão de organização temporal entre os eventos glóticos e supraglóticos revelou-se da seguinte forma: para as plosivas não-vozeadas, um pico de abertura glótica ocorria na soltura oral. Alguns dos sujeitos com deficiência auditiva preservaram o padrão correto de organização temporal interarticulatória entre a laringe e os articuladores superiores. Em outros casos, os falantes com deficiência auditiva omitiam o gesto de abdução/adução para as consoantes não-vozeadas ou produziam tal gesto quando ele não era requerido. Os sujeitos 2 e 3, com inteligibilidade de fala parcial e com fala pouco inteligível respectivamente, demonstraram as maiores diferenças entre vários padrões descritos para sujeitos com audição normal. A duração da oclusão revelou-se consideravelmente mais longa. As medidas de duração da oclusão, que tipicamente é menor para as plosivas vozeadas nas produções de sujeitos com audição normal, ficaram sobrepostas para as produções dos sujeitos com deficiência auditiva. As medidas de VOT também apontaram para a não-diferenciação das vozeadas e não-

¹³ Transiluminação da laringe ou foto-glotografia é um método indireto para registrar a abertura glotal. É usado para registrar a abdução e adução da glote durante a produção de consoantes não-vozeadas e pode também ser utilizada para examinar as vibrações glotais. <http://www1.ldc.lu.se/logopedi/departament/andy/Procedures/Transillumination> (fev. 2003) .

vozeadas para os sujeitos 2 e 3, para os pares /p x b/ e /t x d/ que apresentaram valores médios e desvio padrão (entre parênteses), respectivamente, de 8 (3,9) ms e 11 (3,3) ms e 20 (3,0) ms e 21 (3,5) ms para o sujeito 2 e 29 (6,9) ms e 25 (6,8) ms e 69 (6,3) ms e 59 (6,7) ms para o sujeito 3. Embora os autores não tenham apresentado análise estatística desses dados, pode-se notar que as diferenças de *VOT* entre os pares vozeado/não-vozeado são bem menores que aquelas verificadas nas produções do sujeito com audição normal. Observe-se que os valores para os sujeitos com audição normal foram de 84 (8,1) ms e 15 (3,5) ms para a oposição /p x b/ e 121 (3,8) ms e 23 (4,3) ms para a oposição /t x d/.

A incoordenação entre os gestos laríngeos e supralaríngeos e a conseqüente não diferenciação no *VOT* entre as vozeadas e não-vozeadas podem explicar a freqüência em que o contraste vozeado e não-vozeado é produzido inadequadamente na fala de indivíduos com deficiência auditiva. Em relação à descrição acústica, explica as alterações encontradas no *VOT* (Monsen, 1976; Madureira et al, 2002).

Outros estudos sobre a produção de fala de sujeitos com deficiência auditiva buscam investigar aspectos da coarticulação: pesquisas acústicas e articulatórias revelam que indivíduos com perdas auditivas profundas apresentam em sua fala menos efeitos do contexto se comparados a indivíduos com audição normal.

Tye-Murray (1987) verifica os efeitos da coarticulação na fala de indivíduos com alteração auditiva. A partir da cinefluorografia, investiga se as posturas de fechamento dos lábios (plosivas bilabiais) e do dorso da língua (plosivas velares) são afetadas pelo contexto de cinco vogais diferentes. Sua intenção é descobrir se o que é proposto pela literatura procede, ou seja, se a fala de sujeitos com deficiência auditiva consiste numa série de eventos isolados, porque eles não movem continuamente seus articuladores de uma posição para outra como fazem os sujeitos com audição normal (Öhman, 1966). A autora trabalhou com a produção de cinco sujeitos com perdas de audição profunda (média das freqüências da fala maior que 90 dBNA). Os resultados mostram que, para os dois sujeitos com o pior nível de inteligibilidade de fala, não ocorreram influências do contexto de vogal sobre as posturas de fechamento, ou seja, como a literatura já

previa, esses falantes produzem fala como uma série de eventos separados. Ela sugere, então, que, além dos métodos de reabilitação utilizados, sejam consideradas as consequências da ausência de informação auditiva durante os primeiros anos de aquisição da linguagem.

Em estudo posterior, Tye-Murray (1991) investiga o deslocamento da língua durante a realização de diversos gestos de abertura, na produção de *beet*, *dean*, *ban*, *din*, *bob* e *tune*. Os sujeitos eram três adultos do sexo masculino com deficiência auditiva adquirida no primeiro ano de vida e dois com audição normal, do sexo feminino. Ela observou que sempre ocorre o deslocamento da língua durante os gestos de abertura, de modo que não se pode caracterizar a fala dos sujeitos com deficiência auditiva por esse não-deslocamento. Além disso, ela verificou que todas as trajetórias foram similares, não importando a variação das vogais, com exceção das sílabas iniciadas pela alveolar para um dos sujeitos com deficiência auditiva. Tais trajetórias iguais ou estereotipadas sugerem que alguns falantes com deficiência auditiva contraem os músculos da língua de tal maneira que o corpo da língua se movimenta de forma similar para todas as vogais. Alguns anteriorizam suas vogais e outros posteriorizam. Com relação ao deslocamento da mandíbula, sua pesquisa sugere que não há um deslocamento excessivo isoladamente, mas um deslocamento excessivo da mandíbula em relação a um deslocamento mais restrito da língua. Esse posicionamento mandibular, associado a um movimento de língua estereotipado, pode explicar a produção de vogais, que soam como neutras, por sujeitos com deficiência auditiva.

Waldstein, Baum (1991) investigam a extensão da coarticulação antecipatória da fala de nove sujeitos com perda auditiva profunda, congênita ou adquirida na primeira infância, dos quais cinco na faixa etária dos 7 anos e quatro na de 10 anos, comparando os resultados com os de um grupo controle de crianças com audição normal. Segundo os autores, a dificuldade em produzir efeitos da coarticulação de forma apropriada pode ser um fator que contribui para a redução na inteligibilidade de fala geralmente atribuída a indivíduos com perdas auditivas. A integração ou co-produção de pistas adicionais à identidade do fonema nos segmentos adjacentes facilita a sua identificação. Os autores estudaram esses

efeitos na produção de sílabas CV [ti, tu e ki, ku], utilizando os parâmetros acústicos da duração dos segmentos consonantais (em ms) e da frequência do centróide medido em dois pontos da consoante e verificaram que ocorrem efeitos antecipatórios na fala dos sujeitos com deficiência auditiva, embora sejam menores do que os observados nas produções dos sujeitos com audição normal. Os sujeitos deste estudo, apesar de apresentarem perdas auditivas profundas, tinham um nível de inteligibilidade de fala bastante alto, o que sugere que a inteligibilidade de fala pode estar relacionada à presença de efeitos de coarticulação antecipatória que nem sempre dependem apenas do grau de perda auditiva.

Em outro estudo, os mesmos autores (Baum, Waldstein, 1991) trabalham com a coarticulação perseverativa, na produção dos mesmos sujeitos da pesquisa anterior, agora analisando a produção daquelas consoantes em posição final da sílaba [ti, tu, it, ut e ik, uk]. Eles mostraram que, na produção dos dois grupos de sujeitos, com deficiência auditiva e com audição normal, ocorreram evidências de coarticulação perseverativa. Entretanto, para o primeiro grupo, as diferenças encontradas nas frequências do centróide medido no início das consoantes foram menores do que as diferenças encontradas no outro grupo (a razão entre $F2[C]/F2[u]C$ para o grupo com deficiência auditiva foi de 1,20 e para o grupo com audição normal foi de 1,81). Do ponto de vista da medida de duração, não foram encontradas entre os dois grupos diferenças significativas, mas esse é um resultado diferente daquele a que chegaram para os mesmos sujeitos com relação à coarticulação antecipatória. Essa constatação, afirmam os autores revela as diferentes naturezas dos tipos de coarticulação: a antecipatória refletindo mais o planejamento e a perseverativa mais relacionada a restrições mecânicas. Em resumo para eles, ocorrem evidências de coarticulação antecipatória e perseverativa na fala de sujeitos com perda auditiva profunda, embora a magnitude dos efeitos da coarticulação esteja bem reduzida na produção desses sujeitos.

Segundo Okalidou, Harris (1999), temos escassez de estudos na literatura que aborda a questão da coarticulação na fala de indivíduos surdos. Esses autores também acreditam que há várias limitações nos trabalhos dos autores aqui citados, entre elas a heterogeneidade dos sujeitos com deficiência auditiva; a análise de

produções de sujeitos no início da adolescência, período em que alterações vocais próprias das modificações fisiológicas podem estar presentes; e principalmente a análise de uma porção limitada da produção – apenas a sílaba tônica, não levando em consideração a taxa de elocução como uma variável que interfere nos efeitos da coarticulação. Propõem a análise das medidas acústicas à luz da Fonologia Articulatória (Browman, Goldstein, 1986, 1989) para fazer inferências sobre padrões articulatorios da fala dos sujeitos com deficiência auditiva. Além disso, os autores argumentam a favor da realização de estudos de caso, porque é bastante conhecida a variação individual nas produções de sujeitos com deficiência auditiva. Com um *corpus* formado por sílabas sem sentido, com a estrutura dVd e bVb, sendo as vogais /i/, /u/, /a/, incluídas na frase-veículo *a _____ again*, dez repetições de cada frase foram gravadas em duas taxas de elocução: normal e rápida. Antes de iniciar as gravações, os sujeitos leram as frases impressas em sua taxa de elocução normal para que essa fosse registrada com ajuda do metrônomo. Para gravar as repetições com taxa de elocução rápida, o examinador ajustou o metrônomo para uma velocidade 25% maior que a utilizada para a taxa de elocução normal. Os sujeitos passaram por um período de treinamento antes da gravação, de forma a ajustarem suas produções aos parâmetros impostos pelo metrônomo. Foi utilizado um metrônomo digital, com sinais sonoros e visuais, para que os sujeitos com deficiência auditiva se apoiassem nos sinais luminosos. Durante as gravações, a taxa de elocução de cada sujeito foi periodicamente verificada em relação ao padrão inicialmente estabelecido. A análise acústica constou de medidas de (1) duração da vogal neutra (ou *schwa*) no início da frase, da oclusão da consoante, da sílaba tônica, da vogal tônica e da frase e (2) medida da frequência do segundo formante (F2) em diversos pontos do enunciado: no início meio e fim da vogal neutra, no início da transição e no meio da vogal tônica. No caso dos três sujeitos com deficiência auditiva, as durações de todos os segmentos foram significativamente mais longas para a taxa de elocução normal. Já, ao se comparar as medidas de duração dos segmentos entre os grupos, na taxa de elocução normal, pôde-se observar que todos os segmentos (*schwa*, oclusão, vogal tônica e sílaba) produzidos pelos sujeitos com deficiência auditiva apresentaram maior duração em relação aos produzidos pelos sujeitos com audição normal. Comparando a duração entre os grupos, considerando a taxa

de elocução normal para os sujeitos com audição normal e a taxa de elocução rápida para o grupo com deficiência auditiva, dois sujeitos com deficiência auditiva apresentaram durações de sílaba e frase semelhantes aos sujeitos com audição normal e um deles ainda apresentou segmentos um pouco mais alongados.

A comparação entre as durações dos dissílabos (*schwa* + sílaba) para cada sujeito nas duas taxas de elocução e nos dois contextos de consoante, bilabial e alveolar, não revelou diferenças para nenhum sujeito em nenhuma das condições, levando os autores a concluir que nesse estudo, nenhum sujeito apresentou diferenças de duração em virtude do contexto consonantal, considerando os sujeitos com audição normal e com deficiência auditiva. Em relação às variações de F2, os resultados apontam para uma menor variação nas produções dos falantes com deficiência auditiva, considerando a medida da frequência de F2 no centro das diversas vogais, o que está em conformidade com outras pesquisas que demonstram que as vogais produzidas por sujeitos com deficiência auditiva apresentam-se de certa forma neutralizadas. Em relação aos efeitos da consoante sobre o *schwa*, todos os sujeitos com audição normal apresentaram médias significativamente diferentes para a medida de F2. Para os sujeitos com deficiência auditiva o mesmo não foi observado, com exceção de um sujeito que apresentou diferença na frequência de F2 nos contextos bilabial e alveolar. No que se refere aos efeitos da consoante sobre o F2 medido no meio da vogal tônica, na produção dos sujeitos com audição normal, todos os valores de F2 no contexto alveolar com a vogal /u/ foram mais altos em relação ao contexto bilabial, indicando uma anteriorização da constricção do /u/ no contexto alveolar. As produções de um dos sujeitos com deficiência auditiva seguiu esse mesmo padrão e, para os outros dois, valores inversos foram encontrados, ou seja, valores de F2 mais altos para o /u/ subsequente às plosivas bilabiais. As medidas de F2 logo após a soltura das plosivas mostrou que, para os sujeitos com audição normal as médias dos valores de F2 foram significativamente diferentes entre os pares de dissílabos com as vogais posteriores, /ɯ/ e /u/. Para os sujeitos com deficiência auditiva, F2 só apresentou diferenças significativas nos contexto bilabial x alveolar, com a vogal posterior /ɯ/, quando a língua e a mandíbula estão ambas abaixadas. De forma geral, os autores concluem que os efeitos da coarticulação estão menos presentes

nas produções dos sujeitos com deficiência auditiva se comparadas com as produções dos sujeitos com audição normal. A hipótese de que em uma taxa de elocução maior a coarticulação estaria mais evidente nas produções do primeiro grupo não se confirmou a partir das medidas de duração e principalmente de F2. As trajetórias de F2, para esse grupo, tendem a ser planas, enquanto que, para o grupo com audição normal, observa-se uma variação maior. Os autores sugerem que os resultados encontrados nas produções dos sujeitos com deficiência auditiva apontam para um padrão de co-produção da consoante-vogal no qual a precisão da vogal e da consoante está comprometida resultando em gestos articulatórios menos diferenciados. Segundo eles, a produção das consoantes estaria mais comprometida do que a das vogais, uma vez que nas extremidades dos segmentos vocálicos observa-se pouca influência da consoante sobre a vogal indicando pouca coarticulação.

1.4.4 Percepção das plosivas por sujeitos com deficiência auditiva

Perdas auditivas neurossensoriais são sempre acompanhadas de redução na compreensão de fala, sendo a redução na audibilidade da fala em conversação espontânea o principal fator de dificuldade. O uso da amplificação como forma de restaurar a audibilidade é um recurso sempre utilizado na reabilitação de pessoas com deficiência auditiva (Ewing, Ewing, 1971; Sanders, 1977; Ling, 1978, 1989; Pollack, 1985; Boothroyd, 1982; Estabrooks, 1994; Bevilacqua, Formigoni, 1997; Turner, Brus, 2001, Mello e Novaes, 2001).

Dubno et al (1989) estudaram a identificação auditiva do ponto de articulação das plosivas por sujeitos com deficiência auditiva a partir de pistas espectrais do *burst*. Segundo os autores, a identificação fica mais comprometida à medida que a perda auditiva é mais acentuada nas frequências agudas, apontando para a importância da audibilidade de determinadas regiões de frequências. Para sujeitos com perdas planas, desde que a audibilidade dos sons de fala estivesse garantida, a identificação do ponto de articulação das plosivas ficou mais preservada. Entretanto, existem evidências de que a redução da audibilidade não é o único

fator que dificulta o reconhecimento de fala. Erber (1981) afirma que crianças com perda auditiva nunca conseguem discriminar freqüências tão bem quanto crianças com audição normal. A habilidade em discriminar freqüências está bastante associada à habilidade de perceber sons de fala, especialmente no que se refere à identificação dos espectros dos formantes dos fonemas vozeados.

Hall, Wood (1984) afirmam que os limiares de diferenciação de freqüência (*difference limen* para freqüências) são maiores para sujeitos com deficiência auditiva. Além disso, ressaltam que o limiar pode variar em virtude da duração do sinal. Se o estímulo tiver duração igual ou maior que 200 ms, tal limiar de diferenciação de freqüência não sofre alterações, mas deteriora-se significativamente para estímulos com duração menor que 200 ms. Investigando a interação de fatores na discriminação de freqüências para sujeitos com perda auditiva, os autores compararam os resultados de uma tarefa de identificação de diferenças de freqüências, usando tons puros de 500 Hz e 2000 Hz com duração de 200, 50, 20, 10 e 5ms, realizada por um grupo de 10 sujeitos com audição normal e 10 sujeitos com deficiência auditiva moderada. Os limiares de diferenciação de freqüência considerando o estímulo de 200 ms para o grupo de sujeitos com audição normal foi de 1,9 Hz para 500 Hz e 4,4 Hz para 2000 Hz. Houve aumento no limiar de diferenciação à medida que a duração dos segmentos diminuía e a freqüência do estímulo aumentava, chegando até 27 Hz para estímulos de 500 Hz com duração de 5 ms, e até 48 Hz para estímulos de 2000 Hz com a mesma duração. Os resultados dos sujeitos com deficiência auditiva foram muito variáveis: que alguns apresentaram resultados próximos dos sujeitos com audição normal e outros limiares de diferenciação de freqüência mais elevados. De forma geral, todos os sujeitos com deficiência auditiva apresentaram limiares de diferenciação de freqüência mais elevados que os sujeitos com audição normal, os aumentaram com a diminuição da duração e aumento da freqüência, indicando que a relação entre diferenciação de freqüência, duração e freqüência do estímulo apresentou tendência semelhante à observada no grupo com audição normal.

Embora os resultados sobre a resolução de freqüências para tons puros sejam um indicativo das alterações impostas pela deficiência auditiva

neurossensorial, não podem ser transpostos diretamente para a identificação de diferenças em frequências em sons complexos como os da fala, com vários componentes de frequências e amplitudes, os quais variam em tempos que podem ser muito curtos, como, por exemplo, o ruído transiente das plosivas ou as transições observadas na vogal subsequente à plosiva (Wieringen, Pols, 1994).

Uma série de experimentos com fala natural e sintética foram relatados por Pickett, Danaher (1975), os quais tinham como objetivo pesquisar a percepção de fala por sujeitos com deficiência auditiva. No que se refere à fala natural, concluíram que os sujeitos com perdas auditivas neurossensoriais apresentavam: 1) melhor recepção residual para padrões de fala de baixas frequências, tais como vozeamento, murmúrio nasal e o primeiro formante das vogais; 2) maiores dificuldades foram verificadas para as pistas contidas nas frequências médias e altas, tais como o ruído transiente das plosivas, das consoantes fricativas e dos segundo e terceiro formantes das vogais, o que ratifica Miller, Nicely (1955) que, entre outros, demonstraram por meio de experimentos perceptivos com fala filtrada, que as informações relativas ao ponto de articulação estão principalmente localizadas em frequências em torno e acima de 1000 Hz e que as pistas para o contraste de vozeamento estão presentes em frequências mais baixas. Além disso, Pickett, Danaher (1975) verificaram que as consoantes em posição inicial são mais bem percebidas que em posição final; a discriminação de vogais é melhor que de consoantes; que a discriminação do segundo formante é melhor quando a frequência de F2 é mais baixa, ou seja, em vogais posteriores em contraste com anteriores. Para testar a discriminação de formantes, os autores utilizaram fala sintética simulando dois formantes. Após vários experimentos, verificaram que mesmo sujeitos com perdas neurossensoriais podem apresentar habilidades de discriminar informações de frequências de formantes embora, uma variação individual muito grande tenha sido verificada. A discriminação de formantes é, em geral, melhor para formantes de baixa frequência em contraste com os de alta frequência. Entretanto, a reduzida habilidade de discriminação de frequências pode estar relacionada, em parte, aos altos níveis de intensidade necessários para tornar o estímulo audível, por provocarem três tipos de mascaramento: 1) um que se espalha para frequências mais altas, reduzindo a discriminação de frequências

médias quando elas co-ocorrem com um formante baixo (os autores chegaram a essa conclusão com um experimento no qual a variação de F2 foi testada isoladamente e na presença de F1); 2) um central que ocorre quando o estímulo é apresentado dicoticamente; e 3) um retroativo, que afeta a discriminação do estímulo que antecede a vogal. Os mesmos efeitos de mascaramento foram encontrados para os sujeitos com audição normal quando eles foram testados com estímulos em altas intensidades.

As pistas acústicas relacionadas ao ponto de articulação das plosivas são particularmente suscetíveis às perdas auditivas neurosensoriais, conforme afirmam Van Tansell et al (1982). Consideram eles que isso possa ser devido simplesmente à não-audibilidade das pistas e também às distorções impostas no sinal de fala pela própria perda auditiva: baixa resolução de frequência, crescimento da sensação de intensidade alterado, entre outras. Sintetizaram um contínuo com cinco formantes, simulando plosivas vozeadas (/ba/ para /da/, para /ga/) e, em seguida, editaram o sinal de forma a preservar apenas os dois primeiros pulsos glotais subseqüentes ao *burst*. O objetivo foi avaliar a capacidade de identificação do ponto de articulação das plosivas por nove sujeitos com perdas auditivas de leve a moderada, considerando apenas as informações contidas nos primeiros milissegundos da sílaba (Stevens, Blumstein, 1978 ; Blumstein, Stevens, 1979; Kewley-Port, 1982; Kewley-Port, 1983; Kewley-Port et al, 1983). Entretanto, concluíram que, para esse grupo de sujeitos com perdas auditivas moderadas, os resultados não diferiram dos resultados de um grupo de seis sujeitos com audição normal, quando os estímulos estavam em uma intensidade audível. Contudo, esse achado não pode ser generalizado para perdas auditivas mais acentuadas.

Ochs et al (1989) realizaram um experimento para investigar a correlação entre a habilidade de identificar diferenças em frequências (*difference limen*) e a percepção do ponto de articulação das plosivas. Se o limiar de diferenciação de frequências de um ouvinte for maior que a diferença entre a frequência no início da transição e na porção estacionária da vogal subseqüente à plosiva (ΔF), eles afirmam que a pista da transição não poderia ser usada por esse ouvinte para identificação do ponto de articulação, e ele teria de se apoiar nas pistas do *burst*.

Realizaram testes de identificação de sílabas sintetizadas /bi/ /di/ e /gi/ com cinco formantes. O formante 1 apresentava uma transição fixa para os três pontos de articulação, iniciando em 276 Hz e terminando em 307 Hz. O F2 foi manipulado de forma a ser apresentado com a transição apropriada para cada ponto de articulação (/bi/ = de 1648 a 2070 Hz; /di/ = 1903 a 2070 Hz e; /gi/ = de 2206 a 2070 Hz) e com uma transição neutra. A duração da transição também variou em 10, 19, 28 e 44ms. Para o terceiro formante, a transição era plana, mas a frequência no início da vogal variava de acordo com o ponto de articulação. Testaram sujeitos com audição normal (4), com deficiência auditiva (4) de configuração descendente (com limiares tonais normais até 1 kHz) e sujeitos com audição normal (4) submetidos a mascaramento de forma a simular a perda auditiva do grupo com deficiência auditiva. Verificaram que o aumento da duração determinou melhor identificação das três sílabas para os três grupos. Entretanto, não ocorreram evidências de que a transição de F2 tenha sido utilizada para a identificação do ponto de articulação, já que os estímulos com a transição de F2 não-neutra não foram mais bem identificados do que aqueles nos quais F2 era plano. Somente no caso da sílaba /bi/ é que a transição não-neutra mais longa (44ms) de F2 resultou em melhor identificação do ponto de articulação [$F(2,9)=3.81$, $p < 0,06$]. Note-se que esta era a transição com maior ΔF entre as três sílabas - (+ 422) contra (- 167) para /di/ e (-136) para /gi/.

Na seqüência eles realizaram um segundo experimento para investigar se o limiar de diferenciação de frequência seria o fator determinante para o não-uso da transição de F2 para a identificação dos contrastes de ponto. Eles apresentaram para os sujeitos o F2 isolado de cada sílaba com diferentes transições que variavam entre 3 e 400 Hz. Os resultados mostraram, todavia, que o limiar de diferenciação de frequência para a simulação da transição de F2 não diferiu muito para os três grupos, ficando em torno de 100 Hz, o que não explica os resultados do experimento 1. Os autores consideram que esses resultados aparentemente conflitantes podem ser atribuídos a questões de natureza metodológica, uma vez que os estímulos utilizados para a determinação do limiar de diferenciação de F2 não reproduziam fielmente as condições daquele usado para a identificação dos pontos de articulação, ou seja, apresentavam apenas um formante (F2) e não

incluíam a porção do *burst*. Os autores alertam para a importância da utilização de estímulos complexos que se assemelhem mais à fala na realização de avaliação das habilidades psicoacústicas quando o objetivo for buscar explicações para as alterações na percepção de fala em sujeitos com deficiência auditiva.

Tsui, Ciocca (2000) estudaram a percepção do ponto de articulação e da aspiração das plosivas do cantonês¹⁴ em posição inicial por sujeitos com deficiência auditiva e audição normal. Selecionaram 14 sujeitos com deficiência auditiva bilateral severa e 14 com audição normal com idade média de 15 anos, e manipularam 6 palavras monossilábicas iniciadas pelas seis plosivas variando a duração do VOT. Para as plosivas não aspiradas, um intervalo de silêncio foi adicionado de forma que o VOT ficasse com a mesma duração que o seu par aspirado. No caso da aspirada, uma parte do sinal era retirada para que a medida do VOT ficasse igual a de seu par não-aspirado. Dessa forma, cada plosiva apresentava duas condições: uma com o VOT curto e outra com o VOT longo. Os pesquisadores verificaram que a presença do ruído da aspiração é uma pista mais robusta para a percepção de uma plosiva aspirada do que o aumento do VOT tanto para o grupo de sujeitos com deficiência auditiva quanto para o outro grupo. Em relação ao ponto de articulação, os resultados mostraram que para os sujeitos com audição normal a manipulação da medida do VOT não foi suficiente para interferir na identificação do ponto de articulação. Para eles, as pistas contidas no *burst* e na transição dos formantes das palavras monossilábicas usadas como estímulo foram suficientes para desfazer os efeitos do VOT manipulado. Entretanto, para os sujeitos com deficiência auditiva de forma geral, a percepção de ponto de articulação foi muito alterada. As alterações na resolução de frequência são, para os autores, a principal causa da dificuldade de perceber o ponto de articulação. Note-se que os sujeitos desse estudo apresentavam perda auditiva severa, o que corresponde a limiares tonais maiores que variaram de 71 a 90dBNA (média das frequências de 500, 1000 e 2000 Hz).

A resolução temporal também se encontra alterada em sujeitos com

¹⁴ Cantonês é uma das principais línguas do território chinês. As plosivas do cantonês apresentam 3 pontos de articulação – bilabial, alveolar e velar - e distinguem-se pela aspiração e não pelo vozeamento.

deficiência auditiva. O limiar para detecção de um intervalo de silêncio (*gap detection*) é maior para sujeitos com deficiência auditiva se comparado ao limiar encontrado em sujeitos com audição normal (Dorman, Hannley, 1985). Esse limiar, para sujeitos com audição normal, foi de 13,3ms para frequência de 500Hz e de 7,4ms para 4000Hz. Já para um grupo de sujeitos com deficiência auditiva, as médias foram de 22,9 e 11,6 respectivamente. O limiar de diferenciação de duração também está elevado em sujeitos com deficiência auditiva. Enquanto esse fica em torno de 18,8 ms para um sinal de 500 Hz e 17,9 para um sinal de 4000 Hz para sujeitos com audição normal, a média para sujeitos com deficiência auditiva fica em torno de 48 e 41,8 ms respectivamente. Observa-se uma grande variação entre os sujeitos com deficiência auditiva quanto às habilidades de identificação do intervalo de silêncio, sendo que alguns deles chegam a apresentar resultados semelhantes aos de sujeitos com audição normal. Apesar de, geralmente, a resolução temporal estar alterada para sujeitos com deficiência auditiva, há dados que sugerem que ela fica mais preservada do que a resolução de frequência e, portanto, é provável que indivíduos com perdas auditivas severas e profundas apóiem-se mais em pistas de duração do que espectrais.

Parady et al (1981) investigaram a utilização da duração do *VOT* para identificação do contraste de vozeamento por crianças com deficiência auditiva. Estímulos sintéticos foram apresentados num contínuo de /da/ até /ta/ para 21 crianças com diversos graus de perdas auditivas neurossensoriais e 10 crianças com audição normal. As crianças com perdas moderadas apresentaram habilidades semelhantes aos sujeitos com audição normal, ou seja, diferenciaram /d/ de /t/ a partir de um *VOT* de cerca de 28 ms. Para o grupo com perda severa (dez sujeitos), os resultados foram muito variáveis: cinco sujeitos atingiram 100% de identificação de /ta/ com um *VOT* acima de 35 ms. Já outros três sujeitos só identificaram /t/ em 70%, quando o *VOT* era o mais longo do contínuo, ou seja, 60 ms, e os dois últimos sujeitos só atingiram 40% de respostas corretas para /t/ mesmo com o *VOT* mais longo. No grupo das que possuíam deficiência auditiva profunda, os três sujeitos apresentaram resultados muito diferentes: um deles, cuja deficiência auditiva foi adquirida com 3 anos de idade, foi capaz de discriminar /d/ de /t/ como ocorreu com o grupo com audição normal; outro foi capaz de

discriminar a diferença de duração mas não de identificar os fonemas e o terceiro não foi capaz nem de discriminar nem de identificar os fonemas.

Complementando esse trabalho de Parady et al (1981), Johnson et al (1984) realizaram um experimento semelhante, mas que incluía os pontos de articulação bilabial e velar, além do alveolar, já que tinham como objetivo avaliar a influência do ponto de articulação na identificação do vozeamento. Verificaram os pesquisadores que as fronteiras do *VOT* para os três pontos de articulação foram semelhantes para o grupo de sujeitos com audição normal e para o grupo de crianças com perdas auditivas leves. Já para os grupos de sujeitos com deficiências auditivas moderadas e severas, a habilidade de identificar o ponto de articulação pela pista do *VOT* ficou bastante prejudicada em relação aos dois primeiros grupos. Uma das explicações para esse resultado, segundo os autores, é o fato de a intensidade muito elevada de apresentação do estímulo para o grupo de sujeitos com perdas severas poder ter contribuído para a distorção do sinal e para a conseqüente dificuldade maior encontrada. Assim, além das alterações de resolução temporal decorrentes das alterações fisiológicas da cóclea, a alta intensidade com que sujeitos com perdas auditivas severas e profundas recebem o som de fala pode ser um fator adicional de distorção.

A percepção do vozeamento de plosivas em posição final de sílaba por 8 sujeitos com audição normal e 27 com deficiência auditiva, manipulando parâmetros acústicos como a duração da vogal, duração da oclusão, ruído da soltura da oclusão, murmúrio do vozeamento durante a oclusão, e transição da vogal anterior à plosiva foi investigada por Revoile et al (1982). Nessa pesquisa os autores concluíram que muitos dos sujeitos com perda auditiva apóiam-se mais na pista de duração do que em pistas espectrais para identificar vozeamento. Já para os sujeitos com audição normal, a manipulação dos parâmetros de duração não alterou a identificação desse contraste.

Revoile (1999) relata um experimento no qual testou 12 deficientes auditivos profundos, tendo verificado que 6 deles utilizavam a pista de duração da vogal para identificar o contraste de vozeamento em consoantes em posição final. Dos 6 que mostraram resultados pobres na identificação do contraste de vozeamento, 5

também apresentaram dificuldades em discriminar modificações na duração. Em contraste, um limiar de discriminação de duração normal (25ms - *Difference limen* re 200 ms) foi encontrado entre os que usavam a pista de duração da vogal para identificar o contraste de vozeamento.

Há ainda um outro aspecto que é tido como fator importante para a percepção das consoantes plosivas por sujeitos com deficiência auditiva: a relação entre a energia da consoante em relação à vogal subsequente, conhecida como relação CV. Picheny et al (1986) mostraram que a “fala clara” é mais facilmente percebida por sujeitos com alterações auditivas. Eles fizeram a análise acústica comparativa entre a “fala clara” e a “fala espontânea” e concluíram que a primeira caracteriza-se pela diminuição da velocidade, pela inserção de pausas entre as palavras, pelo alongamento da duração dos sons e pelo aumento da intensidade das plosivas. Um dos aspectos acústicos que se modifica na “fala clara” é a energia da consoante em relação à vogal seguinte, que ficaria mais saliente, ou seja, a relação de energia C/V aumenta, o que facilitaria a sua audibilidade e, portanto, aumentaria as condições de percepção para indivíduos com deficiência auditiva.

Uma série de trabalhos foi realizada para investigar não só o papel da energia relativa da consoante na percepção das plosivas por sujeitos com deficiência auditiva, mas também o modo como a manipulação dessa relação a partir de modificações na amplificação do sinal por aparelhos de amplificação sonora individual poderia melhorar a percepção de consoantes por sujeitos com deficiência auditiva. Para Freyman, Nerbonne (1989), por exemplo, a relação CV é expressa como a diferença em dB entre a energia da consoante e da vogal adjacente. Os autores estudaram a importância da relação CV para a inteligibilidade das consoantes /p/, /t/, /k/, /tʃ/, /ʃ/, /ʒ/, /ʒ/, /ʒ/, a partir da manipulação digital da relação CV das produções de 10 falantes, de forma a igualar todas as produções, tornando a diferença entre a energia da consoante e da vogal equivalentes em todas elas. Compararam a inteligibilidade das produções naturais dos 10 sujeitos com as sílabas manipuladas. Os estímulos foram apresentados para 50 sujeitos com audição normal. Tanto para os estímulos

naturais como para os estímulos manipulados, os resultados mostraram que a inteligibilidade teve pouca relação com a relação CV. Os melhores resultados na identificação das consoantes na condição natural foram para as produções do falante que apresentava a maior diferença de energia entre a consoante e a vogal, ou seja, a menor relação CV (média entre todas as consoantes de -19,1 dB (9,5)). As médias apresentadas para /p/, /t/, /k/ e desvio padrão para as produções dos 10 sujeitos foram respectivamente, - 8,9 (6,7) dB; -14,9 (2,2) dB e -14,8 (2,2) dB. Entretanto, não se observou constância nos dados em relação ao ponto de articulação: para dois sujeitos, o ponto bilabial foi o que apresentou a relação CV maior; comparando-se o ponto alveolar com o velar, para 6 sujeitos, o alveolar apresentou CV maior que o velar e, para outros quatro, o ponto velar é que apresentou a relação CV maior que a alveolar. Desse modo, dizem os autores que a relação entre a energia da consoante e da vogal provavelmente interfere na inteligibilidade não porque seja por si uma referência, mas porque, quando a consoante tem maior energia, sua audibilidade fica mais garantida, principalmente em situações de ruído, uma vez que a relação sinal-ruído (para a consoante) fica mais favorável.

Com o objetivo de avaliar a importância da relação de energia entre a consoante e a vogal para a identificação das plosivas /p/, /t/ e /k/, em sílabas com a vogal /ɪ/, Sammeth et al (1999) fizeram um trabalho no qual, em vez de aumentar a energia da consoante para diminuir a relação CV, diminuíram a energia da vogal. Escolheram esse procedimento para tentar esclarecer se a melhora da identificação com a diminuição da relação CV deve-se à relação em si ou à audibilidade da consoante, como sugerem Freyman, Nerbonne (1989). A vogal /ɪ/ foi selecionada tendo em vista que é a que apresenta a maior intensidade entre as vogais do inglês e, portanto, poderia ter sua amplitude reduzida sem que ficasse inaudível para os sujeitos com deficiência auditiva. Participaram do estudo dois sujeitos com audição normal e seis com perdas auditivas descendentes, de leve a moderada, sem história que sugerisse comprometimento retro-coclear. Os estímulos foram manipulados de forma a serem apresentados em três condições: apenas o *burst* (separado da vogal), a sílaba com a energia da vogal reduzida em 6 dB e com a energia da vogal reduzida em 12 dB. Os resultados mostraram que

não houve melhora na identificação das plosivas em virtude da redução da relação CV. Os autores concordam com Freyman, Nerbonne (1989) afirmando que, uma vez garantida a audibilidade da consoante, não existe o benefício da diminuição da relação em si. Ressalvam, porém, que dado o número limitado de sujeitos e de sons de fala estudados, não é possível afirmar com certeza que a diminuição da relação CV nunca contribui para a melhor identificação da consoante depois que o limiar de audibilidade foi alcançado. Uma outra constatação de importância nesse estudo refere-se à piora da identificação das plosivas pelos sujeitos com deficiência auditiva quando o *burst* foi apresentado isoladamente, o que indica que, para sujeitos com esse tipo de perda, a informação da transição para a vogal exerce um papel importante na identificação da plosiva.

Os estudos que investigam a percepção das plosivas por sujeitos com deficiência auditiva apresentam resultados muito variáveis. De forma geral, pode-se dizer que revelam que a percepção das plosivas é bastante alterada pela deficiência auditiva e que, em geral, tal dificuldade aumenta com o grau da perda. Um aspecto que chama a atenção é que, geralmente, variações individuais são referidas, levando a supor que a percepção dos sons plosivos, assim como de outros sons de fala, é afetada por diversos fatores, não estando apenas vinculada ao grau de perda. Outro fato a ser considerado é que a variação nos resultados para sujeitos com audição normal aponta para diferenças individuais na forma como os ouvintes se utilizam das pistas acústicas, mesmo para sujeitos que não apresentam deficiência auditiva.

2 IDENTIFICAÇÃO DAS PLOSIVAS

Como faz parte dos objetivos deste trabalho estudar a percepção das plosivas pelo sujeito com deficiência auditiva, é importante avaliar as habilidades de percepção das plosivas no mesmo contexto fonético considerado para elaboração do *corpus* a ser analisado acusticamente, ou seja, em posição inicial de palavras que contrastam apenas pela plosiva. Desse modo, realizamos um estudo complementar, objetivando elaborar um procedimento que possibilite a avaliação das habilidades de identificação dos contrastes de vozeamento e ponto de articulação das plosivas de sujeitos com deficiência auditiva de grau severo e profundo.

Muitos testes já foram desenvolvidos com a finalidade de avaliar o reconhecimento de palavras monossílabas e dissílabas, a identificação de sentenças, a discriminação e identificação de fonemas e aspectos suprasegmentais da fala. Alguns deles são específicos para avaliação de crianças ou idosos, para verificar alterações do processamento auditivo ou habilidades de percepção auditiva da fala em indivíduos com capacidade auditiva mínima (Mendel, Danhauer, 1997). Assim, também, diversos autores brasileiros realizaram trabalhos objetivando desenvolver testes para avaliar a percepção de fala no PB (Pupo, 1981; Rodrigues, 1981; Gama, 1994; Garcia, Pereira, Fukuda, 1999). E alguns procedimentos de avaliação da percepção auditiva foram especificamente implementados para avaliar as habilidades de percepção de fala em deficientes auditivos (Bevilacqua, Tech, 1996; Orlandi, 1996; Delgado, Bevilacqua, 1999).

Delgado, Bevilacqua (1999), elaboraram uma lista de palavras para avaliação da percepção auditiva de crianças com deficiência auditiva, e os resultados encontrados apontam para a diminuição do número de acertos com o aumento da perda auditiva e para diferenças entre os três grupos de crianças avaliadas, (com perdas moderadas, severas e profundas) quanto ao número de acertos por modo de articulação. Longone, Borges (1997, 1998) estudaram o reconhecimento de fala em 30 sujeitos com perda auditiva neurosensorial de configuração descendente. A análise das substituições dos fonemas observadas nas respostas dos sujeitos

submetidos ao teste de reconhecimento de palavras monossílabas revelou que o maior número de erros recaiu sobre os sons consonantais e, entre eles, sobre os fonemas fricativos e líquidos. Em geral existe uma concordância que deficientes auditivos com perdas de moderadas a profundas tem menor acesso a contrastes de fala e pistas espectrais para identificação de consoantes do que de vogais (Boothroyd, 1984; Revoile et al, 1991).

Segundo Boothroyd (1982), crianças com perda auditiva de até 60 dBNA (com auxílio de aparelhos de amplificação sonora individual e envolvidas em programas que enfatizam o desenvolvimento da função auditiva) podem ouvir todas as características da fala. Quando a perda está entre 60 e 90 dB, podem perceber auditivamente todas as pistas acústicas, com exceção daquelas relacionadas ao ponto de articulação. Com a perda maior que 90 dB, a maioria dos deficientes auditivos tem acesso apenas a pistas de ritmo e entonação e a alguma informação das vogais e do modo de articulação das consoantes vozeadas. Boothroyd (1985, 1996) aponta as vantagens da utilização de testes de escolha forçada para avaliar a percepção e produção de contrastes de fala em sujeitos com perdas auditivas de grau severo ou profundo. Segundo ele, a identificação ou reconhecimento de sons de fala, sem nenhuma restrição de contexto, pode ser muito difícil para esses sujeitos, e os resultados ruins não dão nenhuma informação adicional sobre sua habilidade em perceber, auditivamente, contraste de fala.

Para avaliar, em nossa pesquisa as habilidades de identificação e discriminação desses contrastes pelo sujeito com deficiência auditiva cujas produções foram submetidas à análise acústica neste trabalho, foi elaborado um teste de discriminação de palavras¹⁵, que utiliza um programa de computador¹⁶ o qual permite a apresentação simultânea de:

¹⁵ No procedimento aqui descrito a tarefa consiste em identificar a palavra falada entre quatro opções. Nesta tarefa estão envolvidas habilidades de discriminação e identificação auditiva. Entretanto, neste estudo vamos nos referir a esta tarefa apenas com o termo discriminação de forma a diferenciá-la, posteriormente das tarefas de identificação descritas no capítulo 3.

¹⁶ O programa foi elaborado por Rogério Lugó, colaborador do LIAAC – Laboratório Integrado de Análise Acústica e Cognição – PUC/SP.

- arquivos de áudio (contendo o *corpus*¹⁷ elaborado para este estudo, composto pelas palavras *pata*, *bata*, *tata*, *data*, *cata* e *gata* inseridas em frase- veículo *Diga_____baixinho*), com as produções do sujeito com audição normal;
- arquivos de vídeo (figuras e frases impressas representando os arquivos de áudio - *visual basic*);
- registro das respostas e o tempo de resposta para cada item.

Foram selecionadas seis figuras para representar cada uma das palavras. *Gata* e *pata* foram representadas por figuras dos respectivos animais. A palavra *data* pela figura de um calendário no qual está sendo marcado um dos dias. *Tata*, abreviação do vocábulo tataravó, foi representada pela figura de uma mulher idosa e a palavra *cata* pela figura de um goleiro catando a bola no gol. Finalmente, *bata* foi representada pela figura de um martelo batendo num prego.

O programa conta com um módulo de apresentação dos itens do teste, que estão representados pela figura e pela frase impressa abaixo de cada figura, de forma que se possa garantir o conhecimento dos vocábulos e sua associação à figura correspondente antes do início do teste (Figura 9).

¹⁷ Os procedimentos de elaboração e gravação do *corpus* e a caracterização dos sujeito com deficiência auditiva e audição normal que participaram da gravação estão no capítulo 3.

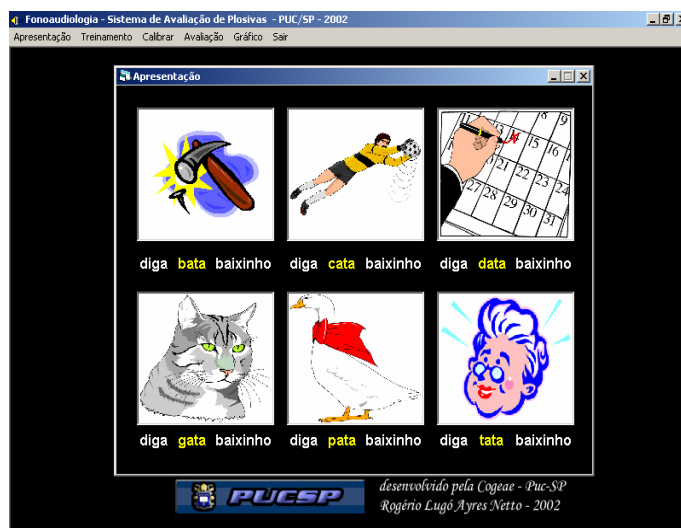


Figura 9 - Módulo de apresentação dos itens do teste

Também, antes do início do teste, o uso de um módulo de treinamento (Figura 10), no qual são utilizados outros vocábulos – *caixa* (pronunciado como [kajã]), *casa*, *capa* e *cara* - representados por figuras, tem a finalidade de garantir a compreensão da tarefa e o manuseio do programa.

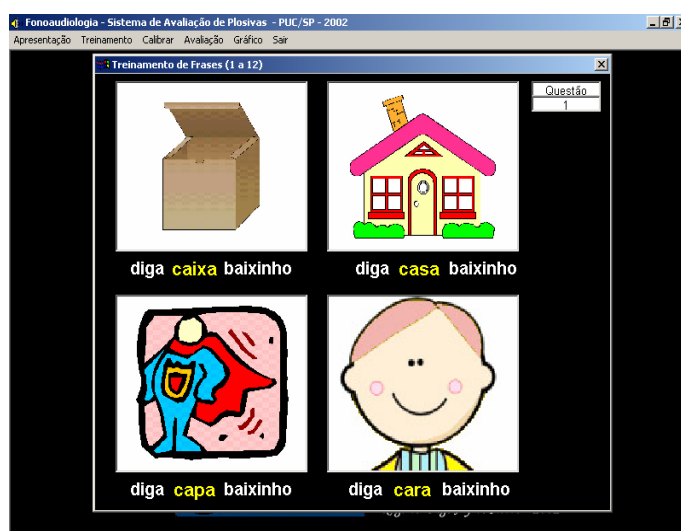


Figura 10 - Módulo de treinamento

Uma tela de calibração permite, com a geração de um tom puro de 1kHz, o ajuste da intensidade de apresentação do som, no caso de o teste ser aplicado em

cabine acústica, com o uso de um audiômetro. Todos os arquivos de áudio (dez repetições de cada vocábulo) foram equalizados em estúdio para evitar que diferenças entre eles pudessem vir a comprometer a audibilidade no caso dos sujeitos com deficiência de audição. Isso foi realizado de forma homogênea, ou seja, sem alterar as características da gravação original (intensidade relativa entre os sons de fala). O programa também permite o registro de dados de identificação do sujeito e das condições de apresentação do teste.

Quando se inicia o teste propriamente dito, o programa faz, a cada novo teste, a aleatorização da apresentação dos itens bem como do quadrante da tela do computador em que as figuras irão aparecer. As telas do teste contêm quatro itens, apresentados em três combinações, considerando o contraste de ponto de articulação e de vozeamento. Assim sendo, em uma tela é apresentado o contraste bilabial x alveolar nas ocorrências surda e sonora (*pata, bata, tata, data*, Figura 11); na outra tela, o contraste bilabial x velar nas ocorrências surda e sonora (*pata, bata, cata, gata*, Figura 12); e, numa terceira tela, o contraste alveolar x velar nas ocorrências surda e sonora (*tata, data, cata, gata*, Figura 13).

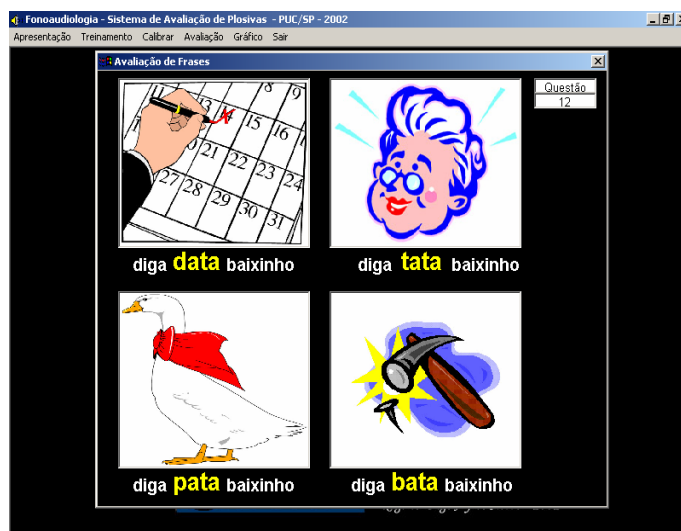


Figura 11- Oposição bilabial x alveolar nas ocorrências vozeada e não-vozeada

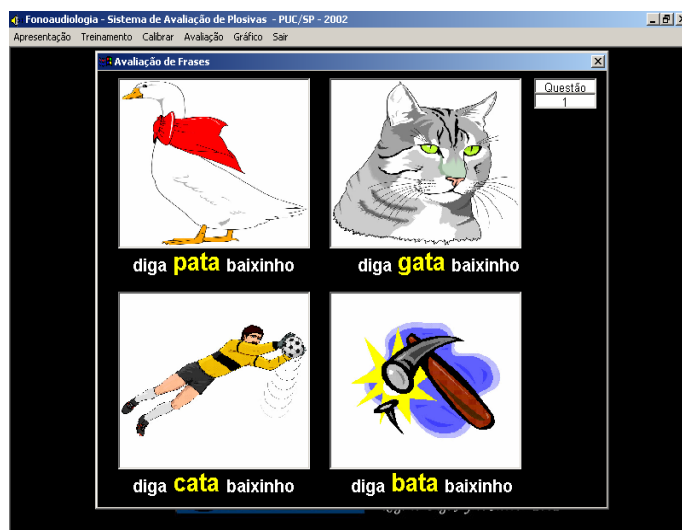


Figura 12 - Oposição bilabial x velar nas ocorrências vozeada e não-vozeada

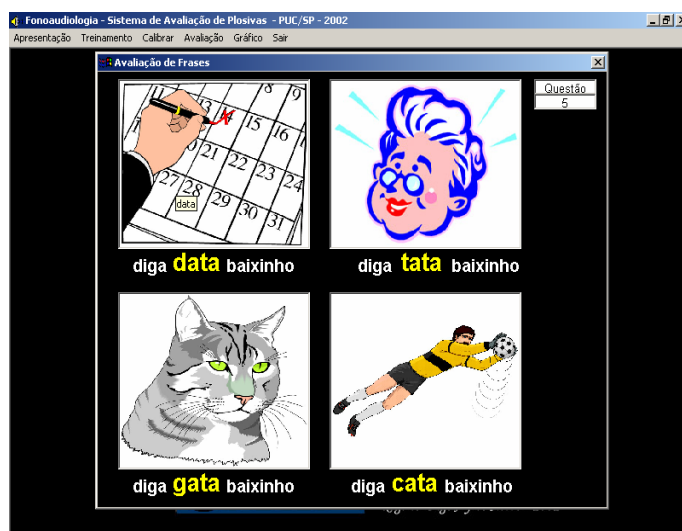


Figura 13 - Oposição alveolar x velar nas ocorrências vozeada e não-vozeada

Cada oposição é apresentada 10 vezes em cada combinação, resultando em 40 apresentações para cada agrupamento de palavras. O teste, portanto, consta de 120 apresentações. Exemplificando: a palavra *pata* aparece 10 vezes combinada com *bata*, *tata* e *data* e 10 vezes combinada com *bata*, *cata* e *gata*.

Cada palavra, no total do teste, é apresentada 20 vezes. A repetição dos itens e a aleatorização das apresentações são necessárias para minimizar a chance de que respostas aleatórias sejam tomadas como resultados.

O programa gera uma tela com o resumo dos resultados apresentados em tabelas segundo os três agrupamentos de palavras, as porcentagens e descrição de erros e acertos. Além da tela dos resultados, é gerado um relatório a cada avaliação contendo: item testado, grupo em que foi apresentado, arquivo de som utilizado, resposta dada, intervalo de tempo entre a apresentação e a resposta e descrição do erro quando ele ocorre.

2.1 Sujeitos sem queixas auditivas – metodologia

Com a finalidade de verificar a adequação do procedimento para a avaliação da percepção de fala de sujeitos com deficiência auditiva, este foi inicialmente aplicado a uma população de crianças sem queixas auditivas. O objetivo foi verificar a adequação do instrumento elaborado, tanto em relação aos itens (palavras e respectivas representações pictóricas), como em relação à forma de aplicação. Procurou-se investigar se a tarefa (escutar aquelas palavras e indicar com o *mouse* a figura correspondente na tela do computador) poderia ser realizada sem dificuldades por crianças e a partir de qual faixa etária.

Dessa forma, foram selecionadas 70 crianças de 6 a 8 anos e 5 meses, sem queixas auditivas ou referentes à linguagem, sem comprometimentos ou dificuldades de fala identificadas pelos professores. As crianças freqüentavam a última série do ensino infantil ou da primeira série do ensino fundamental, sendo 53 matriculadas em escola regular pública e 17 em escola particular. Quanto ao sexo 39, eram do feminino e 31 do masculino. Não foram incluídas na amostra quatro crianças que tiveram muita dificuldade em aprender a tarefa, seja em relação à compreensão da ordem, seja no que se refere à associação das palavras às figuras correspondentes. Além disso, uma outra criança não colaborou e uma outra não foi incluída por não estar bem fisicamente no momento da avaliação. Não foram eliminadas as 13 crianças que nunca haviam manipulado o *mouse* do

computador, tendo sido substituída essa prática pelo ato de apontar a figura na tela. Nesses casos, o examinador mantinha o indicador do *mouse* no centro da tela do monitor e só o movimentava na direção apontada pela criança após a clara indicação do item selecionado¹⁸.

Antes do início do teste, foram utilizados os módulos de apresentação e treinamento para garantir a compreensão tanto das palavras quanto da tarefa a ser realizada, bem como o reconhecimento das figuras correspondentes a cada vocábulo. O teste só foi efetivamente iniciado depois de a examinadora ter certeza de que a criança havia compreendido bem o que deveria ser feito, (por meio do bom desempenho no módulo de treinamento do programa) e de que ela associava cada vocábulo à figura correspondente (a criança reconhecia corretamente todos os vocábulos falados pela examinadora, apontando corretamente a figura no módulo de apresentação dos itens do teste). Cabe observar que, antes da apresentação, também era esclarecido para a criança que a palavra *tata* se referia à abreviação ou “apelido” da tataravó, momento em que se perguntava à criança se ela sabia o que era tataravó e se sabia o nome de sua tataravó. A grande maioria delas não teve nenhuma dificuldade em identificar corretamente os itens *pata*, *gata* e *data*. Quanto aos outros, algumas crianças identificaram-nos corretamente na primeira apresentação, mas para a grande maioria foi necessário que a examinadora os ensinasse, com ajuda de exemplos e perguntas como: *Quem cata a bola no gol? - Você sabe bater o prego com o martelo? – A tataravó é muito velhinha!*

Considerando o número de repetições necessárias e, com o objetivo de evitar o fator cansaço, foram realizadas pausas a cada 40 apresentações ou sempre que solicitado. O tempo de aplicação do teste variou, levando em média cerca de 20 minutos, incluindo as duas pausas. Durante as pausas e no final do teste, a criança recebia um adesivo com motivos infantis. A idéia inicial é que ela os colasse em um cartão que era entregue antes do início do teste. Entretanto, a maioria das crianças preferiu guardar os adesivos.

¹⁸ Uma próxima versão do teste poderá considerar o uso do *touch screen* para facilitar a tarefa.

O teste foi aplicado em uma sala da própria escola, silenciosa, e com fone acoplado a um micro-computador, na intensidade de maior conforto para cada sujeito.

2.2 Sujeitos sem queixas auditivas – resultados

Os resultados encontrados para o grupo de crianças sem queixas auditivas foram analisados com relação ao desempenho de cada sujeito e com relação à frequência do acerto dos itens do teste e dos arquivos de áudio.

Na análise estatística, o teste de Cochran demonstrou que não houve diferenças significativas nos acertos nos 10 arquivos de áudio correspondentes às dez repetições em cada item, sendo cada um apresentado duas vezes. Isso foi analisado para de verificar se havia, entre eles, algum que pudesse interferir no resultado e que, portanto, deveria ser substituído. Para o item *pata* [Cochran's $Q_{(19)}=10,502$; $p=0,939$]; para o item *bata* [Cochran's $Q_{(19)}=25,409$; $p=0,146$]; para *tata* [Cochran's $Q_{(19)}=24,417$, $p=0,181$]; para *data* [Cochran's $Q_{(19)}=13,671$, $p=0,803$]; para *cata* [Cochran's $Q_{(19)}=28,533$, $p=0,074$] e para *gata* [Cochran's $Q_{(19)}=20,754$, $p=0,351$].

Com relação ao desempenho dos sujeitos, verificamos que o teste não revelou dificuldades para esse grupo de crianças. Considerando todos os itens do teste, a maioria das crianças com sem queixas auditivas, (70%) obteve porcentagem de acerto superior a 98 %, o que corresponde a dois itens errados, e 100% das crianças avaliadas obtiveram porcentagem de acerto igual ou superior a 95%, o que correspondeu a 6 erros em 120. O número de sujeitos por porcentagem de acerto está apresentado no gráfico abaixo (figura 14).

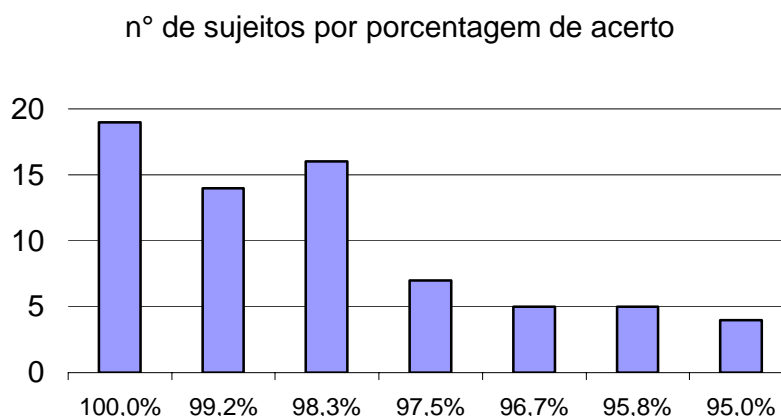


Figura 14 - Gráfico representativo da distribuição dos sujeitos sem queixas auditivas por porcentagem de acerto (n = 70)

As médias de acerto por item estão apresentadas na tabela 1. Convém lembrar que cada item foi apresentado 20 vezes. Pode-se notar que para todos os itens a média de acerto foi alta, embora exista uma diferença entre os itens: o item com maior média de acerto foi *pata*, seguido de *gata*, *data*, *tata*, *cata* e *bata*. O teste de Friedman aponta que existem diferenças significativas entre os acertos dos itens ($\chi^2_{(5)}=19,261$, $p=0.002$).

Tabela 1 - DISTRIBUIÇÃO DO NÚMERO DE ACERTOS POR ITEM, CONSIDERANDO O TOTAL DE 70 SUJEITOS SEM QUEIXAS AUDITIVAS

	Mínimo	Máximo	Média	Desvio Padrão
PATA	18	20	19,84	0,44
BATA	17	20	19,51	0,65
TATA	17	20	19,65	0,66
DATA	17	20	19,71	0,57
CATA	17	20	19,59	0,79
GATA	18	20	19,75	0,55

Os itens com maior tendência de erro foram as palavras relacionadas a verbos de ação, (*cata* e *bata*), e a abreviação de tataravó, (*tata*). As duas primeiras são mais difíceis de serem representadas por meio de figuras e a última

é menos familiar. A comparação das médias de acertos (*one way ANOVA*) demonstrou que não houve diferença entre o desempenho em cada item considerando o fator sexo. Para *pata*: [$F_{(1,68)}=1,371$, $p=0,286$]; *bata*: [$F_{(1,68)}=0,33$, $p=0,856$]; *tata*: [$F_{(1,68)}=1,142$, $p=0,289$]; *data*: [$F_{(1,68)}=0,232$, $p=0,632$]; *cata*: [$F_{(1,68)}=1,080$, $p=0,302$]; *gata*: [$F_{(1,68)}=0,053$, $p=0,819$]. Entretanto, considerando o fator faixa etária, encontramos diferença significativa no que se refere ao item *bata* [$F_{(2,67)}=4,430$, $p=0,016$]. O teste *a posteriori* de Scheffe indicou diferença ($p=0,029$) no desempenho das crianças na faixa etária de 6 a 6 anos e 11 meses em relação aos outros dois grupos (de 7 a 7 anos e 11 meses e de 8 anos a 8 anos e 5 meses). Neste grupo, ($n=20$) a média de acerto para o item *bata* foi de 19,15. Para os outros itens, não foram encontradas diferenças significativas relacionadas à faixa etária. A partir destes resultados, que sugerem que as crianças mais novas apresentaram maior dificuldade em identificar o item *bata* convém que o teste, da forma em que se encontra, seja utilizado com crianças com idade a partir de 7 anos. Considerando tal restrição, é possível afirmar que o teste pode ser utilizado para investigar a discriminação das plosivas do PB em sujeitos com deficiência auditiva, com idade superior a 7 anos.

Para a faixa etária estudada no grupo sem queixas auditivas, em início de alfabetização, o apoio da escrita não foi relevante, ou seja, as crianças com audição normal não se apoiaram na leitura das palavras relativas a cada item. Para a aplicação em sujeitos com deficiência auditiva, usamos como critério selecionar sujeitos que já dominassem a leitura, de forma que o apoio da escrita seja utilizado.

2.3 Sujeitos com deficiência auditiva – metodologia

Participaram do estudo 13 sujeitos com deficiência auditiva neurosensorial congênita ou adquirida antes dos dois anos de idade, de vários graus de perda auditiva e configuração audiométrica plana ou descendente. A faixa etária variou de 8 anos e 2 meses a 19 anos e 2 meses. Todos os sujeitos estão ou estiveram inseridos em programas de atendimento fonoaudiológico que incluem em seus

objetivos o uso funcional da audição residual e, portanto, fazem uso regular de aparelhos de amplificação sonora individual. Todos freqüentam escola e todos puderam reconhecer os itens do teste também pela escrita. O quadro 1 apresenta a caracterização dos sujeitos e os limiares para tom puro da melhor orelha bem como a média para as freqüências de 0,5; 1 e 2 kHz (PTA). Os limiares audiométricos estão representados no gráfico apresentado na figura 15, agrupados por grau de perda auditiva – moderadamente severa, severa e profunda (Silman, Silverman, 1997). É importante ressaltar que o sujeito número nove é o mesmo sujeito com deficiência auditiva descrito no capítulo 3, que trata da metodologia.

QUADRO 1 - CARACTERIZAÇÃO DOS SUJEITOS COM DEFICIÊNCIA DE AUDIÇÃO: SEXO, IDADE, LIMIARES AUDIOMÉTRICOS POR VIA AÉREA DA MELHOR ORELHA EM DB NA, NAS FREQUÊNCIAS DE 0.25 A 8 kHz E MÉDIA DAS FREQUÊNCIAS DE 0.5; 1 E 2 kHz – PTA, EM dB NA

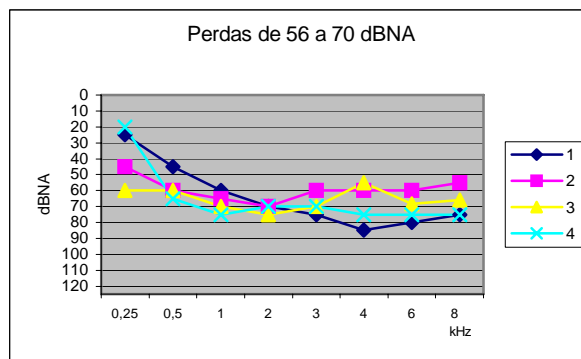
sujeito	sexo	idade	0,25	0,5	1k	2k	3k	4k	6k	8k	PTA
1	F	9a9m	25	45	60	70	75	85	80	75	58
2	M	8a2m	45	60	65	70	60	60	60	55	65
3	F	14a9m	60	60	70	75	70	55	68	66	66
4	F	14a	20	65	75	70	70	75	75	75	70
5	M	9a6m	30	55	90	85	75	75	110	100↓	77
6	F	18a	65	80	75	80	70	80	110	90	78
7	M	13a2m	65	70	85	85	75	75	80	95	80
8	M	15a	85	85	85	80	70	70	80	95	83
9	F	16a	90	90	100	80	75	80	120↓	100↓	90
10	M	19a2m	70	80	100	105	115	120↓	120↓	100↓	95
11	M	14a9m	80	85	90	110	115	120	120↓	100↓	95
12	M	18a1m	70	90	100	105	120	120↓	120↓	100↓	98
13	F	16a1m	90	90	95	115	110	115	120↓	100↓	100

Legenda:

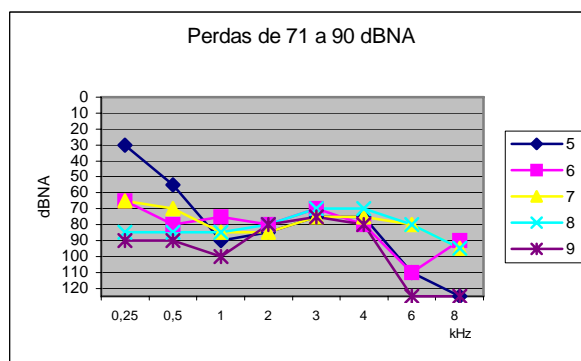
A seta (↓) ao lado do valor indica que não houve resposta para a intensidade referida naquela freqüência.

PTA é a abreviação de *pure tone average*, ou seja, média dos limiares tonais nas freqüências de 0,25, 0,5 e 1 kHz que são consideradas as mais importantes para a percepção auditiva de fala.

A. Perdas moderadamente severas – de 56 a 70 dBNA



B. Perdas severas – 71 a 90 dBNA



C. Perdas profundas – > 90 dBNA

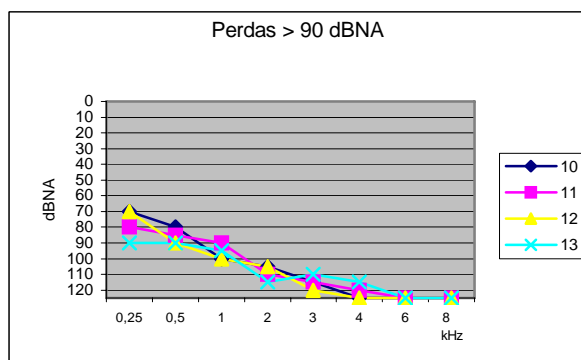


Figura 15 - Gráfico representativo dos limiares tonais dos sujeitos com deficiência auditiva agrupados por grau de perda auditiva: (A) perdas moderadamente severas; (B) perdas severas; (C) perdas profundas

O teste foi aplicado em cabina acústica, com o computador acoplado ao audiômetro, com o monitor e o *mouse* dentro da cabina. Após a calibração, o nível de apresentação selecionado foi o de maior conforto para cada sujeito e variou de 65 a 75dB NPS, em campo livre. Todos os sujeitos estavam usando seus

aparelhos de amplificação sonora individual da forma habitual, que foram previamente conferidos pela examinadora para verificação das baterias e qualidade do som.

A escrita, para esse grupo, foi usada como apoio no módulo de apresentação dos itens, para garantir a compreensão dos vocábulos, antes do início do teste. Como no grupo de crianças com audição normal, o teste só foi iniciado depois de garantida a correta identificação dos itens do teste, com apoio da leitura oro-facial e da escrita. O teste também só foi iniciado após bom desempenho no módulo de treinamento. Não foram encontradas dificuldades na aplicação do teste nesse grupo, provavelmente em virtude da faixa etária ser mais elevada e também pelo fato de estes sujeitos estarem mais habituados a situações de avaliação semelhantes. Mesmo assim, foram tomados cuidados em relação à interrupção do teste a cada 40 apresentações, para evitar o cansaço.

2.4 Sujeitos com deficiência auditiva – resultados

Os resultados encontrados para o grupo de sujeitos com deficiência auditiva foram tratados de forma diferente em relação aos revelados pelas crianças sem queixas auditivas. Nesse caso, não se trata mais de avaliar o procedimento, mas sim de analisar os acertos e erros como indícios da habilidade de perceber os contrastes de vozeamento e ponto de articulação das plosivas do PB.

A porcentagem de acerto foi calculada para cada sujeito, para o contraste de ponto de articulação nas três oposições apresentadas no teste (40 apresentações para cada oposição):

1. condição A: bilabial x alveolar - (BxA)
2. condição B: bilabial x velar - (BxV)
3. condição C: alveolar x velar- (AxV)

A porcentagem de acerto do contraste de voz foi calculada para cada sujeito, considerando todas as apresentações do teste (120 no total). Considerou-se acerto o contraste de voz ter sido corretamente identificado, independentemente do acerto

quanto ao ponto de articulação.

Como nesse teste é obrigatória a escolha de um entre quatro itens (dois de mesmo ponto de articulação, nas ocorrências surda e sonora), existe 50% de chance de respostas corretas serem dadas ao acaso. Para eliminar esse efeito, as porcentagens de acerto foram calculadas utilizando-se uma fórmula matemática sugerida por Boothroyd (1985) que elimina do resultado o acerto casual. Dessa forma, os resultados foram obtidos com a seguinte fórmula:

$$\text{Porcentagem de acerto} = \frac{\text{total de acertos} - \text{chance de acerto}}{\text{total de itens} - \text{chance de acerto}}$$

Os resultados estão dispostos na tabela 2, seguindo a ordem crescente das médias dos limiares das frequências de 0,5; 1 e 2 kHz (PTA) dos sujeitos, que estão identificados com números de 1 a 13.

Tabela 2 - DISTRIBUIÇÃO DOS SUJEITOS COM DEFICIÊNCIA AUDITIVA SEGUNDO PORCENTAGEM DE ACERTO NA DISCRIMINAÇÃO DO PONTO DE ARTICULAÇÃO, POR OPOSIÇÕES APRESENTADAS, PONTO (TOTAL) E VOZEAMENTO

sujeitos	Grau de perda	BxA	BxV	AxV	Ponto	Voz
1	MS	65	95	95	85	92
2	MS	80	100	100	93	100
3	MS	50	65	35	50	61
4	MS	100	85	80	88	96
5	S	50	50	10	36	75
6	S	80	100	100	93	96
7	S	100	100	85	95	100
8	S	85	95	95	92	80
9	S	35	95	60	63	0
10	P	35	50	55	53	0
11	P	40	80	20	46	45
12	P	15	40	20	25	48
13	P	0	25	0	0	0

Legenda:

BxA - contraste bilabial x alveolar

BxV - contraste bilabial x velar

AxV - contraste alveolar x velar

Ponto - contraste de ponto de articulação (total)

Voz - contraste de vozeamento

Grau de perda: MS - moderadamente severa, S – severa, P - profunda

No gráfico apresentado na figura 16, é possível visualizar a distribuição dos sujeitos por porcentagem de acerto na discriminação do contraste de voz e de ponto de articulação, considerando o resultado global. Nota-se que ocorre, de forma geral, uma diminuição na porcentagem de acerto na discriminação dos dois contrastes conforme aumenta o grau de perda auditiva. Entretanto, isso não é verdadeiro para todos os casos, uma vez que observamos também perdas menores e piores resultados, como ocorreu com os sujeitos números 3 e 5. Em contrapartida, os casos 8 e 6, ainda que apresentem grau de perda auditiva mais acentuado, revelaram uma porcentagem alta de acertos.

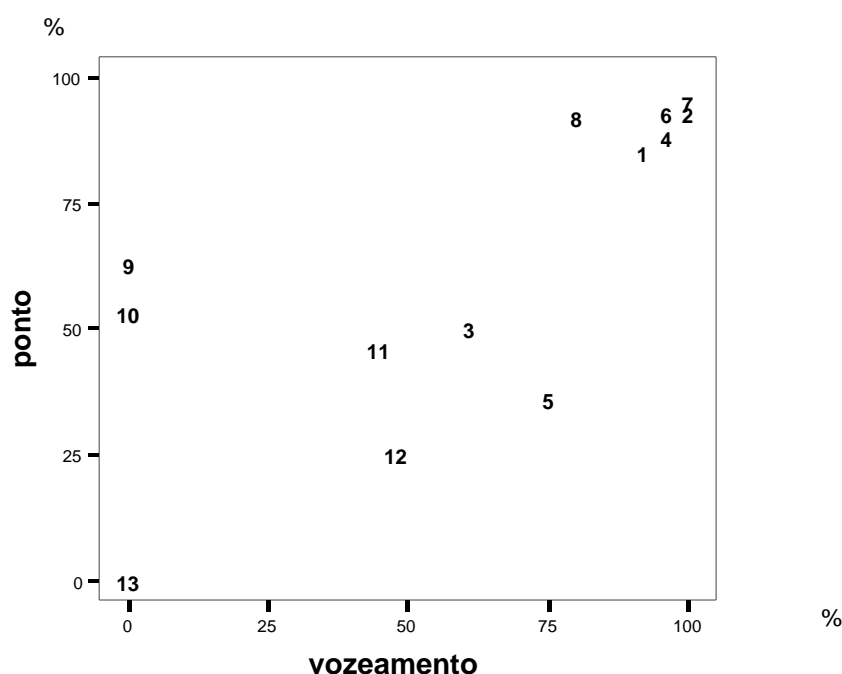


Figura 16 - Gráfico representativo da distribuição dos sujeitos por porcentagem de acerto na discriminação dos contrastes de ponto de articulação e voz

Para oito dos sujeitos estudados, a porcentagem de acerto na discriminação do contraste de voz foi mais alta que a porcentagem de acerto na discriminação do contraste de ponto de articulação (sujeitos 1, 2, 3, 4, 5, 6, 7, e 12). Um dos sujeitos apresentou apenas 1% de diferença (sujeito 11), três apresentaram maior porcentagem de acerto para o contraste de ponto de articulação (sujeitos 8, 9 e 10)

e um não identificou nenhum dos contrastes acima do nível da chance (sujeito 13). Esse último também é o que apresenta o maior grau de perda auditiva. No gráfico apresentado na figura 17, estão representadas as médias de acerto para o contraste de voz e ponto de articulação. O teste de Wilcoxon demonstrou que não ocorreram diferenças significativas entre os acertos na discriminação do contraste de vozeamento e ponto de articulação, considerando os 13 sujeitos estudados ($Z = -0,549$, $p = 0,583$).

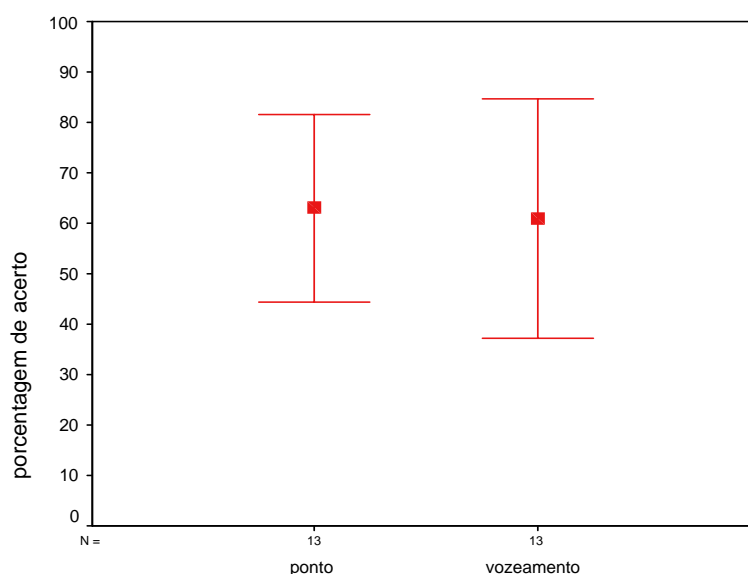


Figura 17 - Gráfico representativo das médias das porcentagens de acerto na discriminação do ponto de articulação e vozeamento pelo grupo de sujeitos com deficiência auditiva

Os resultados da discriminação do ponto de articulação por oposição estão representados no gráfico apresentado na figura 18. Podemos observar que, de forma geral, as porcentagens mais altas de discriminação do ponto de articulação ocorreram na oposição bilabial x velar, seguida da alveolar x velar e da bilabial x alveolar, que foi a oposição que apresentou porcentagens mais baixas de discriminação. No entanto, também nesse caso pode ser verificada a variabilidade entre os sujeitos, embora se observe que quanto maior é o grau de perda menor é a porcentagem de acerto.

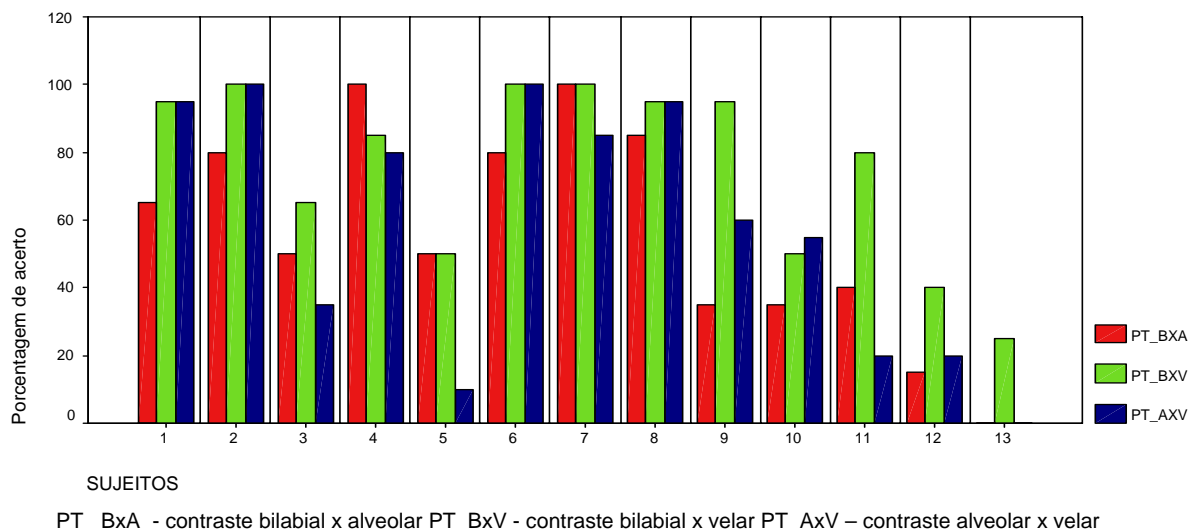
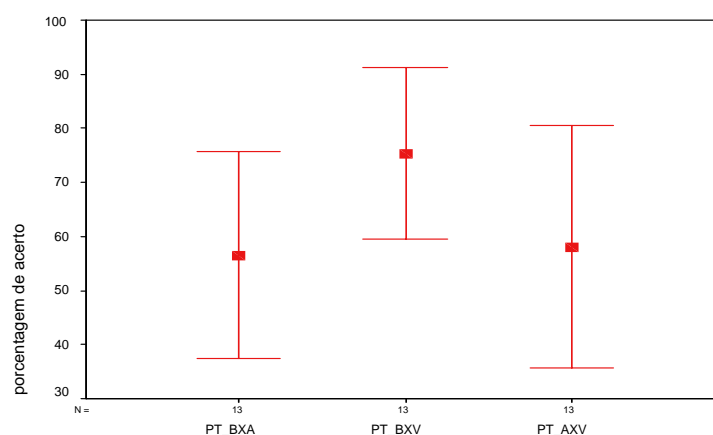


Figura 18 - Gráfico representativo da distribuição das porcentagens de acerto na identificação do contraste de ponto de articulação para cada sujeito, para cada contraste apresentado

No gráfico mostrado pela figura 19, estão representadas as médias de acerto na discriminação do contraste de ponto de articulação, por oposição apresentada. Fica evidente o maior percentual de acerto na discriminação do contraste de ponto de articulação quando a oposição é a bilabial x velar. O teste de Friedman apontou para uma diferença significativa entre as porcentagens de acerto para os três pontos de articulação ($\chi^2_{(2)}=8,933$, $p=0,011$).



PT_ BxA - contraste bilabial x alveolar PT_BxV - contraste bilabial x velar PT_AxV – contraste alveolar x velar

Figura 19 - Gráfico representativo da média das porcentagens de acerto na discriminação do ponto de articulação para cada oposição apresentada

Notamos uma tendência de menor porcentagem de acerto na discriminação dos contrastes de ponto de articulação e voz com o aumento do grau da perda auditiva. Entretanto, a variação das respostas dos sujeitos, mesmo entre aqueles com o mesmo grau de perda auditiva, é muito grande, como indica o desvio padrão e o coeficiente de variação na tabela 3, a seguir:

Tabela 3 - MÉDIAS DAS PORCENTAGENS DE ACERTO PARA CADA OPOSIÇÃO SEGUNDO OS GRUPOS DE SUJEITOS POR GRAU DA PERDA AUDITIVA

Grau de perda		BxA	BxV	AxV	Ponto	Voz
MS	Média (d.p.)	74 (21)	86 (15)	78 (30)	79 (17)	87 (18)
	N	4	4	4	4	4
	cf.var.	0,28	0,17	0,38	0,21	0,21
S	Média (d.p.)	70 (27)	88 (21)	70 (37)	76 (26)	70 (41)
	N	5	5	5	5	5
	cf.var.	0,39	0,23	0,53	0,34	0,58
P	Média (d.p.)	23 (18)	49 (23)	24 (23)	31 (24)	23 (27)
	N	4	4	4	4	4
	cf.var.	0,80	0,47	0,96	0,77	1,17
Total	Média (d.p.)	57 (32)	75 (26)	58 (37)	63 (31)	61 (39)
	N	13	13	13	13	13
	cf.var.	0,56	0,35	0,64	0,49	0,55

Legenda:

BxA - contraste bilabial x alveolar

BxV - contraste bilabial x velar

AxV - contraste alveolar x velar

Ponto - contraste de ponto de articulação (total)

Voz - contraste de vozeamento

Grau de perda: MS - moderadamente severa; S - severa, P - profunda

N - número de sujeitos por grupo

d.p. - desvio padrão

cf.var. - coeficiente de variação

Esses resultados estão de acordo com a literatura que indica que as habilidades de discriminação de sons de fala não podem ser previstas apenas pelo grau da perda auditiva (Sammeth et al, 1996; Turner, Brus, 2001), embora, de forma geral, possamos dizer que as dificuldades em perceber contrastes de sons da fala se intensifiquem com o aumento da perda auditiva (Boothroyd, 1982; Longone, Borges, 1997; Delgado, Bevilacqua, 1999).

A comparação entre as médias de acerto entre os grupos de sujeitos por grau de perda auditiva (*one way Anova*) mostrou que, para o contraste de ponto de

articulação, ocorreram diferenças significativas entre os grupos [$F_{(2,10)}=5,365$, $p=0,26$]. O teste *a posteriori* de Scheffe indicou que ocorreram diferenças significativas entre o grupo com perda moderadamente severa e o grupo com perda profunda ($p=0,048$). Para contraste de voz, observamos resultado semelhante: a comparação de médias demonstrou diferença significativa entre os grupos [$F_{(2,10)}=4,569$, $p=0,039$]. O teste *a posteriori* de Scheffe apontou para uma diferença significativa entre o grupo com perda profunda e o grupo com perda moderadamente severa ($p=0,047$).

Na figura 20, estão apresentados os gráficos que ilustram a distribuição das porcentagens de acerto segundo o grau de perda auditiva por contraste de ponto de articulação (A) e para o contraste de voz (B).

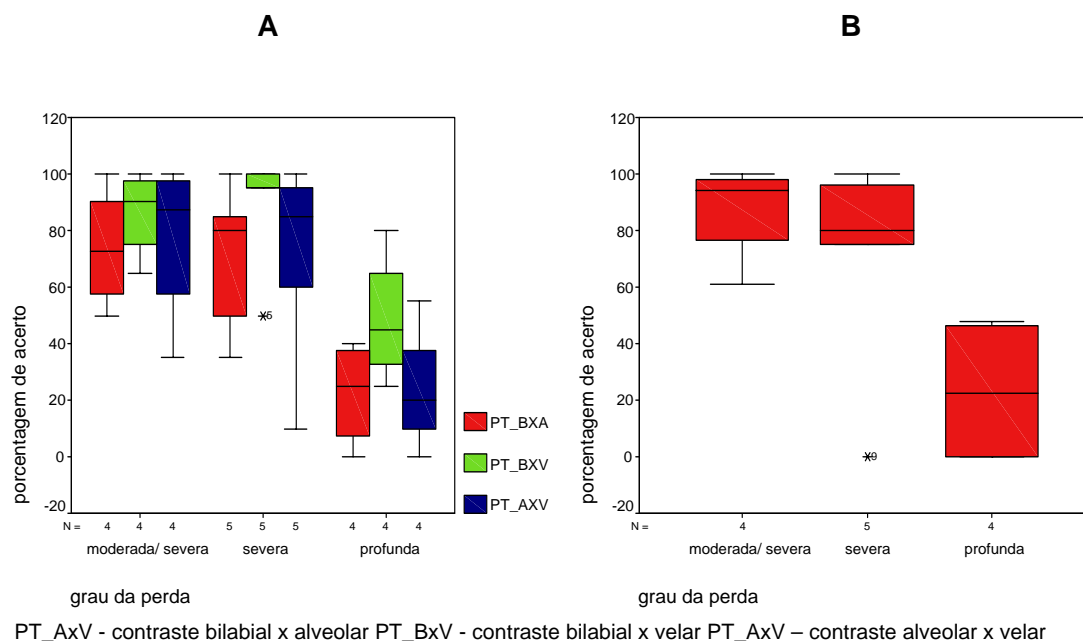


Figura 20 - Gráficos representativos das médias das porcentagens de acerto para os diferentes graus de perda auditiva por contraste de ponto de articulação nas três oposições apresentadas (A) e por contraste de voz (B)

Considerando o contraste de ponto de articulação por oposição apresentada, a análise estatística mostrou que, para as oposições bilabial x alveolar [$F_{(2,10)}=6,413$, $p=0,016$] e bilabial x velar [$F_{(2,10)}=4,926$, $p=0,032$], ocorreram diferenças significativas entre os grupos. Já para o contraste alveolar x velar, não foram observadas diferenças significativas entre os grupos de sujeitos com diferentes

graus de perda auditiva [$F_{(2,10)}=3,595$, $p=0,067$]. Para a oposição bilabial x alveolar, o teste *a posteriori* de Scheffe revelou que não houve diferenças significativas entre os grupos com perdas moderadamente severas e severas, mas sim entre estes e o grupo com perdas profundas ($p=0,031$ e $p=0,035$ respectivamente). Para a oposição bilabial x velar, ocorreram diferenças significativas entre o grupo com perda severa e o grupo com perda profunda ($p=0,050$).

Os dados aqui apresentados evidenciam a diversidade existente entre os sujeitos com deficiência auditiva no que se refere à habilidade de identificar os contrastes das plosiva em posição inicial em sílaba tônica, inclusive entre sujeitos com mesmo grau de perda auditiva.

Verificamos uma correlação significativa (negativa) entre o grau de perda e a porcentagem de acertos na discriminação dos contrastes de voz e ponto, que diminuiu à medida que o grau da perda auditiva era maior (Correlação de *Pearson*, para o contraste de voz: $r=-664$; $p=0,013$; e para contraste de ponto de articulação: $r= -364$ $p=0,020$). Para o contraste de ponto de articulação a correlação mais forte foi para a oposição bilabial x alveolar ($r= -662$; $p=0,007$), seguida da oposição alveolar x velar ($r= -590$; $p=0,017$) e da bilabial x velar ($r= -583$; $p=0,018$)

Tais resultados todavia, devem ser considerados com cautela, uma vez que se referem a um número reduzido de sujeitos. A diversidade nos resultados entre os sujeitos apontam para a importância de se estudar cada caso, na busca de compreender as habilidades de percepção de sons de fala de cada sujeito com deficiência auditiva.

Cabe observar que os dados referentes ao sujeito de número nove (que é o mesmo sujeito selecionado para as tarefas de produção de fala deste trabalho e está caracterizado no capítulo 3 - Métodos) serão reapresentados no capítulo 4 (Resultados), e considerados como parte dos resultados das tarefas de percepção das plosivas.

3 MÉTODOS

Apresentaremos, a seguir, a descrição da elaboração e gravação do *corpus*, a caracterização dos sujeitos envolvidos nesta pesquisa, os procedimentos para realização das tarefas de identificação e discriminação auditiva das plosivas e os procedimentos adotados para realização das medidas acústicas.

3.1 Elaboração do *corpus*

Para os propósitos desta pesquisa, foi elaborado um *corpus* de seis palavras do PB, dissílabas, paroxítonas, que correspondem ao padrão acentual mais freqüente no PB (Albano, 1995), seguindo o padrão CVCV, em que a consoante inicial corresponde a uma das seis plosivas do PB, seguidas da terminação [____], formando pares mínimos. Como já explicitamos, as palavras *pata*, *tata*, *cata*, *bata*, *data* e *gata* foram inseridas na frase-veículo *Diga _____ baixinho*.

Essas palavras que compuseram o *corpus* foram selecionadas com o uso dos seguintes critérios: (1) deveriam ser passíveis de representação pictórica, de forma a possibilitar a elaboração de um instrumento de avaliação da discriminação das plosivas do PB, no qual a tarefa requerida seria que o sujeito apontasse a figura correspondente à palavra escutada; (2) deveriam ser familiares a crianças e sujeitos com deficiência de audição. No entanto, criar uma lista de pares mínimos envolvendo as seis plosivas do PB que atendesse a tais critérios não foi tarefa fácil, sendo a seleção aqui apresentada a que nos pareceu melhor.

O *corpus*, produzido por dois sujeitos caracterizados a seguir, foi analisado acusticamente e também utilizado para a realização das diferentes tarefas de percepção auditiva propostas neste estudo, incluindo a elaboração do instrumento para a avaliação da discriminação das plosivas do PB descrito no capítulo 2.

3.1.1 Procedimento de Gravação do *corpus*

Os dados foram gravados no estúdio de gravações da Faculdade de Comunicação e Filosofia da PUC/SP, em gravador digital (DAT), TCD-D8 *Sony*, com microfone *AudioTechnica* – ATM 25, com impedância de 600 Ω , posicionado em cima da mesa à frente do falante, num suporte de forma a ficar a 10 cm de sua boca.

Foram gravadas 10 repetições de cada frase-veículo contendo cada uma das seis palavras dissílabas paroxítonas que compuseram o *corpus*. Assim sendo, cada sujeito gravou 60 repetições.

Os dois sujeitos leram as frases impressas em fonte Arial, tamanho 48, em cartões apresentados um a um, em ordem aleatória, através do vidro da cabina acústica.

Os dados foram digitalizados no *CSL-4300B* da *Kay Elemetrics*, numa taxa de amostragem de 22 kHz.

3.1.2 Sujeitos

Para a gravação do *corpus*, foram selecionadas duas pessoas do sexo feminino, uma com audição normal e outra com deficiência auditiva.

O sujeito com audição normal (AN) é uma mulher de 35 anos de idade (na época da gravação), nascida e residente na cidade de São Paulo, que não apresenta alterações de audição, voz e articulatórias.

O sujeito com deficiência auditiva (DA) é uma jovem de 16 anos (na época da gravação), nascida e residente na cidade de São Paulo, que apresenta deficiência auditiva neurossensorial, bilateral, de grau severo à direita e profundo à esquerda, decorrente de meningite aos 18 meses de idade. Aos 24 meses, iniciou atendimento fonoaudiológico e uso de amplificação, sem interrupção até a data da gravação. Sempre esteve envolvida no serviço de atendimento fonoaudiológico

que privilegia o uso da audição residual e da língua oral. Na época da gravação, freqüentava a 7ª série do ensino fundamental. Os dados referentes à avaliação audiológica são apresentados na sequência. Os limiares tonais, com e sem amplificação, estão no quadro 2. Cabe lembrar que esse é o sujeito número 9 mencionado no capítulo 3, que descreve o procedimento de discriminação das plosivas.

QUADRO 2 - LIMIARES TONAIS (dBNA) NA ORELHA DIREITA (OD) E ORELHA ESQUERDA (OE), MEDIDOS COM FONES E EM CAMPO LIVRE, COM AMPLIFICAÇÃO

	250	500	1000	2000	3000	4000	6000	8000	Hz
OD	90	90	100	80	75	80	120↓	100↓	dB NA
OE	90	110	120	120↓	120↓	120↓	120↓	100↓	dB NA
AMP		45	30	30	30	40			dB NA

Legenda :

OD – orelha direita

OE – orelha esquerda

AMP – em campo livre com uso de aparelho de amplificação sonora individual

A seta (↓) ao lado do valor indica que não houve resposta para a intensidade referida naquela frequência.

Os resultados da imitanciometria revelaram curva timpanométrica tipo A e ausência de reflexo acústico estapediano contralateral, bilateralmente. Na logaudiometria, foram obtidos: Limiar de Detecção de Fala em 70 dBNA à direita e 90 dBNA à esquerda e somente para a orelha direita, Limiar de Reconhecimento de Fala na intensidade de 80 dBNA e Índice de Reconhecimento de Fala de 60% para a lista de palavras dissílabas e 40% para monossílabas.

3.2 Tarefas de percepção

Foram realizadas quatro tarefas de percepção auditiva: a identificação das produções do sujeito com deficiência auditiva por sujeitos com audição normal, a identificação das produções do sujeito com audição normal por sujeitos com

audição normal, a discriminação/identificação¹⁹ das produções do sujeito com audição normal pelo sujeito com deficiência auditiva e a discriminação/identificação das produções do sujeito com deficiência auditiva por ele mesmo.

3.2.1 Sujeitos com audição normal e as produções do falante com deficiência auditiva

Os sujeitos que participaram da tarefa de identificação das produções do falante com deficiência auditiva foram 60 estudantes de graduação (do primeiro ano do curso de Fonoaudiologia) da PUC/SP, que ainda não estavam familiarizados com a fala de deficientes auditivos.

Das dez produzidas pelo sujeito deficiente auditivo foi escolhida a quinta, porque essa seria a menos afetada por fatores como cansaço — que poderiam acontecer nas listas finais, ou hiperarticulação, como poderia acontecer nas primeiras produções.

As seis sentenças foram apresentadas aleatoriamente para os sessenta sujeitos numa sala silenciosa, que usavam fones de ouvido acoplados ao computador. A intensidade da apresentação foi ajustada no nível de maior conforto para cada sujeito. Previamente à realização da tarefa, cada sujeito escutou três outras palavras com características semelhantes às que compõem o *corpus*, inseridas na mesma frase-veículo, produzidas nas mesmas condições pelo sujeito deficiente auditivo, para que pudesse familiarizar-se com a tarefa que iria realizar. Os sujeitos foram orientados a escrever em folha apropriada o que tinham acabado de escutar, mesmo que não soasse como uma palavra conhecida.

¹⁹ A tarefa de percepção auditiva realizada por meio do procedimento descrito no capítulo 2, tem uma característica diversa daquela realizada pelos sujeitos com audição normal, identificando as produções das plosivas na fala do sujeito com deficiência auditiva e com audição normal. A primeira é uma múltipla escolha de um entre quatro itens e entre dois contrastes (vozeamento e ponto de articulação). A segunda implica a identificação de sons, sem nenhuma restrição de contexto, a não ser as próprias restrições da língua. Dessa forma, embora as duas tarefas envolvam habilidades de discriminação e identificação de sons, neste trabalho vamos usar o termo *tarefa de identificação* para as tarefas 1 e 2 *tarefa de discriminação* para as tarefas 3 e 4, descritas adiante.

3.2.2 Sujeitos com audição normal e as produções do falante com audição normal

As produções do sujeito com audição normal foram apresentadas para outros 60 sujeitos estudantes da Faculdade de Fonoaudiologia da PUC/SP sem alterações auditivas.

O mesmo procedimento descrito no item 3.2.1 (Sujeitos com audição normal identificando as produções do sujeito com deficiência auditiva) foi realizado com as produções do falante com audição normal, que foram apresentadas a 60 sujeitos, dispensando as três apresentações prévias, uma vez que, nesse caso, não havia necessidade de habituação à produção da fala do sujeito com audição normal.

3.2.3 Sujeito com deficiência auditiva e as produções do sujeito com audição normal

Para realização da tarefa de discriminação auditiva das plosivas produzidas pelo sujeito com audição normal, foi utilizado o procedimento desenvolvido paralelamente a este trabalho, cujo detalhamento encontra-se no capítulo 2, e que é designado como *procedimento de discriminação de contrastes de vozeamento e ponto de articulação das plosivas do PB*²⁰. Trata-se de um teste de múltipla escolha entre quatro itens, no qual o sujeito deve indicar com o *mouse* na tela de computador a figura correspondente ao item escutado. Os itens apresentados auditivamente correspondem às 10 repetições gravadas pelo sujeito com audição normal²¹.

3.2.4 Sujeito com deficiência auditiva e suas próprias produções

Para realização da tarefa de discriminação de suas próprias produções por parte do sujeito com deficiência auditiva, foi feita uma adaptação no procedimento

²⁰ Cabe-nos lembrar, mais uma vez, que o sujeito com deficiência auditiva deste estudo e o sujeito de número 9 do grupo dos que possuem deficiência auditiva, que foi apresentado no capítulo 2, são a mesma pessoa.

²¹ Tal teste foi concebido para ser utilizado nesta pesquisa, mas também na clínica fonoaudiológica, como procedimento de avaliação da percepção auditiva para sons de fala. Desse modo, outros pares mínimos serão incluídos, a saber: mata, nata, lata, rata e chata.

de discriminação das plosivas do PB: substituímos os arquivos de áudio que continham as produções do sujeito com audição normal pelos arquivos que continham as produções do sujeito com deficiência auditiva.

3.3 Análise Acústica

A análise acústica do *corpus* gravado e digitalizado foi realizada por meio do *software Multispeech* da *Kay Elemetrics*, modelo 3700, versão 2.3, no Laboratório Integrado de Análise Acústica e Cognição (LIAAC) da Pontifícia Universidade Católica de São Paulo. As seguintes medidas foram realizadas: (1) tempo de ataque do vozeamento ou *voice onset time (VOT)*; (2) frequência de F1 no início da vogal subsequente (*onset F1*); (3) frequência fundamental no início da vogal subsequente (*f0 no onset da vogal*); (4) duração das consoantes (absoluta e relativa), duração da frase e dos segmentos adjacentes; (5) energia da consoante em relação à vogal subsequente; (6) frequência do primeiro, segundo e terceiro formantes das vogais adjacentes (F1, F2 e F3 das vogais); (7) transição do primeiro, segundo e terceiro formantes em direção à vogal subsequente (transição de F1, F2 e F3), medida a cada pulso do primeiro terço da vogal subsequente.

Segue a descrição dos procedimentos utilizados na realização das medidas. Estão detalhados os parâmetros utilizados para gerar os espectrogramas de banda larga, os espectros - FFT²², LPC²³, o Histórico do Formante (*Formant History*), os contornos de *pitch* e de energia e a marcação dos períodos de vozeamento (*voice period marks*).

3.3.1 Parametrização usada na realização das medidas

²² Abreviatura para *Fast Fourier Transform*, algoritmo que realiza a Transformada de Fourier em microcomputadores, com muita rapidez.

²³ Abreviatura para *Linear Predictive Coding* – codificação por predição linear, usado basicamente para medir os formantes.

Para a geração dos gráficos e realização das medidas foram utilizados os seguintes parâmetros:

- Espectrograma de banda larga: *analysis size* - 75 pontos; *pre-emphasis* - 0,98; *window weighting* - *Hanning*.
- LPC: para realização da medida da frequência dos formantes das vogais: *frame length* - 10 ms para medida dos formantes na porção estacionária da vogal e 20 ms para a medida da transição dos formantes (/Fn/)t); *filter order* - 24; *pre-emphasis* - 0,98; *window weighting* - *Hanning*.
- FFT: para sobreposição do LPC, para determinação da frequência dos formantes: *analysis size* - 1024 pontos; *pre-emphasis* - 0,98; *window weighting* - *Hanning*.
- Marcador de períodos (*Voice Period Marks*): análise invertida, *analysis range* - (70 - 350 Hz).
- Histórico do Formante (*Formant History*): *frame length* - 10 ms; *frame advance* - 10 ms; *pre-emphasis* - 0,98; *filter order* - 24; *window weighting* - *Hanning*.
- Contorno de *pitch* (*Pitch Countour*): *analysis range* - 70 a 350 Hz; *Linear display range* - (0 - 350 Hz); *Log display range* - (50 - 350 Hz).
- Contorno de energia (*Energy contour*): *frame length* - 10 ms; *display range* - (0 - 80 dB).

3.3.2 Procedimento para a realização das medidas acústicas

Os procedimentos adotados foram os seguintes:

VOT – Essas medidas foram realizadas a partir da forma da onda, com a observação simultânea do espectrograma de banda larga. O VOT foi medido considerando como ponto inicial da soltura articulatória da plosão até o ponto em que o vozeamento inicia-se (Lisker, Abramsom, 1964). O início do vozeamento, no caso das plosivas não vozeadas, VOT positivo, foi considerado no ponto em que os dois primeiros formantes da vogal subsequente eram identificados claramente. O momento correspondente na forma da onda foi marcado no ponto de cruzamento zero, imediatamente anterior à curva ascendente do primeiro ciclo regular. No caso do VOT negativo, terminamos a medida no ponto em que a vogal

anterior deixava de apresentar os dois primeiros formantes, ou um deles, sendo esse determinado também a partir da forma da onda, no ponto de cruzamento zero no último ciclo regular da vogal anterior. Devido às diferenças encontradas na produção de fala do sujeito com deficiência auditiva (ausência de pré-vozeamento nas plosivas sonoras), três medidas foram realizadas para esse sujeito: a duração da manutenção da barra de sonoridade após a vogal precedente; a duração entre a interrupção da barra de sonoridade e a plosão da consoante e o *VOT* positivo. A figura 21 ilustra a medida do *VOT* positivo na produção do sujeito com deficiência auditiva.

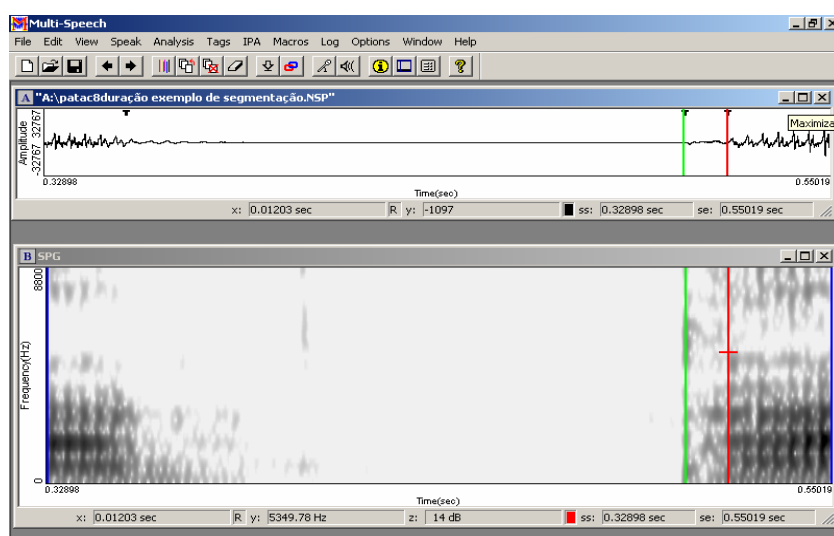


Figura 21- exemplo da medida de *VOT* da plosiva [p] - repetição 9 da frase *diga pata baixinho* produzida pelo sujeito com deficiência auditiva – DA

Frequência no onset de F1 - A seleção do ponto para essa medida no início do primeiro formante foi feita a partir da forma da onda com base na inspeção do espectrograma de banda larga. Foi marcado primeiro pico do primeiro ciclo regular da vogal subsequente, com a correspondente visualização do primeiro formante no espectrograma. Uma vez determinado o ponto, a medida da frequência foi feita a partir do LPC, com sobreposição do FFT. Um exemplo de como foi realizada a medida pode ser observado na figura 26, que também ilustra a medida da transição dos formantes.

f0 no início da vogal subsequente - Para a realização da medida da frequência fundamental no início da vogal subsequente a cada plosiva estudada selecionou-se o trecho desde antes do início da consoante até o final da vogal tônica. Foi gerado o *Voice Period Marks* – marcador de períodos e, quando necessário, a marcação dos períodos glotais foi corrigida manualmente, de forma que ela coincidissem com os picos dos pulsos. Após essa correção, foi gerado o *Pitch Countour* – Contorno de *Pitch* com os respectivos resultados numéricos. O valor correspondente ao início da vogal subsequente foi anotado. A tabela com os resultados numéricos gerados foi impressa para posterior conferência dos resultados.

Medidas de duração - Segundo Barbosa (1999), medir a duração de um segmento de fala não é uma tarefa simples. Mesmo que se defina manualmente os eventos acústicos considerados como as fronteiras do segmento a ser medido, eles nem sempre serão claramente detectáveis na forma da onda ou no espectrograma, em consequência do fenômeno da coarticulação. Alguns eventos tomados como fronteiras são o início e o fim do vozeamento da vogal e o início e o fim da fricção consonantal. Neste estudo, as medidas de duração foram feitas a partir da forma da onda, tendo-se como referência o espectrograma de banda larga.

Embora o objetivo deste estudo seja estudar a produção das plosivas, elas não ocorrem isoladas na fala e, neste *corpus*, estão em posição inicial de palavra inserida em frase-veículo. Assim, além de medir a duração da consoante, mediremos as durações dos segmentos adjacentes, de forma que a medida de duração da consoante possa ser analisada em relação aos outros segmentos do enunciado. Sabemos que no PB a duração está relacionada ao padrão de acento (Massini, 1991; Barbosa, 2000), ou seja, espera-se que a sílaba tônica seja marcada pelo aumento da duração. Ademais, como a duração dos segmentos de fala depende da taxa de elocução, é necessário que se calcule também a duração relativa dos segmentos se quisermos comparar as medidas de duração das produções dos dois sujeitos. Assim foi feita também a medida de duração da frase-veículo para que pudéssemos calcular a duração relativa de cada segmento, isto é, a que porcentagem do enunciado corresponde cada um deles (Gama-Rossi, 2000).

A delimitação dos diversos segmentos medidos foi feita a partir dos seguintes critérios:

- Duração da consoante - O início da obstrução foi considerado quando a vogal precedente acabava, ou seja, no ponto em que, no espectrograma, não se observavam mais os dois primeiros formantes ou um deles e, marcado, na forma da onda, no ponto de cruzamento zero, imediatamente anterior à curva ascendente do primeiro ciclo não regular. O final da consoante foi determinado como o ponto de início da vogal subsequente, ou seja, a partir da visualização no espectrograma dos dois primeiros formantes e, na forma da onda, no ponto de cruzamento zero, imediatamente anterior à curva ascendente do primeiro ciclo regular (Figura 22).

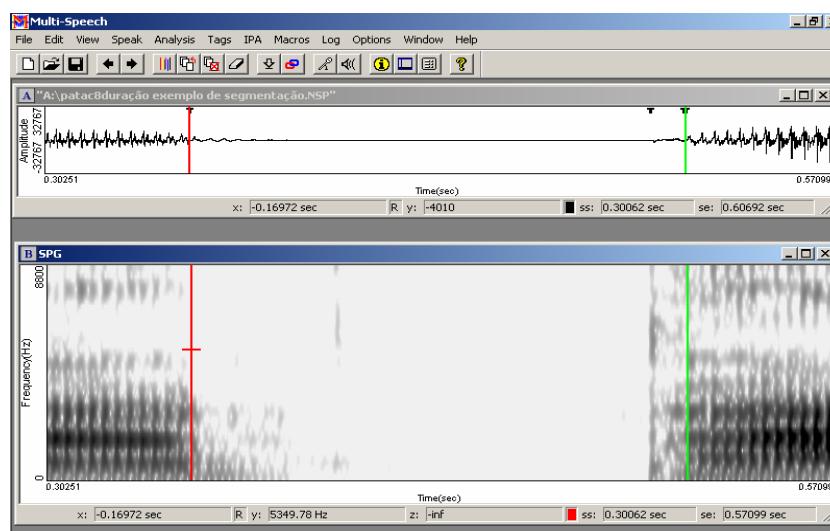


Figura 22- Exemplo da medida da duração da consoante [p] – repetição 8 da frase *diga pata baixinho* produzida pelo sujeito com deficiência auditiva – DA

- Duração da vogal adjacente - O início da vogal foi marcado na forma da onda, no ponto de cruzamento zero, imediatamente anterior à curva ascendente do primeiro ciclo regular, confrontando no espectrograma com a visualização clara dos dois primeiros formantes. O final da vogal foi marcado na forma da onda, no ponto de cruzamento zero, imediatamente anterior à curva ascendente do último ciclo regular no ponto onde no espectrograma ainda era possível observar os dois primeiros formantes (Figura 23).

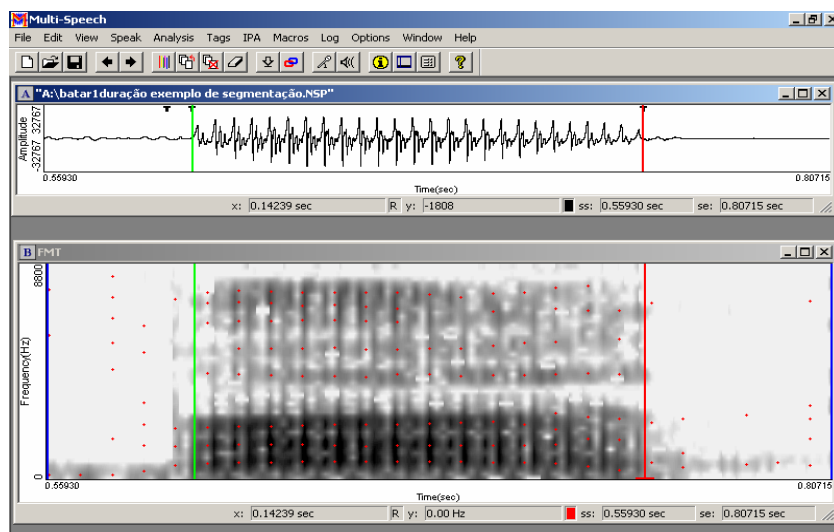


Figura 23- Exemplo de medida de duração da vogal adjacente – vogal tônica [a] - repetição 1 da frase *diga bata baixinho* produzida pelo sujeito com audição normal - AN

- Duração da frase - O início da frase foi considerado como o início da primeira vogal ([i] da palavra *diga*), uma vez que não é possível determinar o início da obstrução da africada [dʒ]. O início da vogal [i] seguiu os mesmos critérios descritos para a vogal adjacente. Nas figuras 24 e 25, os cursores indicam os pontos iniciais e finais de duas frases e as marcações na forma da onda, referem-se à delimitação dos segmentos para a realização não só das medidas de duração como também das medidas de *VOT*, energia da consoante e da vogal.

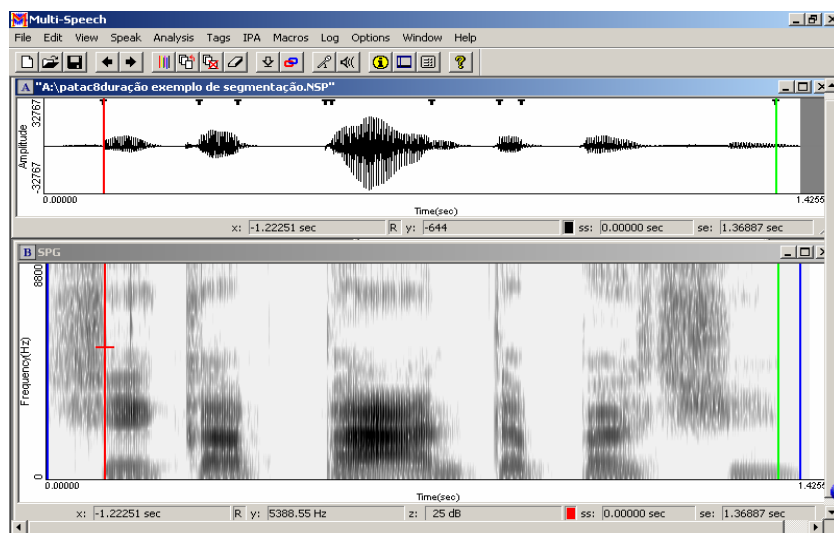


Figura 24- Exemplo de segmentação da frase - *diga pata baixinho* - para realização das medidas de duração, VOT e energia – produzida pelo sujeito com deficiência auditiva - DA

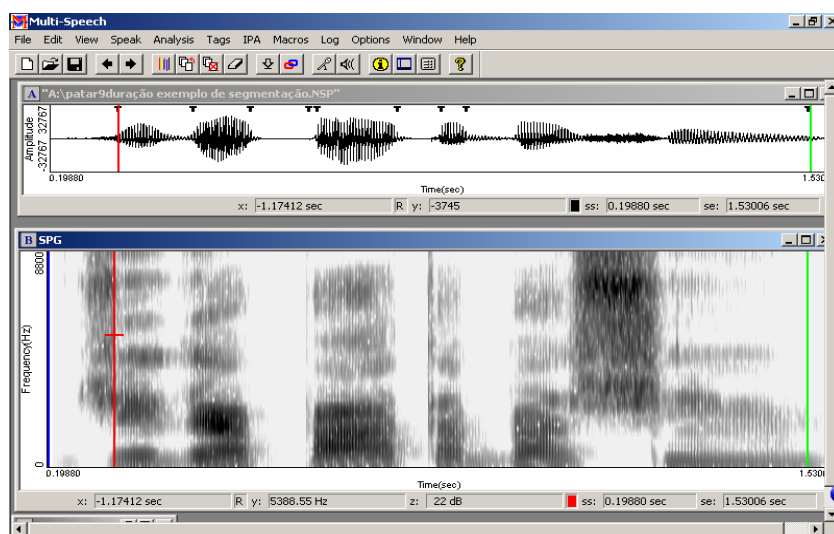


Figura 25 – Exemplo de segmentação da frase - *diga pata baixinho* - para realização das medidas de duração, VOT e energia – produzida pelo sujeito com audição normal - AN

Energia da consoante e das vogais adjacentes - As consoantes e as vogais segmentadas segundo o mesmo critério já exposto no procedimento de medida da duração foram selecionadas a partir da forma da onda e do espectrograma de banda larga. Em seguida, foi gerado o Contorno de Energia (*Energy Contour*) referente à porção selecionada. Depois, eram gerados os resultados estatísticos e

o valor da *root mean square* (rms) da energia em dB era anotado e a tabela gerada impressa para posterior conferência e tabulamento.

Frequência dos formantes das vogais adjacentes - Foram extraídos os formantes da vogal [a], nas suas ocorrências próximas à plosiva. Na posição pós-tônica da palavra *diga*, que antecede a plosiva, na posição tônica e pós-tônica da palavra alvo. A partir do espectrograma de banda larga, do *Formant History* e da forma da onda, foi selecionado o ponto da vogal que corresponde ao seu segmento estacionário, o que correspondeu, de forma geral, ao seu ponto médio. Após a seleção do ponto, as frequências do primeiro, segundo e terceiro formantes foram determinadas com base no LPC, com sobreposição do FFT.

Transição dos formantes - A transição dos formantes foi medida seguindo o procedimento proposto por Kewley-Port (1982). As frequências dos formantes 1, 2 e 3 foram medidas a partir do início da vogal subsequente até o ponto que correspondia a cerca de 40% da duração da vogal. Cabe observar que, por ocasião da realização dessa medida os dados referentes à duração já estavam tabulados e analisados. No caso das produções do sujeito com audição normal essa porcentagem corresponde a 56 ms em média e para o sujeito com deficiência auditiva, 80 ms. O LPC foi gerado em cada pico de máximo de cada ciclo da vogal durante esses tempos respectivamente, o que resultou no caso do sujeito com audição normal, 11 medidas e, no caso do sujeito com deficiência auditiva, 20 medidas. Essa diferença decorre não apenas da duração mais longa da vogal tônica no caso do sujeito com deficiência auditiva, mas também em virtude da frequência fundamental desse último ter se apresentado mais alta e, portanto, determinar um maior número de ciclos por segundo. Os parâmetros utilizados foram os mesmos sugeridos pela autora: LPC: *frame length* - 20 ms; *filter order* - 24, que está de acordo com a frequência de amostragem de nossos dados; tipo de *window weighting* - *Hanning* (a autora usa *Haming*). Uma diferença em relação ao procedimento da autora citada se refere ao fato que ela mediu os formantes a cada 5 ms. Como em nosso caso o posicionamento do cursor para geração do LPC foi feito manualmente, optou-se por posicioná-lo no pico de cada ciclo, de forma a facilitar a tarefa. No caso do sujeito com audição normal, ocorreu quase uma correspondência a esse valor uma vez que, 11 medidas foram feitas em 56 ms, na

média. No caso do sujeito com audição normal, esse procedimento aumentou o número de pontos medidos (20 medidas). O FFT foi sobreposto, usando os mesmos parâmetros utilizados na medida do formantes das vogais. Na figura 26 pode ser visualizada na forma da onda a marcação dos pontos, a partir dos quais foi gerado o LPC e o FFT para a medida dos formantes em cada um deles. Além disso, pode-se visualizar o cursor no pico do primeiro formante obtido no primeiro pulso da vogal subsequente à plosiva velar [g] produzida pelo sujeito com deficiência auditiva.

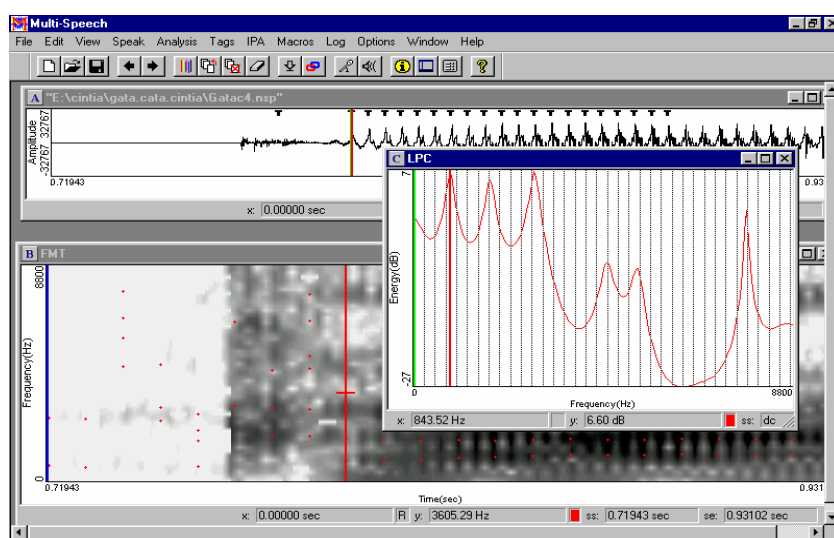


Figura 26 - Exemplo da realização da medida da transição dos formantes – LPC gerado a partir do primeiro pulso da vogal tônica [a] da palavra *gata* da repetição 4 produzida pelo sujeito com deficiência auditiva

Apresentaremos no próximo capítulo os resultados encontrados nas tarefas de identificação/discriminação auditiva e na análise acústica.

4 RESULTADOS

Após termos realizado a revisão da literatura acerca do tema desta pesquisa, cuja metodologia foi apresentada no capítulo anterior, procedemos, neste capítulo, à apresentação dos resultados das tarefas de identificação e discriminação para os sons plosivos e das medidas acústicas realizadas para a análise da produção das plosivas pelos dois sujeitos, um com deficiência auditiva e o outro com audição normal, cuja caracterização encontra-se no item 3.1.2.

4.1 Percepção das plosivas

Os resultados das tarefas de identificação auditiva das consoantes plosivas em posição inicial de sílaba tônica de palavras paroxítonas do PB serão aqui mostrados. Inicialmente, explicitamos os resultados da primeira tarefa realizada (tarefa 1), que diz respeito à percepção, por parte de um grupo de sujeitos com audição normal, das produções do sujeito com deficiência auditiva, com cuja fala aquele grupo não está familiarizado. Depois, são mencionados os resultados da segunda tarefa (tarefa 2), que consiste na percepção por um grupo de sujeitos com audição normal, das produções do sujeito com audição normal. Finalmente, serão apresentados os resultados das duas tarefas realizadas pelo sujeito com deficiência auditiva por meio do procedimento de avaliação da discriminação das plosivas do PB, desenvolvido em paralelo a este estudo: a primeira refere-se à discriminação das produções do sujeito com audição normal (tarefa 3) e a segunda, à discriminação de suas próprias produções (tarefa 4). Para essa última tarefa foi realizada uma adaptação no procedimento de discriminação das plosivas desenvolvido, tendo sido substituídos os arquivos de áudio contendo as produções do sujeito com audição normal pelos arquivos em que estão gravadas as produções do sujeito com deficiência auditiva.

Poucas produções de plosivas vozeadas foram identificadas corretamente: [d] 11,7%(7 acertos); [b] 6,7%(4 acertos). As produções da plosiva vozeada [g] não foram identificadas corretamente nenhuma das vezes. Ao analisarmos as confusões de identificação, observamos que a maior parte das produções da bilabial vozeada [b] foi identificada como uma bilabial não-vozeada [p] (78,3% - 46 vezes) e a da velar vozeada [g], como a velar não-vozeada [k] (91,7% - 55 vezes). Entretanto, a alveolar [d], apesar de ser a vozeada identificada maior número de vezes como tal, foi identificada como seu par não-vozeado apenas 56,7% (34) das vezes. Em 8,3% das apresentações (5), foi identificada como a bilabial não vozeada [p] e em 3,3% (2) como a vozeada [b]; em 13,3% (8) das apresentações foi identificada como a velar não vozeada [k] e em 6,7% como uma consoante não plosiva (Tabela 4).

Resultados da tarefa 2

Identificação pelo grupo com audição normal das plosivas produzidas pelo sujeito com audição normal

Todas as produções do sujeito ouvinte foram identificadas corretamente (100% de acerto) pelos 60 sujeitos com audição normal.

Resultados das tarefas 3 e 4

Discriminação realizada pelo sujeito com deficiência auditiva

Os resultados das duas tarefas de discriminação das plosivas do PB realizadas pelo sujeito com deficiência auditiva estão apresentadas nas tabelas 5 e 6. Primeiramente, os resultados do teste de discriminação das plosivas das produções do sujeito com audição normal e, na tabela 6, os resultados do teste adaptado. Com relação à discriminação do ponto de articulação nas produções do

sujeito com audição normal pelo sujeito com deficiência de audição (tarefa 3), podemos observar que o contraste mais bem discriminado foi o do ponto de articulação bilabial x velar - 95% de acerto, seguido da oposição alveolar x velar. O pior resultado foi para a oposição bilabial x alveolar – 35%.

No que se refere ao contraste de vozeamento, verificou-se que o sujeito com deficiência auditiva não o discrimina auditivamente (apenas 0,2% de acerto acima do nível de chance). Os dados sobre a tarefa de percepção na qual o sujeito com deficiência auditiva deveria identificar suas próprias produções (tarefa 4) demonstram, de forma geral, um menor número de respostas corretas em todos os contextos, tanto para o contraste de vozeamento quanto para o ponto de articulação. Com relação ao contraste de ponto de articulação não ocorreram respostas corretas (descontando o nível de acerto casual) para as oposições bilabial/alveolar e alveolar/velar. Observamos 20% de respostas corretas no contraste bilabial/velar. Cabe-nos ressaltar que tal porcentagem de acerto refere-se às respostas corretas para os itens correspondentes às bilabiais. Como pode ser visto na tabela 6, na oposição bilabial x alveolar, quase todas os itens foram identificados como bilabiais, sendo apenas 4 itens iniciados com velares identificados corretamente.

Tabela 5 - RESPOSTAS DO SUJEITO COM DEFICIÊNCIA AUDITIVA PARA A TAREFA DE DISCRIMINAÇÃO DAS PLOSIVAS PRODUZIDAS PELO SUJEITO COM AUDIÇÃO NORMAL

APRESENTAÇÃO

R E S P O S T A				
	pata	bata	tata	data
	pata	7	3	5
	bata	1	5	0
	tata	2	1	3
	data	0	1	2
	pata	bata	cata	gata
	pata	9	4	0
	bata	1	5	0
	cata	0	1	3
	gata	0	0	7
				tata
				data
				cata
				gata
				tata
				data
				cata
				gata

Porcentagem de acerto (acima da chance) por contraste de ponto de articulação nas três oposições – bilabial x alveolar; bilabial x velar; alveolar x velar

oposição	ponto
bilabial x alveolar	35%
bilabial x velar	95%
alveolar x velar	60%

Porcentagem de acerto (acima da chance) para o contraste de sonoridade

oposição	acerto
sonoridade	0,2%

Tabela 6 - RESPOSTAS DO SUJEITO COM DEFICIÊNCIA AUDITIVA PARA A TAREFA DE DISCRIMINAÇÃO DAS PLOSIVAS PRODUZIDAS POR ELE PRÓPRIO

APRESENTAÇÃO

R E S P O S T A				
	pata	bata	tata	data
	pata	3	1	1
	bata	2	5	4
	tata	1	3	2
	data	4	1	3
				1
	pata	bata	cata	gata
	pata	8	3	5
	bata	2	7	5
	cata	0	0	0
	gata	0	0	0
				4
	tata	data	cata	gata
	tata	0	5	4
	data	2	1	2
	cata	6	0	0
	gata	2	4	4

Porcentagem de acerto (acima da chance) por contraste de ponto de articulação nas três oposições – bilabial x alveolar; bilabial x velar; alveolar x velar

oposição	acerto
bilabial x alveolar	0%
bilabial x velar	20%
alveolar x velar	0%

Porcentagem de acerto (acima da chance) para o contraste de sonoridade

oposição	acerto
sonoridade	0,1%

Quanto ao contraste de sonoridade, também não ocorreu acerto acima do nível do casual (0,1% de respostas corretas). Em relação aos resultados da tarefa de discriminação das produções do sujeito com audição normal, houve uma certa concordância quanto ao ponto de articulação – melhores resultados no contraste bilabial x alveolar – embora com uma grande diminuição do acerto no caso da

discriminação das próprias produções. Também com relação às suas próprias produções, o sujeito com deficiência auditiva não discriminou o contraste de vozeamento acima do nível de acerto por chance.

4.2 Produção das plosivas do PB

Os resultados das medidas acústicas realizadas com a finalidade de estudar a produção das plosivas considerando o sujeito com deficiência auditiva (DA) e o com audição normal (AN) serão apresentados a seguir. Em primeiro lugar são expostos os dados referentes ao tempo de ataque do vozeamento - VOT^{24} , frequência de primeiro formante no início da vogal subsequente (*onset* de F1), frequência fundamental no início da vogal subsequente (f_0 no *onset* da vogal). Em seguida, mostramos os dados sobre a duração da consoante e dessa medida em relação a outros segmentos, a saber: duração da frase, duração das vogais adjacentes (pós-tônica da palavra anterior, tônica da palavra e pós-tônica da palavra, duração da sílaba tônica). Posteriormente vêm os resultados referentes à energia da consoante em relação à vogal subsequente, energia das vogais adjacentes e frequência dos formantes (medidos na porção estacionária das vogais adjacentes). Finalmente, encontram-se os dados relacionados à transição dos formantes F1, F2 e F3 da consoante em direção à vogal subsequente.

Os dados foram submetidos à análise estatística para a comparação de médias por meio da análise de variância com um fator (*one way ANOVA*), com aplicação do teste *a posteriori* de Scheffe, quando da ocorrência de diferença significativa entre as médias.

4.2.1 Medidas de *VOT*

Os resultados das medidas de *VOT* das plosivas produzidas pelo falante com audição normal (tabela 7) mostraram que as plosivas vozeadas apresentam

²⁴ Os resultados das medidas de *VOT* e duração foram parcialmente analisados por Madureira et al (2002).

valores de *VOT* negativos, sendo que as bilabiais apresentaram os valores mais altos em módulo e as velares os mais baixos. A média para os valores das bilabiais é -90 ms; para as dentais/alveolares foram encontrados valores médios em -77 ms; para as velares, - 66 ms.

As plosivas não vozeadas, para esse mesmo sujeito, apresentaram valores positivos de *VOT*, sendo que a velar [k] teve valores mais altos que a dental/alveolar [t] e essa última teve valores mais altos que a bilabial [p]. Isso era esperado devido aos aspectos fisiológicos de produção descritos anteriormente. As médias para a medida do *VOT* das consoantes são: para as velares 31 ms, para as dentais/alveolares 19 ms e para as bilabiais 11 ms.

Tabela 7 - MEDIDAS DE *VOT* (ms) DAS PLOSIVAS PRODUZIDAS PELO SUJEITO COM AUDIÇÃO NORMAL

	PAta	BAta	TAta	DAta	CAta	GAta
média (d.p)	11 (3)	-90 (7)	19 (2)	-77 (6)	32 (5)	-66 (6)
cf. var.	0,29	-0,07	0,11	-0,08	0,16	-0,09

Legenda:
d.p.- desvio padrão
cf.var. - coeficiente de variação

Os valores de *VOT* produzidos pelo sujeito com audição normal diferem de acordo com o ponto de articulação e vozeamento. A análise estatística revela diferenças significativas entre todos os valores [$F_{(5,54)}=1116,85$, $p=0,000$]. O teste *a posteriori* de Scheffe aponta para o fato de os valores de *VOT* dos pares vozeados/não-vozeados diferirem significativamente, sendo entre [p, b] $p=0,000$, entre [t, d] $p=0,000$ e entre [k, g] $p=0,000$. Considerando a oposição por ponto de articulação, verificamos o seguinte: para a oposição [p, t], $p=0,046$; para [b, d], $p=0,000$; para [p, k], $p=0,000$; para [b, g], $p=0,000$; para [t, k], $p=0,000$; para [d, g], $p=0,002$.

Entretanto, para o sujeito deficiente auditivo, os valores de *VOT* (tabela 8) para as bilabiais (considerando as não-vozeadas) sobrepõem-se parcialmente aos valores das alveolares, conforme demonstra o gráfico apresentado na figura 27.

Tabela 8 - MEDIDAS (ms) PARA O SUJEITO COM DEFICIÊNCIA AUDITIVA: MANUTENÇÃO DA BARRA DE SONORIDADE (M.B.S.); INTERRUÇÃO DA BARRA DE SONORIDADE APÓS A VOGAL PRECEDENTE (I.B.S.); VOT POSITIVO (VOT+)

A. bilabiais

	PAta	PAta	PAta	BAta	BAta	BAta
	M.B.S.	I.B.S.	VOT+	M.B.S.	I.B.S.	VOT+
média (d.p.)	26 (12)	163 (29)	13 (2)	31 (12)	152 (14)	12 (3)
cf.var.	0,45	0,18	0,15	0,37	0,09	0,24

B. alveolares

	TAta	TAta	TAta	DAta	DAta	DAta
	M.B.S.	I.B.S.	VOT+	M.B.S.	I.B.S.	VOT+
média (d.p.)	36 (14)	167 (10)	15 (5)	35 (12)	164 (10)	14 (3)
cf.var.	0,39	0,06	0,34	0,33	0,06	0,19

C. velares

	CAta	CAta	CAta	GAta	GAta	GAta
	M.B.S.	I.B.S.	VOT+	M.B.S.	I.B.S.	VOT+
média (d.p.)	16 (3)	157 (31)	37 (6)	19 (4)	144 (25)	35 (4)
cf.var.	0,21	0,19	0,16	0,24	0,17	0,11

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

As medidas de manutenção da barra de sonoridade e de interrupção da barra de sonoridade após a vogal precedente foram realizadas com o objetivo de investigar se ocorria um maior tempo de manutenção da barra de vozeamento nas plosivas nas quais ele deveria ocorrer, ou seja, nas plosivas vozeadas do PB. Contudo, não foram observadas diferenças significativas entre essas medidas, considerando-se os pares de vozeadas e não-vozeadas nas produções desse sujeito com deficiência auditiva. A análise estatística, considerando a manutenção da barra de sonoridade (M.B.S.), nos permitiu verificar que houve diferenças significativas entre as plosivas, mas não entre os pares vozeados/não vozeados [$F_{(5,54)} = 6,612$, $p=0,000$; teste *a posteriori* de Scheffe: para comparação de [p] e [b], $p=0,961$; para [t] e [d], $p=1,000$; para [k] e [g], $p=0,997$]. Em relação à medida da interrupção da barra de sonoridade, não ocorreram diferenças significativas entre

as seis plosivas [$F_{(5,54)}=1,567$, $p=0,185$]. Esse resultado indica que, do ponto de vista dessas duas medidas, não se observou na produção do sujeito com deficiência de audição diferenciação entre as plosivas vozeadas e não-vozeadas.

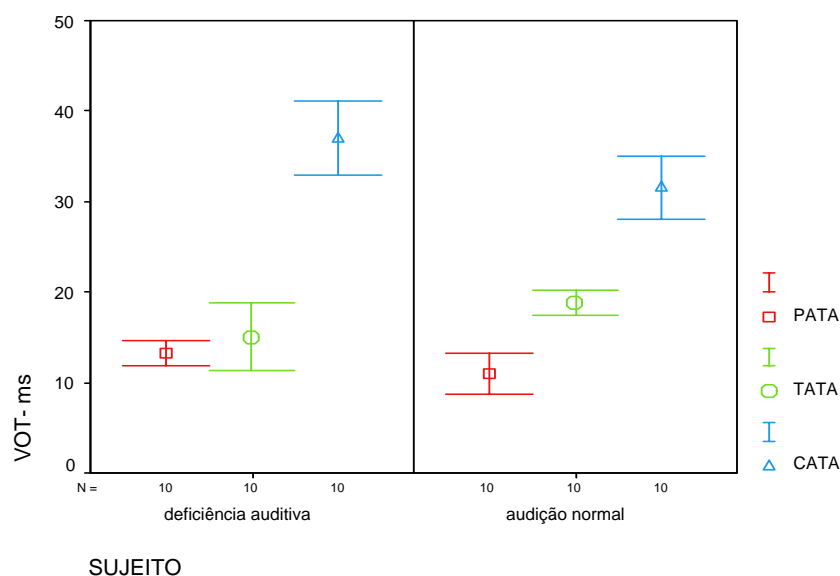


Figura 27 - Gráfico representativo dos valores de VOT para [p, t, k] produzidas pelos dois sujeitos

Nas produções do sujeito com deficiência auditiva, não foram encontrados valores de *VOT* negativo ou zero para as plosivas vozeadas, que se comportaram como se fossem não-vozeadas (figura 28). Os valores de *VOT* foram ligeiramente maiores para as plosivas não-vozeadas; essa diferença porém, não foi significativa. A análise estatística aponta para a diferenciação nas medidas de *VOT* nas produções do sujeito com deficiência auditiva apenas no que se refere ao ponto de articulação velar, que se diferenciou do ponto bilabial e alveolar, como já foi apontado também no gráfico da figura 27 [$F_{(5,54)}=85,177$, $p=0,000$]. O teste *a posteriori* de Scheffe para a comparação dos pares vozeados/não-vozeados apresenta os seguintes resultados: para a oposição [p] e [b], $p=0,979$; para [t] e [d], $p=0,998$; para [k] e [g], $p=0,870$. Já para a comparação entre os pontos de articulação: [p] e [t], $p=0,958$; para [p] e [k], $p=0,000$; para [t] e [k], $p=0,000$; [b] e [d], $p=0,867$; para [b] e [g], $p=0,000$; para [d] e [g], $p=0,000$.

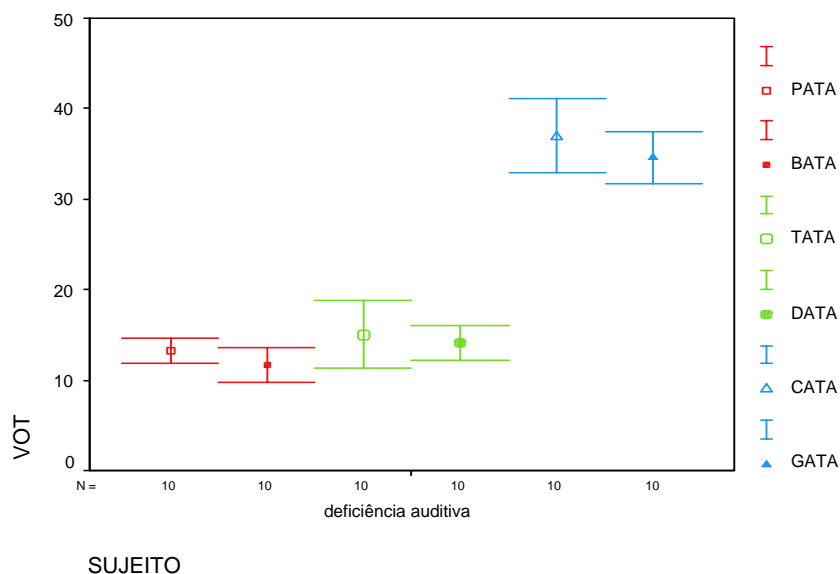


Figura 28 - Gráfico representativo dos valores de *VOT* para as bilabiais, alveolares e velares produzidas pelo sujeito com deficiência auditiva

Na análise estatística, não se encontrou diferença significativa entre as medidas de *VOT* da plosiva bilabial vozeada [b] produzida pelo sujeito com deficiência auditiva, e da bilabial não-vozeada [p] produzida pelo sujeito com audição normal. Também entre os valores de *VOT* da velar vozeada [g] produzida pelo sujeito com deficiência auditiva e a velar não-vozeada [k] produzida pelo sujeito com audição normal, não foi identificada diferença significativa. Entretanto, encontramos diferença significativa [$t_{(9)}=4,616$, $p=0,001$] entre as medidas de [d] produzida pelo sujeito com deficiência auditiva e de [t] produzida pelo falante com audição normal (figura 29). Cabe-nos lembrar que não há diferenciação entre a bilabial e alveolar produzida pelo sujeito com deficiência auditiva.

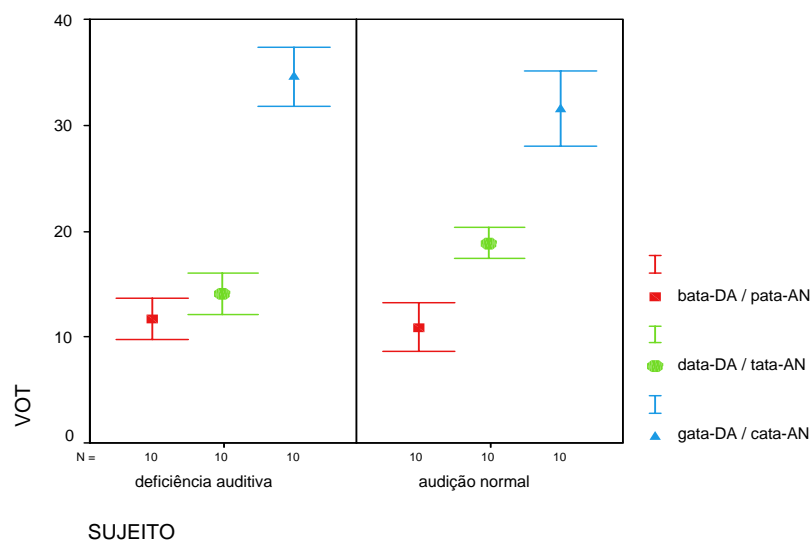


Figura 29 - Valores de VOT (ms) para [b, d, g] do sujeito com deficiência auditiva e valores de VOT (ms) para [p, t, k] do sujeito com audição normal

No gráfico apresentado na figura 30 pode-se observar uma síntese dos valores de VOT das seis plosivas estudadas para os dois sujeitos. Fica clara a diferença entre os sujeitos no que se refere às plosivas vozeadas. Também evidencia-se a não-diferenciação do VOT das bilabiais e alveolares na produção do sujeito com deficiência de audição:

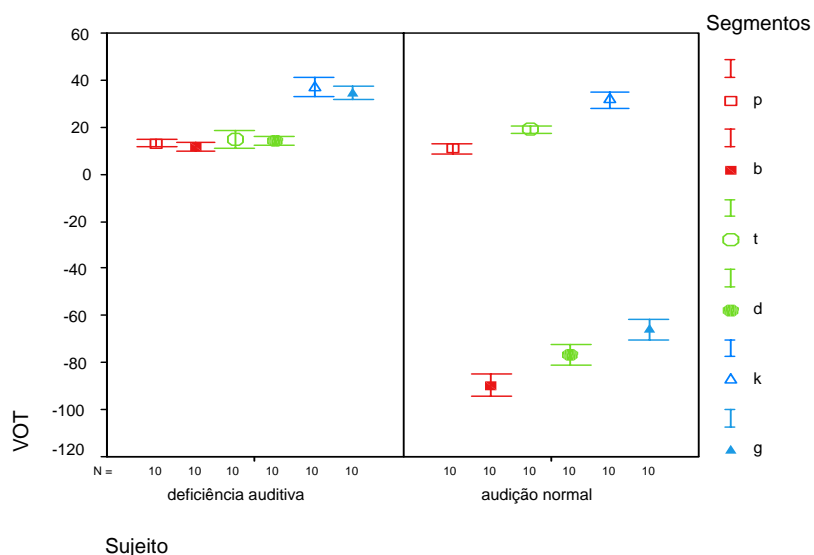
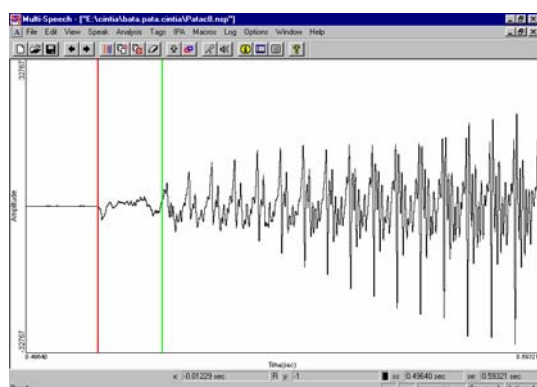


Figura 30 - Gráfico representativo dos valores de VOT para os dois sujeitos considerando as seis plosivas estudadas

A seguir, há alguns exemplos de espectrogramas com as medidas de VOT para os dois sujeitos (figuras 31, 32, 33, 34, 35 e 36). À esquerda estão os dados referentes ao sujeito com deficiência auditiva e à direita os do sujeito com audição normal.

VOT de [p] na palavra *pata*

Sujeito com deficiência auditiva



Sujeito com audição normal

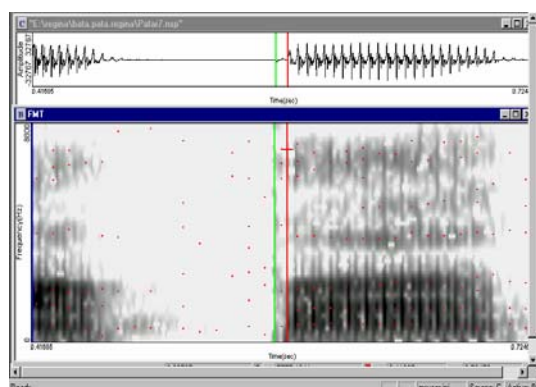
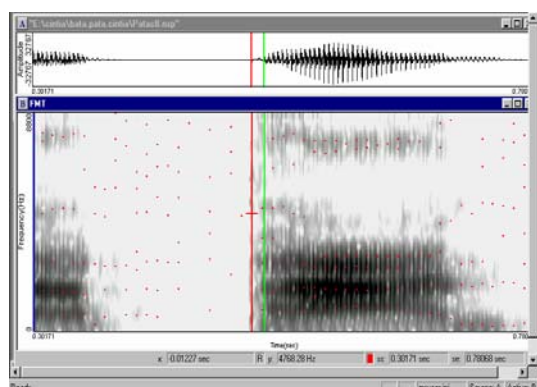
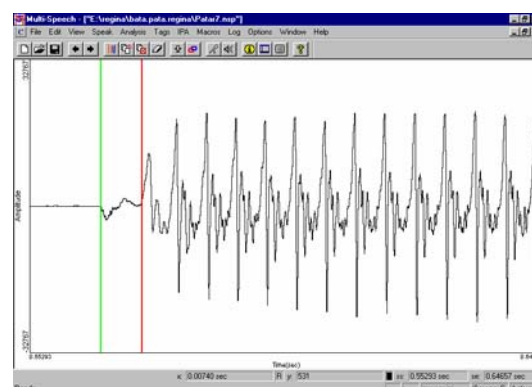
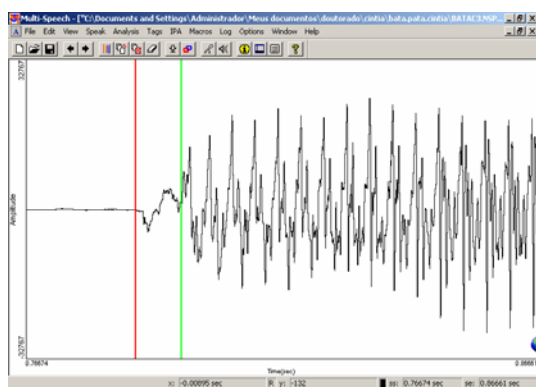


Figura 31 - Espectrogramas - medida de VOT na produção da bilabial não-vozeada [p]

VOT de [b] na palavra *bata*

Sujeito com deficiência auditiva



Sujeito com audição normal

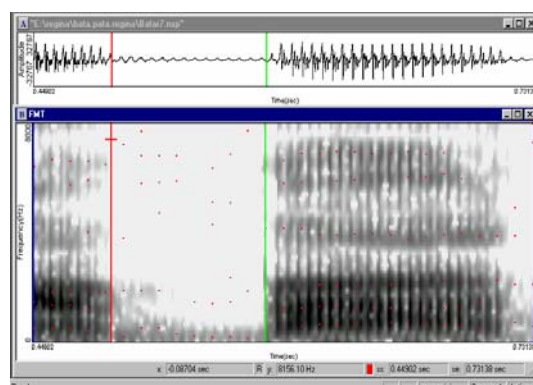
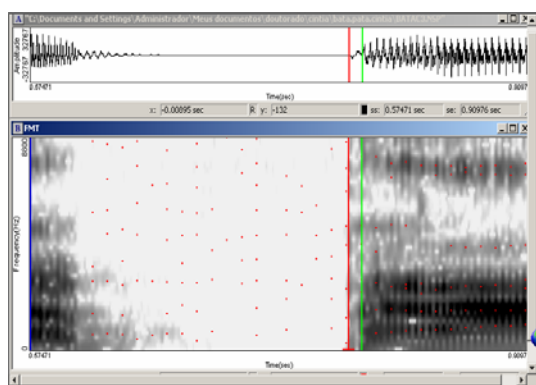
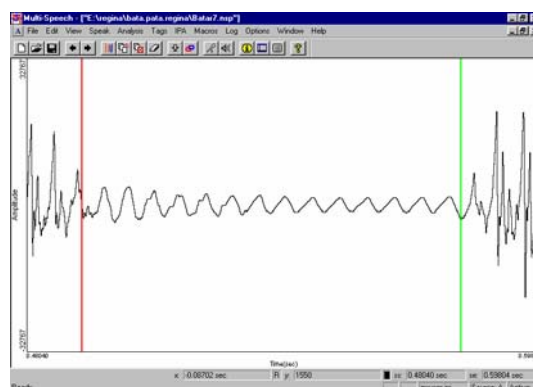
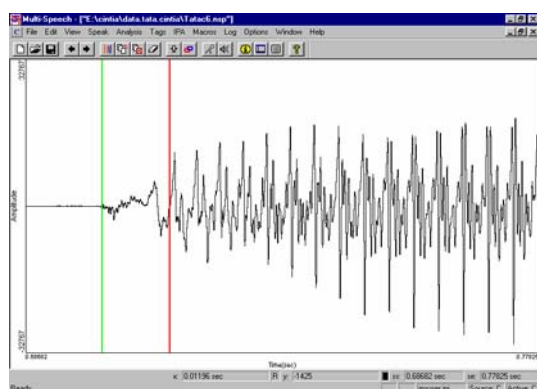


Figura 32 - Espectrogramas - medida de VOT na produção da bilabial vozeada [b]

VOT de [t] na palavra *tata*

Sujeito com deficiência auditiva



Sujeito com audição normal

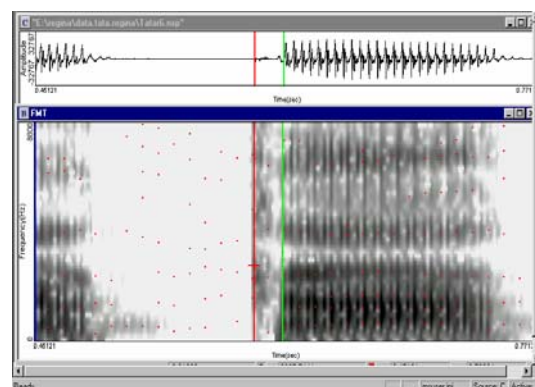
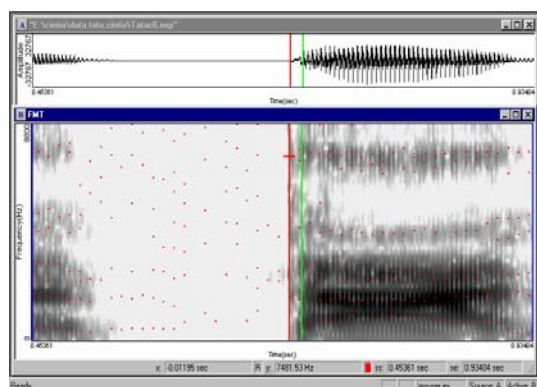
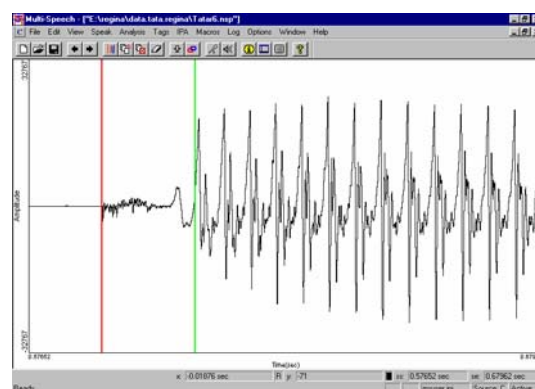
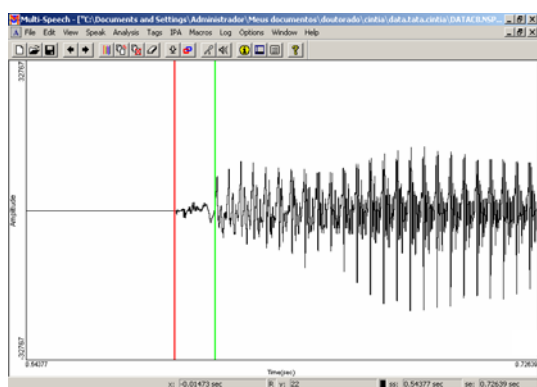


Figura 33 - Espectrogramas - medida de VOT na produção da alveolar não-vozeada [t]

VOT de [d] na palavra *data*

Sujeito com deficiência auditiva



Sujeito com audição normal

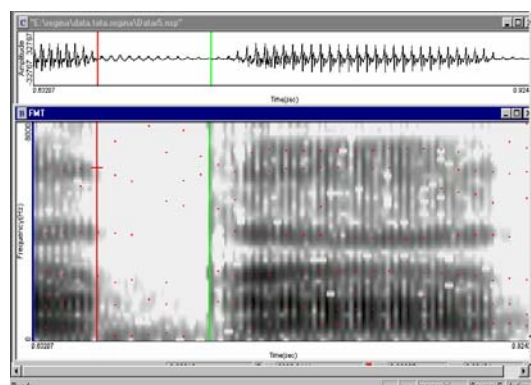
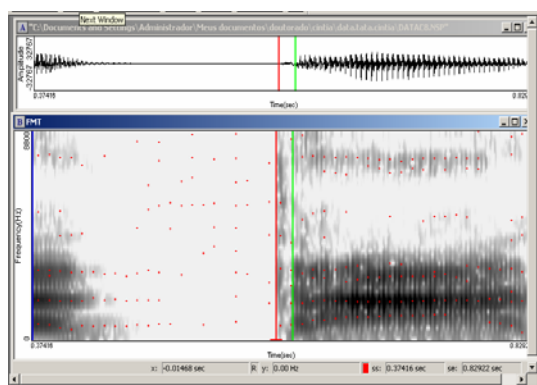
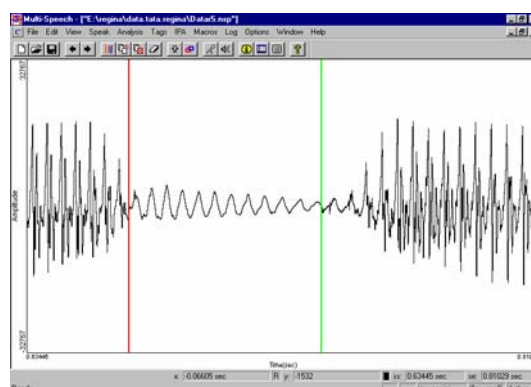
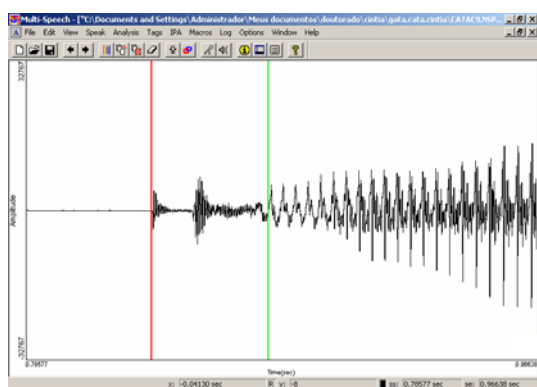


Figura 34 - Espectrogramas - medida de VOT na produção da alveolar vozeada [d]

VOT de [k] na palavra *cata*

Sujeito com deficiência auditiva



Sujeito com audição normal

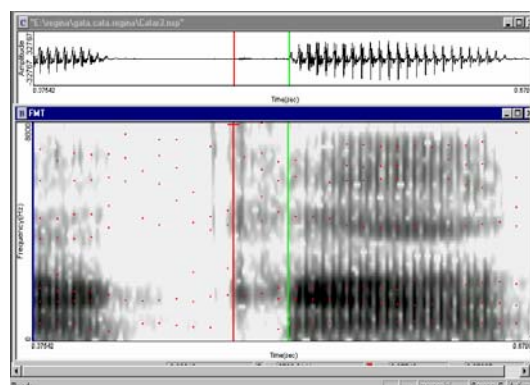
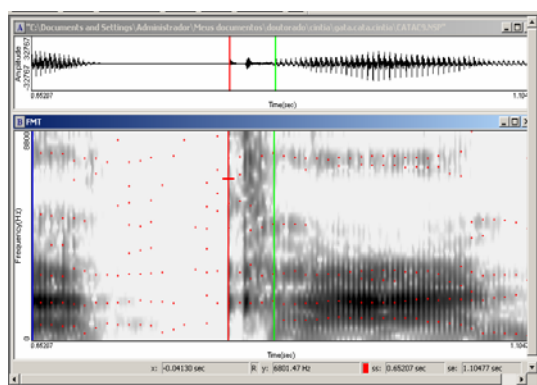
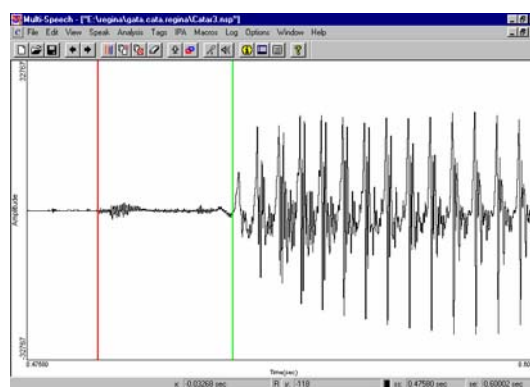
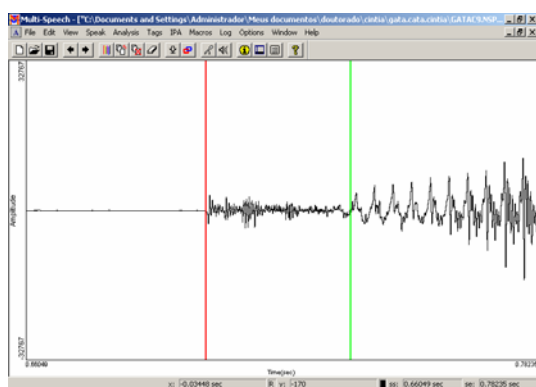


Figura 35 - Espectrogramas - medida de VOT na produção da velar não-vozeada [k]

VOT de [g] na palavra *gata*

Deficiência auditiva



Audição normal

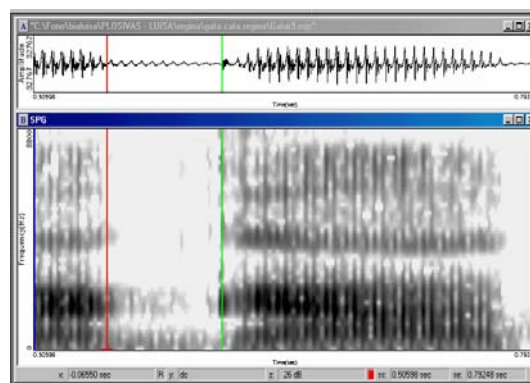
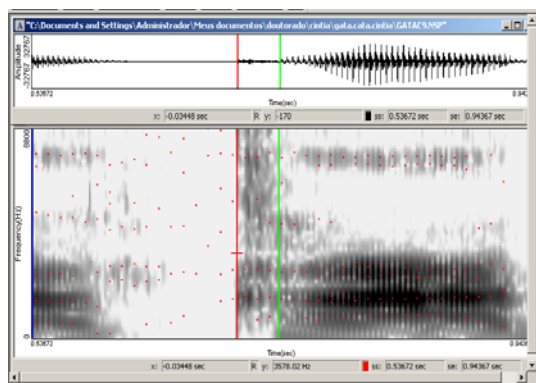
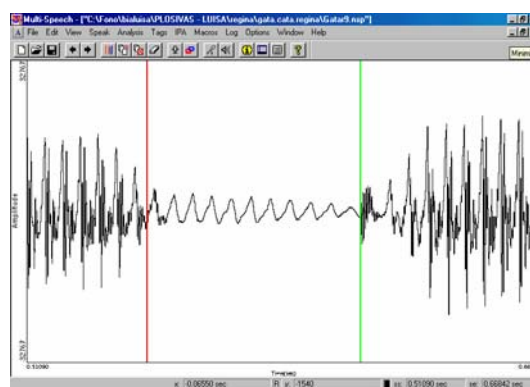


Figura 36 - Espectrogramas - medida de VOT na produção da velar vozeada [g]

4.2.2 Medidas do onset de F1

Os resultados da medida da frequência no *onset* de F1, para o sujeito com audição normal seguiram a tendência da literatura, ou seja, os valores da frequência no *onset* do F1 foram mais altos após plosivas não-vozeadas (Kluender, 1991; Shimizu, 1996; Benki, 2001). A comparação de médias mostrou diferenças significativas entre os valores de F1 no início da vogal subsequente entre os pares vozeados/não-vozeados [$F_{(5,54)} = 119,970$, $p = 0,000$]. O teste *a posteriori* de Scheffe revelou que para a comparação entre [p] e [b], $p = 0,013$; entre [t] e [d], $p = 0,000$; e entre [k] e [g], $p = 0,000$.

Para o sujeito com deficiência auditiva, a frequência no *onset* do F1 das plosivas vozeadas e não-vozeadas não apresentou diferenças significativas entre os pares vozeados e não-vozeados, ocorrendo apenas algumas diferenças por ponto de articulação [$F_{(5,54)} = 16,505$, $p = 0,000$] (tabela 9). No teste *a posteriori* de Scheffe observamos que para a comparação entre [p] e [b] $p = 0,106$; entre [t] e [d] $p = 0,997$; e entre [k] e [g] $p = 1,000$.

Tabela 9 - VALORES DA FREQUÊNCIA NO ONSET DE F1 PARA OS DOIS SUJEITOS

Sujeitos		Onset F1 [pa]	Onset F1 [ba]	Onset F1 [ta]	Onset F1 [da]	Onset F1 [ca]	Onset F1 [ga]
AN	média (d.p.)	706 (28)	665 (34)	634 (20)	546 (18)	615 (16)	485 (19)
	cf. var.	0,04	0,05	0,03	0,03	0,03	0,04
DA	média (d.p.)	830 (74)	751 (60)	712 (36)	697 (45)	861 (55)	857 (66)
	cf. var.	0,09	0,08	0,05	0,06	0,06	0,08

Legenda:

d.p. - desvio padrão

cf. var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Considerando os dois sujeitos, podemos dizer que todos os valores da frequência no *onset* de F1 foram mais altos nas produções do sujeito com deficiência auditiva se comparados aos valores das produções do outro falante ($p < 0,001$), como se verifica no gráfico da figura 37. Convém observar que isso depende de características individuais e representa uma variação comumente

encontrada entre falantes.

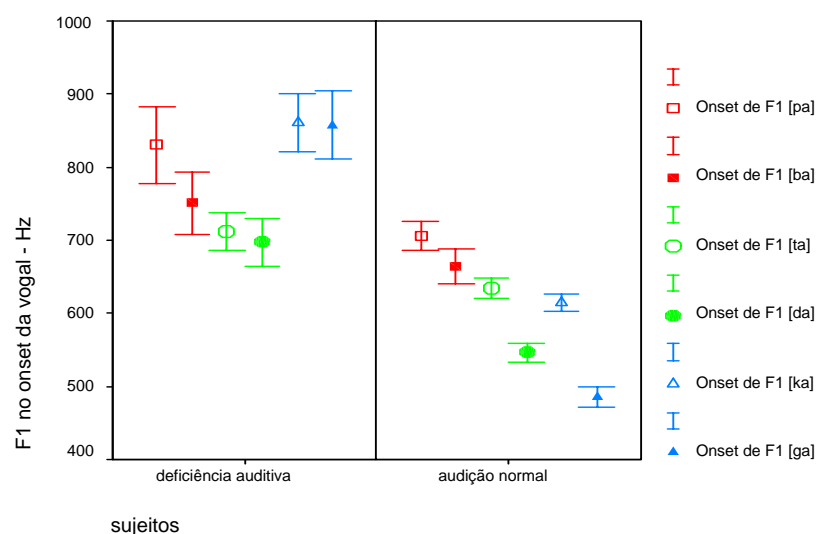


Figura 37 - Gráfico representativo da distribuição dos valores da frequência no *onset* do F1 para os dois sujeitos

4.2.3 Medidas de f0 no início da vogal subsequente

Os valores da frequência fundamental no início da vogal subsequente encontrados nas produções do sujeito com audição normal estão na tabela 10. Podemos observar que eles são compatíveis com o esperado, isto é, o f0 sofre uma queda no início do vozeamento da vogal quando ela sucede uma plosiva vozeada (Shimizu, 1996; Holt et al, 2001). As diferenças encontradas para os pares vozeados e não-vozeados nos três pontos de articulação são estatisticamente significativas [$F_{(5,54)}=30,511$, $p=0,000$]. No teste *a posteriori* de Scheffe, encontramos para a comparação entre [p] e [b], $p=0,000$; entre [t] e [d], $p=0,000$; entre [k] e [g], $p=0,000$. Não foram observadas diferenças significativas entre os três pontos de articulação nas vozeadas e não-vozeadas.

Tabela 10 - FREQUÊNCIA FUNDAMENTAL (f0) NO *ONSET* DA VOGAL SUBSEQÜENTE PARA OS DOIS SUJEITOS

Sujeitos		Pata f0	Bata f0	Tata f0	Data f0	Cata f0	Gata f0
AN	média (d.p.)	213 (17)	165 (80)	223 (25)	174 (8)	214 (14)	172 (7)
	cf. var.	0,08	0,05	0,11	0,05	0,07	0,04
DA	média (d.p.)	254 (19)	252 (15)	263 (13)	260 (22)	244 (5)	236 (12)
	cf. Var	0,08	0,06	0,05	0,08	0,02	0,05

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Para o sujeito com deficiência auditiva, não houve diferenças significativas entre os valores de f0 nos pares vozeado e não-vozeado. Apenas ocorreu diferença significativa entre alguns pontos de articulação [$F_{(5,54)}=4,387$, $p=0,002$]. Pelo teste *a posteriori* de Scheffe, considerando a comparação entre os pares vozeado/não-vozeado, encontramos para [p] e [b], $p=1,000$; para [t] e [d], $p=0,999$; e para [k] e [g], $p=0,921$.

No gráfico da figura 38, a distribuição dos valores de f0 para os dois sujeitos pode ser bem visualizada. Ficam evidentes as diferenças associadas ao contraste de vozeamento nas produções do sujeito com audição normal e, por outro lado, a ausência de distinção nas produções do sujeito com deficiência auditiva:

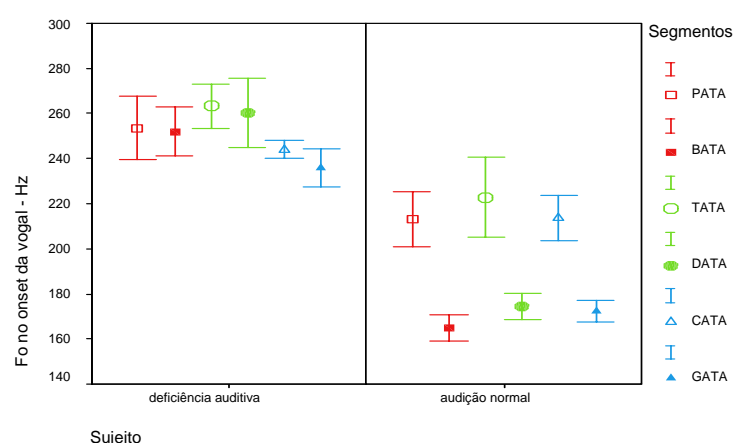


Figura 38 - Gráfico representativo dos valores de f0 no início (*onset*) da vogal subsequente para os dois sujeitos nos seis contextos estudados

A seguir, estão as medidas de f_0 no meio da vogal tônica subsequente à plosiva para o sujeito com audição normal e para o sujeito com deficiência auditiva (tabela 11).

Tabela 11 - FREQUÊNCIA FUNDAMENTAL (f_0) NO MEIO DA VOGAL SUBSEQÜENTE PARA OS DOIS SUJEITOS

Sujeito		f_0 meio [pa]	f_0 meio [ba]	f_0 meio [ta]	f_0 meio [da]	f_0 meio [ka]	f_0 meio [ga]
AN	média(d.p.)	185(5)	181(5)	187(5)	180(5)	187(5)	181(5)
	cf.var.	0,03	0,03	0,03	0,03	0,03	0,03
DA	média (d.p.)	221(8)	223(8)	216(10)	213(8)	226(7)	229(8)
	cf.var.	0,04	0,04	0,04	0,04	0,03	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Para o sujeito com audição normal, o valor de f_0 no meio da vogal tônica subsequente às consoantes plosivas foi, em média de 184 Hz e não foram encontradas diferenças significativas entre os valores de acordo com o contexto da consoante. Se compararmos as medidas de f_0 no início e no meio da vogal, podemos observar que, para todos os contextos de consoante plosiva vozeada, os valores de f_0 foram mais baixos no início e, para as consoantes não-vozeadas, os valores foram mais altos no início.

Já para o sujeito com deficiência auditiva, todas as medidas das frequências de f_0 no início da vogal subsequente foram mais altas se comparadas ao centro da vogal, apresentando um comportamento semelhante ao observado no contexto das consoantes não-vozeadas para o sujeito com audição normal. Esses dados podem ser visualizados no gráfico da figura 39:

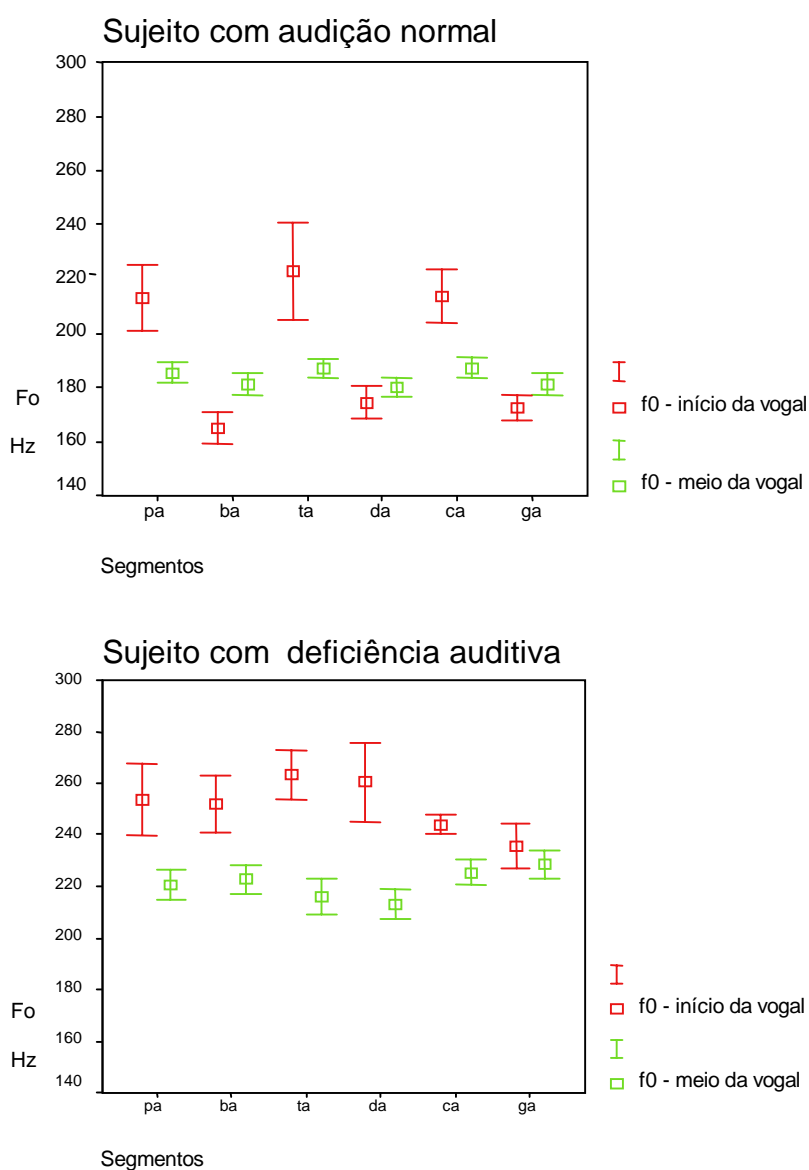


Figura 39 - Gráficos representativos das frequências de f0 medida no início (*onset*) da vogal e no meio da vogal subsequente às seis plosivas estudadas para os dois sujeitos

4.2.4 Medidas de duração das consoantes plosivas

Duração absoluta

Os valores das medidas de duração das plosivas produzidas pelo sujeito com audição normal estão de acordo com os dados encontrados na literatura, isto é,

foram encontrados valores de duração maiores para as plosivas não-vozeadas em relação aos seus pares vozeados (Barbosa,1996; Veloso,1995) [$F_{(5,54)} = 64,371$, $p = 0,000$]. No teste *a posteriori* de Scheffe, observou-se diferença significativa entre [p] e [b], $p=0,000$; entre [t] e [d], $p=0,000$; e entre [k] e [g], $p=0,000$. Quanto ao ponto de articulação, não foram encontradas diferenças significativas entre os valores no grupo de vozeadas e no grupo de não-vozeadas, com exceção da duração da vozeada [g], que se apresentou menor que [b] ($p=0,000$). Os valores detalhados estão apresentados na tabela 12 e ilustrados no gráfico da figura 40 a seguir:

Tabela 12 – DURAÇÃO (ms) DAS PLOSIVAS PRODUZIDAS PELO SUJEITO COM AUDIÇÃO NORMAL

	Pata	Bata	Tata	Data	Cata	Gata
média (d.p.)	118 (5)	99 (6)	120 (7)	91 (5)	116 (8)	83 (6)
cf.var.	0,05	0,06	0,06	0,06	0,07	0,07

Legenda:
d.p.- desvio padrão
cf.var. - coeficiente de variação

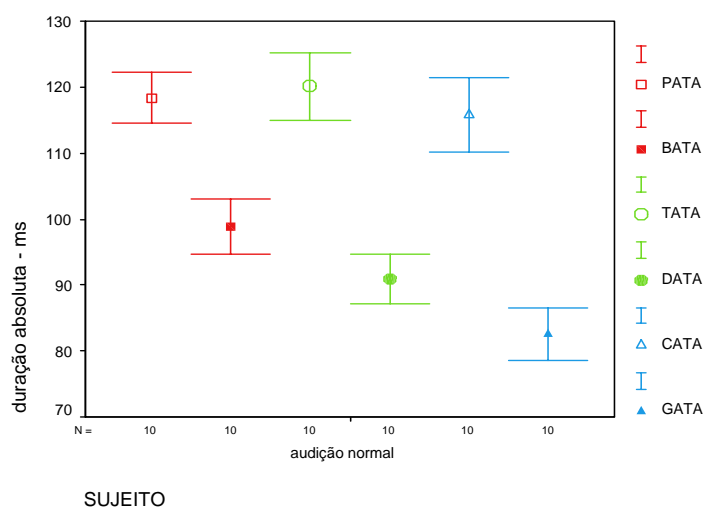


Figura 40 - Gráfico representativo das medidas de duração (ms) das plosivas produzidas pelo sujeito com audição normal

A diferença entre as medidas de duração dos pares bilabiais, alveolares e velares é 19, 29 e 33 ms, respectivamente.

Comparando os valores de duração das plosivas vozeadas e seus pares não-

vozeados, produzidas pelo sujeito com deficiência auditiva (tabela 13 e gráfico da figura 41), não foram encontradas diferenças significativas.

Tabela 13 - DURAÇÃO (ms) DAS PLOSIVAS PRODUZIDAS PELO SUJEITO COM DEFICIÊNCIA AUDITIVA

	Pata	Bata	Tata	Data	Cata	Gata
média (d.p.)	202 (26)	195 (11)	215 (16)	211 (13)	210 (26)	198 (25)
cf.var.	0,13	0,06	0,08	0,06	0,12	0,13

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

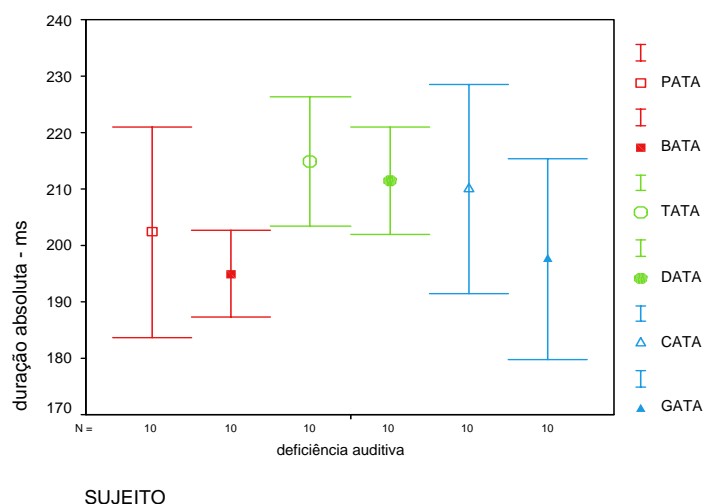


Figura 41 - Gráfico representativo das medidas de duração (ms) absoluta das plosivas produzidas pelo sujeito com deficiência auditiva

Esses dados apontam para a não-diferenciação por parte do sujeito com deficiência auditiva do contraste de vozeamento se considerarmos o parâmetro da duração da plosiva. Além disso, revelam um alongamento da consoante. Para investigarmos se tal alongamento restringe-se a uma tendência geral na produção do sujeito com deficiência auditiva, consideramos a duração das consoantes em relação aos outros segmentos do enunciado, ou seja, verificamos a duração relativa, calculando a que porcentagem da frase corresponde cada segmento medido. Para tanto, foram realizadas medidas de duração das frases, para os dois

sujeitos. Os valores médios das frases-veículo, envolvendo as seis plosivas estudadas, estão na tabela 14. Na análise estatística, não foram encontradas diferenças significativas entre as medidas de duração da frase, considerando-se cada sujeito.

Tabela 14 - MEDIDAS DE DURAÇÃO DAS FRASE (ms) PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		diga pata...	diga bata...	diga tata..	diga data...	diga cata...	diga gata...
AN	média (d.p.)	1183 (33)	1205 (73)	1247 (41)	1201 (41)	1227 (33)	1229 (41)
	cf.var.	0,03	0,06	0,03	0,03	0,03	0,03
DA	média (d.p.)	1266 (47)	1233 (69)	1299 (58)	1321 (106)	1303 (81)	1.303 (38)
	cf.var.	0,04	0,06	0,04	0,08	0,06	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

A relação entre as medidas de duração das frases produzidas foi calculada por meio da divisão dos valores médios encontrados para ambos os sujeitos (duração das produções do sujeito com deficiência auditiva/duração das produções do sujeito com audição normal). Os resultados foram os seguintes: frase-veículo com a palavra *pata* = 1/1,11; com *bata* = 1/ 1,02; com *tata* = 1/1,04; com *data* = 1/1,10; com *cata* = 1/1,06; com *gata* = 1/1,06.

As frases do sujeito com deficiência auditiva são mais longas, segundo revelam os dados, mas não justificam o alongamento da consoante, para o qual encontramos as seguintes relações (duração da consoante produzida pelo sujeito com deficiência auditiva / duração da consoante produzida pelo sujeito com audição normal): para a consoante [p] = 1/1,71; para [b] = 1/1,97; para [t] = 1/1,79; para [d] = 1/ 2,32; para [k] = 1/1,81; para [g] = 1/1,39. Apresentamos na tabela 15, a seguir, as medidas da duração relativa das consoantes para os dois sujeitos, ilustradas no gráfico da figura 42.

Tabela 15 - DURAÇÃO RELATIVA DAS PLOSIVAS PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		Pata	Bata	Tata	Data	Cata	Gata
AN	média (d.p.)	10 (0,3)	8 (0,5)	10 (1)	8 (0,4)	9 (1)	7 (0,3)
DA	média (d.p.)	16 (2)	16 (1)	17 (0,7)	16 (1)	16 (2)	15 (2)

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

O teste de Kruskal-Wallis aponta para diferenças significativas entre as durações relativas das consoantes plosivas, considerando os dois sujeitos. Para a comparação das produções de [p] [$\Pi^2_{(1)}=12,637$, $p=0,000$]; de [b] [$\Pi^2_{(1)}=12,142$, $p=0,000$]; de [t] [$\Pi^2_{(1)}=13,512$, $p=0,000$]; de [d] [$\Pi^2_{(1)}=10,705$, $p=0,001$]; de [k] [$\Pi^2_{(1)}=11,593$, $p=0,001$]; de [g] [$\Pi^2_{(1)}=13,216$, $p=0,000$].

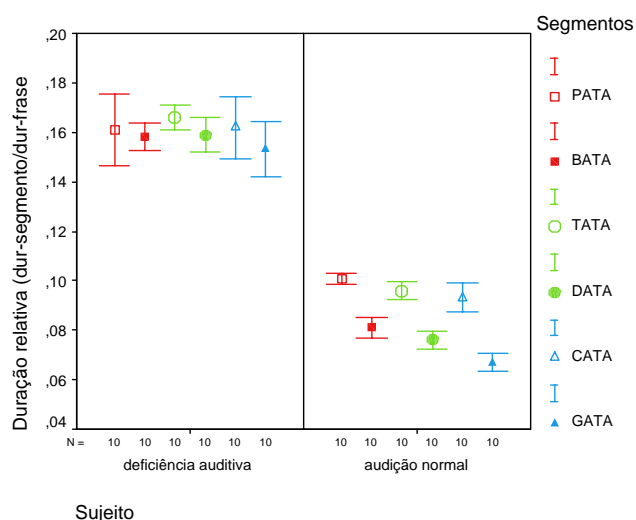


Figura 42 - Gráfico representativo dos valores de duração relativa para as seis plosivas na produção dos dois sujeitos

Para maior clareza quanto ao alongamento das consoantes plosivas na produção do sujeito com deficiência auditiva, os valores de duração relativa dessas e de vários segmentos adjacentes estão bem evidenciados nos gráficos das figuras 43, 43 e 45, nos quais é possível visualizar a maior duração não somente da

consoante, mas também da vogal tônica subsequente e, conseqüentemente, da sílaba tônica da palavra em foco. O aumento da duração, que no PB está relacionado ao padrão de acento (Massini, 1991; Barbosa, 2000) - ou seja, a sílaba tônica é marcada pelo aumento da duração -, parece estar exacerbado na produção do sujeito com deficiência auditiva.

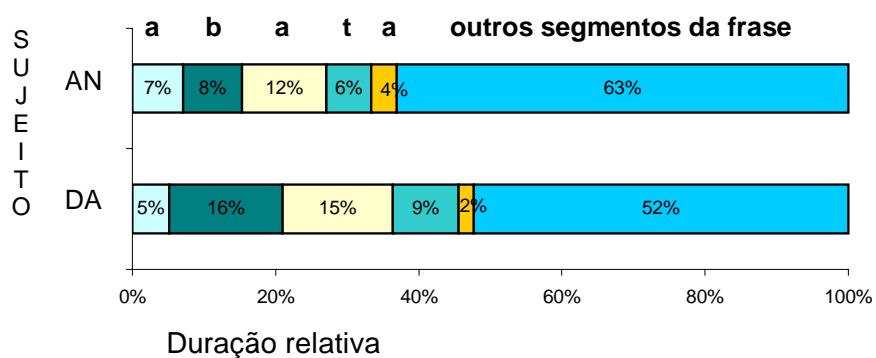
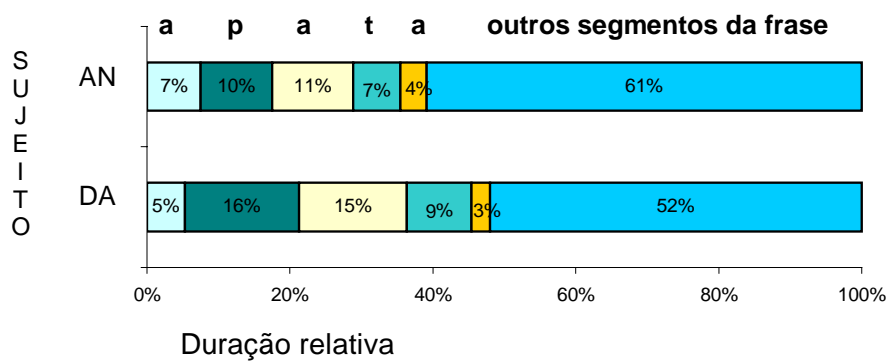
Além do alongamento da consoante plosiva em posição inicial da palavra, podemos observar o alongamento da vogal em posição tônica (tabelas 24 e 25) resultando no aumento da duração da sílaba tônica (tabelas 16 e 17). Ademais, a consoante plosiva [t] na posição pós-tônica da palavra, nas produções do sujeito com deficiência auditiva, encontra-se alongada (tabelas 18 e 19), resultando num alongamento da palavra (tabelas 20 e 21).

Um aspecto que chama a atenção na duração dos segmentos é que o citado alongamento da sílaba tônica da palavra em foco é, de certa forma, “compensado” pelo sujeito com deficiência auditiva, pela evidente redução das vogais em posição átona e pela redução dos segmentos restantes da frase (outros segmentos), representados nos gráficos nas figuras 43, 44 e 45²⁵.

A duração de “outros segmentos” representa a diferença entre a duração total da frase (excluindo a consoante inicial da palavra *diga*) e a duração da palavra acrescida da vogal anterior à palavra.

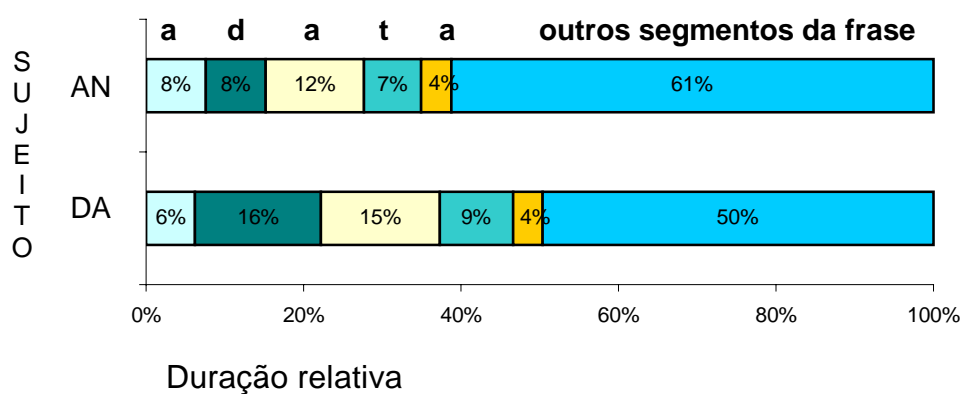
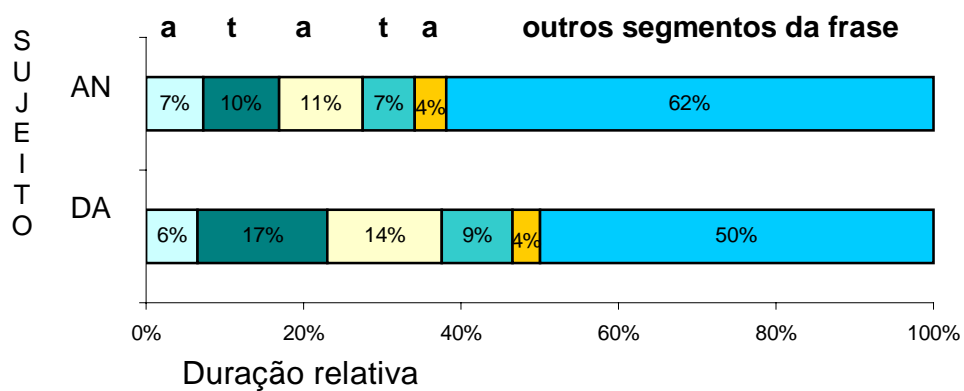
As medidas de duração de cada um desses segmentos citados serão apresentadas a seguir, (tabelas 16, 17, 18, 19, 20, 21, 22, 23, 24 e 25) considerando-se os valores absolutos e relativos. Para as medidas de duração absoluta, foram realizadas comparações das médias das produções para cada um dos sujeitos (intra-sujeito) (*one way* ANOVA, teste *a posteriori* de Scheffe). As medidas de duração relativa foram utilizadas para fazer comparações entre os sujeitos (teste de Kruskal-Wallis).

²⁵ Na figura 24 e 25 no (capítulo 3, pg. 98), que mostra um espectrograma com um exemplo de segmentação da frase para a realização das medidas de duração, este fenômeno também pode ser claramente visualizado. Tais figuras encontram-se no capítulo 3 (p. 98 e 99 respectivamente).



AN – sujeito com audição normal
 DA – sujeito com deficiência auditiva

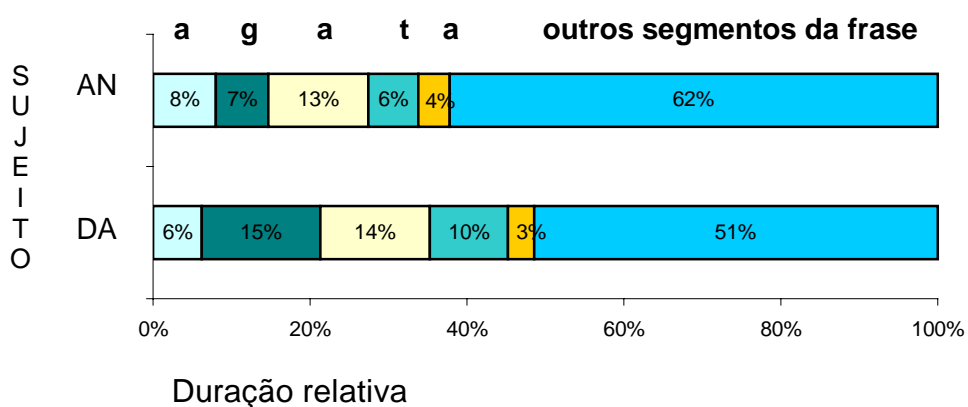
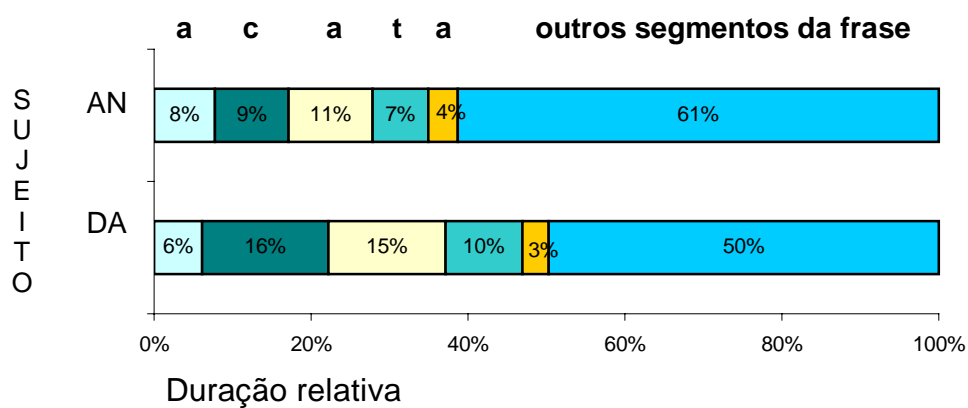
Figura 43 - Gráfico representativo da distribuição dos valores de duração relativa da vogal anterior à plosiva, da plosiva, da vogal subsequente, da plosiva [t] da sílaba átona da palavra, da vogal pós-tônica da palavra e de outros segmentos da frase – bilabiais – para os dois sujeitos



AN – sujeito com audição normal

DA – sujeito com deficiência auditiva

Figura 44 - Gráfico representativo da distribuição dos valores de duração relativa da vogal anterior à plosiva, da plosiva, da vogal subsequente, da plosiva [t] da sílaba átona da palavra, da vogal pós-tônica da palavra e de outros segmentos da frase – alveolares – para os dois sujeitos



AN – sujeito com audição normal

DA – sujeito com deficiência auditiva

Figura 45 -Gráfico representativo da distribuição dos valores de duração relativa da vogal anterior à plosiva, da plosiva, da vogal subsequente, da plosiva [t] da sílaba átona da palavra, da vogal pós-tônica da palavra e de outros segmentos da frase – velares - para os dois sujeitos

A comparação entre as medidas de duração absoluta das sílabas tônicas (tabela 16) nas produções de cada um dos sujeitos não apresentou diferenças significativas para os seis contextos de plosivas para os dois sujeitos.

Tabela 16 - DURAÇÃO ABSOLUTA (ms) DAS SÍLABAS EM POSIÇÃO TÔNICA PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		sílaba [pa]	sílaba [ba]	sílaba [ta]	sílaba [da]	sílaba [ka]	sílaba [ga]
AN	média(d.p.)	253 (12)	241 (15)	253 (13)	241 (13)	246 (12)	238 (10)
	cf.var.	0,05	0,06	0,05	0,05	0,05	0,04
DA	média(d.p.)	382 (28)	385 (16)	403 (22)	410 (25)	404 (33)	380 (30)
	cf.var.	0,07	0,04	0,05	0,06	0,08	0,08

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

A duração relativa das sílabas tônicas na produção do sujeito com deficiência auditiva (tabela 17) encontra-se aumentada se comparada às produções do sujeito com audição normal para todas as consoantes (teste de Kruskal-Wallis: para a comparação da sílaba [pa], $\Pi^2_{(1)}=14,638$; $p=0,000$; para sílaba [ba], $\Pi^2_{(1)}=14,638$; $p=0,000$; para a sílaba [ta], $\Pi^2_{(1)}=14,672$; $p=0,000$; sílaba [da], $\Pi^2_{(1)}=14,649$; $p=0,000$; para a sílaba [ka], $\Pi^2_{(1)}=14,548$; $p=0,000$; para a sílaba [ga] $\Pi^2_{(1)}=14,649$; $p=0,000$).

Tabela 17 - DURAÇÃO RELATIVA DAS SÍLABAS EM POSIÇÃO TÔNICA PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		sílaba [pa]	sílaba [ba]	sílaba [ta]	sílaba [da]	sílaba [ka]	sílaba [ga]
AN	média (d.p.)	21 (1)	20 (2)	20 (1)	20 (1)	20 (1)	19 (0,6)
DA	média (d.p.)	31 (3)	31 (2)	31 (1)	31 (2)	31 (2)	29 (2)

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Para a duração absoluta da consoante [t] na posição pós-tônica (tabela 18), também não foram observadas diferenças significativas entre as medidas de duração da consoante [t] nos seis contextos de consoantes para cada sujeito.

Tabela 18 - DURAÇÃO ABSOLUTA (ms) DAS CONSOANTES [t] DA SÍLABA ÁTONA PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		t (pata)	t (bata)	t (tata)	t (data)	t (cata)	t (gata)
AN	média (d.p.)	78 (7)	76 (15)	82 (4)	87 (7)	87 (7)	79 (15)
	cf.var.	0,09	0,19	0,05	0,08	0,08	0,18
DA	média(d.p.)	115 (11)	113 (13)	117 (8)	124 (16)	128 (15)	129 (15)
	cf.var.	0,10	0,11	0,07	0,13	0,12	0,12

Legenda:
d.p. - desvio padrão
cf.var. - coeficiente de variação
AN - sujeito com audição normal
DA - sujeito com deficiência auditiva

Considerando a duração relativa da consoante [t] na sílaba átona (tabela 19), a análise estatística mostrou que todas as produções do sujeito com deficiência auditiva apresentaram valores maiores de duração se comparadas às produções do sujeito com audição normal (Kruskal-Wallis: [t] na palavra *pata* $\Pi^2_{(1)}=14,961$, $p=0,000$; [t] na palavra *bata* $\Pi^2_{(1)}=14,729$, $p=0,000$; [t] na palavra *tata* $\Pi^2_{(1)}=15,249$, $p=0,000$; [t] na palavra *data* $\Pi^2_{(1)}=7,811$, $p=0,005$; [t] na palavra *cata* $\Pi^2_{(1)}=14,540$, $p=0,000$; [t] na palavra *gata* $\Pi^2_{(1)}=14,113$, $p=0,000$).

Tabela 19 - DURAÇÃO RELATIVA DAS CONSOANTES [t] DA SÍLABA ÁTONA PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		t (pata)	t (bata)	t (tata)	t (data)	t (cata)	t (gata)
AN	média(d.p.)	6 (0,5)	6 (1,2)	7 (0,5)	7 (0,6)	7 (0,50)	6 (0,1)
DA	média(d.p.)	9 (1)	9 (1)	9 (1)	9 (1)	10 (1)	10 (1)

Legenda:
d.p. - desvio padrão
cf.var. - coeficiente de variação
AN - sujeito com audição normal
DA - sujeito com deficiência auditiva

Quanto à duração absoluta das palavras iniciadas pelas seis consoantes estudadas (tabela 20), a análise estatística revela-nos que não houve diferenças significativas entre as produções do sujeito com audição normal. Já para as

produções do sujeito com deficiência auditiva, observamos diferença significativa somente entre as produções [$F_{(5,54)} = 3,558$; $p=0,007$] referentes às palavras *bata*, a mais curta, e à palavra *data*, a mais longa (teste *a posteriori* de Scheffe, $p=0,040$).

Tabela 20 - DURAÇÃO ABSOLUTA (ms) DAS PALAVRAS PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		pata	bata	tata	data	cata	gata
AN	média (d.p.)	375 (16)	360 (25)	385 (19)	374 (18)	380 (17)	366 (22)
	cf.var.	0,04	0,07	0,05	0,05	0,05	0,06
DA	média (d.p.)	541 (34)	524 (30)	566 (33)	584 (42)	576 (49)	552 (34)
	cf.var.	0,06	0,06	0,06	0,07	0,09	0,06

Legenda:
d.p. - desvio padrão
cf.var. - coeficiente de variação
AN - sujeito com audição normal
DA - sujeito com deficiência auditiva

A análise estatística (teste de Kruskal-Wallis) que nos permite comparar as médias de duração relativa das palavras entre os dois sujeitos (tabela 21), aponta novamente para o alongamento da produção do sujeito com deficiência auditiva (para a palavra *pata*, $\Pi^2_{(1)} = 14,752$, $p=0,000$; para *bata* $\Pi^2_{(1)} = 14,449$, $p=0,000$; para *tata* $\Pi^2_{(1)} = 14,438$, $p=0,000$; para *data* $\Pi^2_{(1)} = 14,660$, $p=0,000$; para *cata* $\Pi^2_{(1)} = 14,449$, $p=0,000$; para *gata* $\Pi^2_{(1)} = 4,493$, $p=0,000$). Isso era esperado em virtude do alongamento da sílaba tônica e da consoante [t] na posição pós-tônica.

Tabela 21 - DURAÇÃO RELATIVA DAS PALAVRAS PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito		Pata	Bata	Tata	Data	Cata	Gata
AN	média (d.p.)	32 (0,5)	30 (2)	31 (1,3)	31 (0,5)	31 (1,8)	30 (0,3)
DA	média (d.p.)	43 (3)	43 (2)	44 (2)	44 (2)	44 (3)	42 (2)

Legenda:
d.p. - desvio padrão
cf.var. - coeficiente de variação
AN - sujeito com audição normal
DA - sujeito com deficiência auditiva

Para a medida de duração absoluta de outros segmentos não encontramos diferenças significativas entre as produções de cada um dos sujeitos (tabela 22).

Tabela 22 – DURAÇÃO ABSOLUTA (ms) DE OUTROS SEGMENTOS PRODUZIDOS PELOS DOIS SUJEITOS

Sujeito		outros s. [p]	outros s. [b]	outros s. [t]	outros s. [d]	outros s. [k]	outros s. [g]
AN	média (d.p.)	720 (25)	761 (76)	772 (37)	735 (31)	752 (26)	765 (26)
	cf.var.	0,03	0,10	0,05	0,04	0,03	0,03
DA	média (d.p.)	658 (58)	646 (56)	649 (44)	655 (64)	648 (70)	671 (38)
	cf.var.	0,09	0,09	0,07	0,10	0,11	0,06

Legenda:

outros s. [C] outros segmentos da frase na qual está inserida [C], onde C é uma das plosivas

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Quando, porém, comparamos a duração relativa de outros segmentos entre as produções dos dois sujeitos, verificamos uma inversão no padrão de alongamento dos segmentos até agora descrito. Conforme já foi apontado, em contraposição ao alongamento da sílaba tônica e da consoante [t] na posição pós-tônica, ocorre um encurtamento dos outros segmentos que compõem a frase (tabela 23). A análise estatística mostrou diferenças significativas em todos os contextos de consoante estudados (Kruskal-Wallis: outros segmentos para a palavra *pata*, $\Pi^2_{(1)}=14,571$, $p=0,000$; para a palavra *bata*, $\Pi^2_{(1)}=14,504$, $p=0,000$; para a palavra *tata*, $\Pi^2_{(1)}=14,559$, $p=0,000$; para a palavra *data*, $\Pi^2_{(1)}=14,548$, $p=0,000$; para a palavra *cata*, $\Pi^2_{(1)}=14,449$, $p=0,000$; para a palavra *gata*, $\Pi^2_{(1)}=14,449$; $p=0,000$).

Tabela 23 - DURAÇÃO RELATIVA DE OUTROS SEGMENTOS PRODUZIDOS PELO DOIS SUJEITOS

Sujeito		outros s. [p]	outros s. [b]	outros s. [t]	outros s. [d]	outros s. [k]	outros s. [g]
AN	média (d.p.)	61 (3)	63 (1)	62 (1)	61 (1)	61 (0,3)	62 (0,1)
DA	média (d.p.)	52 (1)	52 (1)	50 (2)	50 (1)	50 (1)	51 (2)

Legenda:

outros s. [C] outros segmentos da frase na qual está inserida [C], onde C é uma das plosivas

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Considerando-se a duração absoluta das vogais adjacentes na produção de cada um dos sujeitos, verifica-se que existem diferenças significativas entre as durações das vogais nas três posições estudadas (pós-tônica anterior à plosiva, tônica e pós-tônica). Todas as vogais tônicas, para os dois sujeitos, são mais longas do que as pós-tônicas da palavra anterior e do que as pós-tônicas da palavra em foco. A análise estatística para as produções do sujeito com audição normal mostrou que: para as vogais adjacentes a [p] [$F_{(2,27)} = 396,364$, $p=0,000$]; para as adjacentes a [b] [$F_{(2,27)} = 396,400$, $p=0,000$]; para as adjacentes a [t] [$F_{(2,27)} = 351,952$, $p=0,000$]; para as adjacentes a [d] [$F_{(2,27)} = 560,271$, $p=0,000$]; para as adjacentes a [k] [$F_{(2,27)} = 586,635$, $p=0,000$] para as adjacentes a [g] [$F_{(2,27)} = 986,793$, $p=0,000$]. Já a análise estatística para as produções do sujeito com deficiência auditiva revelou que: para as vogais adjacentes a [p] [$F_{(2,27)} = 578,018$, $p=0,000$]; para as adjacentes a [b] [$F_{(2,27)} = 1247,131$, $p=0,000$]; para as adjacentes a [t] [$F_{(2,27)} = 358,818$, $p=0,000$]; para as adjacentes a [d] [$F_{(2,27)} = 265,268$, $p=0,000$]; para as adjacentes a [k] [$F_{(2,27)} = 338,286$, $p=0,000$]; para as adjacentes a [g] [$F_{(2,27)} = 425,015$, $p=0,000$]. Os dados referentes às medidas de duração das vogais adjacentes às plosivas figuram na tabela 24.

Se fizermos a comparação entre as medidas de duração da vogal tônica produzida pelo sujeito com deficiência auditiva nos diferentes contextos de consoantes, não encontraremos diferenças entre elas, ou seja, todas as vogais produzidas nessa posição tiveram a mesma duração. Já para o sujeito com audição normal, ocorreram diferenças entre essas medidas de duração nos diferentes contextos de plosivas [$F_{(5,54)} = 11,860$, $p=0,000$]. Verificamos que a duração das vogais que sucedem as plosivas não-vozeadas é menor que a das que sucedem as plosivas vozeadas. Comparando as médias da duração relativa das vogais, foram observadas diferenças significativas para as velares ($p=0,000$) e alveolares ($p=0,008$), mas não para as bilabiais. A diferença em ms foi de 25 ms para as velares, 17 ms para as alveolares e de 7 ms para as bilabiais. Isso explica o fato de termos encontrado diferenças significativas entre as medidas de duração das consoantes vozeadas e não-vozeadas e não entre as sílabas, para o sujeito com audição normal.

Tabela 24 - DURAÇÃO ABSOLUTA (ms) DA VOGAL ANTERIOR À PLOSIVA, DA VOGAL TÔNICA E DA VOGAL PÓS-TÔNICA DA PALAVRA EM FOCO, NOS SEIS CONTEXTOS DE PLOSIVA ESTUDADOS, PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito	vogal anterior	digA pa	digA ba	digA ta	digAda	digA ca	digA ga
AN	média (d.p.)	89 (5)	85 (7)	90 (5)	92 (5)	95 (3)	98 (6)
	cf.var.	0,06	0,08	0,06	0,05	0,04	0,06
DA	média (d.p.)	67 (10)	63 (4)	84 (10)	81 (11)	79 (13)	80 (12)
	cf.var.	0,16	0,06	0,12	0,14	0,16	0,15
Sujeito	vogal tônica	pA	bA	tA	dA	cA	gA
AN	média (d.p.)	135 (11)	142 (11)	133 (9)	150 (10)	131 (8)	156 (7)
	cf.var.	0,08	0,08	0,07	0,06	0,06	0,04
DA	média (d.p.)	191 (14)	190 (11)	188 (14)	199 (17)	194 (17)	182 (14)
	cf.var.	0,07	0,06	0,07	0,09	0,09	0,08
Sujeito	vogal pós-tônica	patA	batA	tatA	datA	catA	gatA
A N	média (d.p.)	44 (3)	43 (4)	50 (5)	47 (5)	46 (4)	49 (4)
	cf.var.	0,07	0,10	0,11	0,11	0,08	0,07
DA	média (d.p.)	33 (8)	27 (7)	46 (12)	50 (17)	44 (9)	43 (6)
	cf.var.	0,23	0,25	0,27	0,33	0,21	0,13

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Embora o padrão de variação da duração entre as vogais siga o esperado para o PB, observamos diferença muito mais acentuada para o sujeito com deficiência auditiva, em relação às produções do sujeito com audição normal, no que se refere ao alongamento da vogal tônica. Na tabela 25, há as médias das durações relativas da vogal adjacente nas três posições estudadas.

O alongamento da vogal tônica e o encurtamento da vogal pós-tônica exacerbados nas produções do sujeito com deficiência auditiva foram confirmados na análise estatística considerando a comparação das medidas de duração relativa entre os dois sujeitos.

Para as vogais tônicas, verificamos que todas as medidas de duração apresentaram-se mais longas nas produções do sujeito com deficiência auditiva (para [a] na sílaba [pa] $\Pi^2_{(1)}=14,132$, $p=0,000$; para [a] na sílaba [ba] $\Pi^2_{(1)}=15,372$, $p=0,000$; para [a] na sílaba [ta] $\Pi^2_{(1)}=14,972$, $p=0,000$; para [a] na sílaba [da] $\Pi^2_{(1)}=14,902$, $p=0,000$; para [a] na sílaba [ka] $\Pi^2_{(1)}=15,091$, $p=0,000$; para [a] na sílaba

[ga] $\Pi^2_{(1)} = 6,400$, $p = 0,011$).

Já para as vogais anteriores às plosivas, foi possível observar que todas as produzidas pelo sujeito com deficiência auditiva apresentaram valores de duração menores se comparados às produções do sujeito com audição normal (para [i] anterior ao [p] $\Pi^2_{(1)} = 15,139$, $p = 0,000$; para [i] anterior ao [b] $\Pi^2_{(1)} = 14,370$, $p = 0,000$; para [i] anterior ao [t] $\Pi^2_{(1)} = 6,891$, $p = 0,009$; para [i] anterior ao [d] $\Pi^2_{(1)} = 11,982$, $p = 0,001$; para [i] anterior ao [k] $\Pi^2_{(1)} = 8,981$, $p = 0,003$; para [i] anterior ao [g] $\Pi^2_{(1)} = 12,313$, $p = 0,000$).

Para as vogais em posição pós-tônica, os valores de duração foram menores nas produções do sujeito com deficiência auditiva somente nos contextos das bilabiais e velares: para as pós-tônicas da palavra *pata* $\Pi^2_{(1)} = 10,528$, $p = 0,001$; para as pós-tônicas da palavra *bata* $\Pi^2_{(1)} = 15,150$, $p = 0,000$; para as pós-tônicas da palavra *cata* $\Pi^2_{(1)} = 5,007$, $p = 0,025$; para as pós-tônicas da palavra *gata* $\Pi^2_{(1)} = 10,231$, $p = 0,001$. Para as vogais pós-tônicas das palavras iniciadas pelas alveolares, não ocorreram diferenças significativas.

Tabela 25 - DURAÇÃO RELATIVA DA VOGAL ANTERIOR À PLOSIVA, DA VOGAL TÔNICA E DA VOGAL PÓS-TÔNICA DA PALAVRA EM FOCO, NOS SEIS CONTEXTOS DE PLOSIVA ESTUDADOS, PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito	vogal anterior	digA pa	digA ba	digA ta	digAda	digA ca	digA ga
AN	média (d.p.)	7 (0,5)	7 (0,9)	7 (0,5)	8 (0,5)	8 (0,4)	8 (0,5)
DA	média (d.p.)	5 (0,9)	5 (0,2)	6 (0,9)	6 (0,6)	6 (1,3)	6 (1)
	vogal tônica	pA	bA	tA	dA	cA	gA
AN	média (d.p.)	11 (1)	12 (1)	11 (1)	12 (1)	11 (1)	13 (1)
DA	média (d.p.)	15 (1)	15 (1)	14 (1)	15 (1)	15 (1)	14 (1)
	vogal pós-tônica	patA	batA	tatA	datA	catA	gatA
AN	média (d.p.)	4 (0,3)	4 (0,5)	4 (0,5)	4 (0,4)	4 (0,3)	4 (0,3)
DA	média (d.p.)	3 (0,6)	2 (0,5)	4 (0,9)	4 (1)	3 (0,6)	3 (0,4)

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

As diferenças entre as durações das vogais adjacentes podem ser visualizadas no gráfico da figura 46, a seguir:

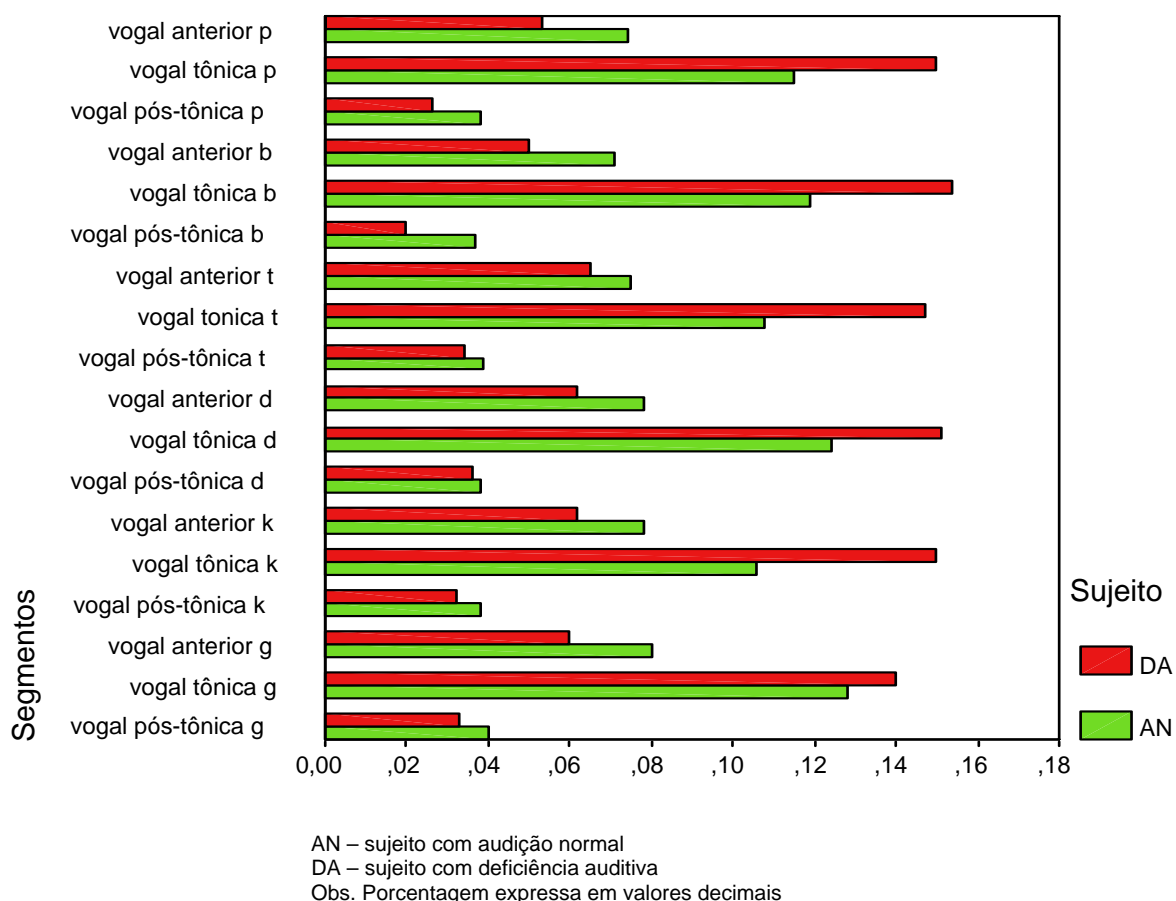


Figura 46 - Gráfico representativo da distribuição dos valores médios de duração relativa da vogal [a] nas três posições estudadas, nos seis contextos, para os dois sujeitos

Calculamos ainda as diferenças entre as medidas de duração da vogal na posição tônica e na posição pós-tônica da palavra em foco, as quais são apresentadas na tabela 26, uma outra forma de comparar as produções dos dois sujeitos já que passaremos a considerar a relação entre os dois segmentos:

Tabela 26 - DIFERENÇA ENTRE A DURAÇÃO (ms) DA VOGAL TÔNICA E DA PÓS-TÔNICA DA PALAVRA NOS DIFERENTES CONTEXTOS DE CONSOANTE PARA OS DOIS SUJEITOS

Sujeito		tôn – pós [p]	tôn – pós [b]	tôn – pós [t]	tôn – pós [d]	tôn – pós [k]	tôn – pós [g]
AN	média (d.p.)	91 (11)	99 (9)	83 (8)	103 (10)	84 (9)	107 (7)
	cf.var.	0,12	0,09	0,10	0,10	0,11	0,07
DA	média (d.p.)	158 (15)	163 (12)	142 (13)	149 (13)	150 (20)	139 (15)
	cf.var.	0,09	0,07	0,09	0,09	0,14	0,11

Legenda:

tôn – pós - diferença entre a duração da vogal tônica e da pós-tônica da palavra iniciada por [C] onde C é uma das plosivas

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

A comparação entre os sujeitos mostra que as diferenças entre as medidas de duração da vogal /a/ na posição tônica e pós-tônica são maiores para o sujeito com deficiência auditiva do que para o sujeito com audição normal confirmando o alongamento da tônica e, por outro lado, a brevidade da pós-tônica (Kruskal-Wallis: no contexto [p] $\Pi^2_{(1)}=14,286$, $p=0,000$; no contexto [b] $\Pi^2_{(1)}=14,329$, $p=0,000$; no contexto [t] $\Pi^2_{(1)}=14,296$, $p=0,000$; no contexto [d] $\Pi^2_{(1)}=14,307$, $p=0,000$; no contexto [k] $\Pi^2_{(1)}=14,307$, $p=0,000$; no contexto [g] $\Pi^2_{(1)}=14,307$, $p=0,000$).

4.2.5 Medidas de energia da consoante em relação à vogal subsequente

A medida de energia (*root mean square* – rms em dB) da consoante e das vogais adjacentes foi realizada para todos os contextos de plosivas e para os dois sujeitos. Para o sujeito com audição normal, observamos uma diferenciação no padrão da energia relativa entre as consoantes vozeadas e não-vozeadas, ou seja, uma diferença menor entre a energia da plosiva vozeada e da vogal subsequente para os três pontos de articulação [$F_{(5,54)}=68,805$, $p=0,000$]. Por meio do teste *a posteriori* de Scheffe, vimos diferenças significativas entre as sílabas [pa] e [ba] ($p=0,000$), entre [ta] e [da] ($p=0,000$) e entre [ka] e [ga] ($p=0,000$). Como a medida incluiu todo o período de oclusão da consoante, tal diferença provavelmente está refletindo a presença da barra de vozeamento nas plosivas vozeadas produzidas pelo sujeito com audição normal. Essa diferença aumenta em 1dB entre a vozeada

bilabial (-11 dB) e a alveolar (-12 dB) e em 2 dB entre a alveolar e a velar (-13 dB). Já para as plosivas não-vozeadas, o contrário pôde ser observado: a maior diferença foi encontrada para a bilabial (-26 dB), com a diminuição da diferença em 1 dB para a alveolar (-25 dB) e dessa para a velar (-23 dB) com a diminuição em 2 dB. A análise estatística não aponta para diferenças significativas entre as medidas de energia considerando o ponto de articulação. Entretanto, deve-se observar que diferenças de mais de 1dB são identificadas pelo ouvido humano. A Tabela 27 apresenta a relação C/V entre as plosivas e a vogal subsequente para os dois sujeitos.

Tabela 27 - MEDIDA DA ENERGIA DA CONSOANTE PLOSIVA EM RELAÇÃO À VOGAL SUBSEQÜENTE (rms em dB) PARA OS DOIS SUJEITOS

Sujeito		P- A	B- A	T- A	D- A	C- A	G- A
AN	média (d.p.)	-26 (2)	-11 (3)	-25 (2)	-12 (1)	-23 (4)	-13 (2)
	cf.var.	-0,08	-0,28	-0,09	-0,11	-0,19	-0,12
DA	média (d.p.)	-27 (3)	-27 (3)	-28 (5)	-28 (3)	-26 (3)	-26 (3)
	cf.var.	-0,11	-0,10	-0,16	-0,10	-0,13	-0,12

Legenda:

C-A – diferença de energia entre a consoante e a vogal subsequente onde C é uma das plosivas

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

Para o sujeito com deficiência auditiva, também nesse parâmetro não foram encontradas diferenças estatisticamente significativas entre os pares de plosivas vozeadas e não-vozeadas e nem diferenças em dB. Quanto ao ponto de articulação, ocorreu o mesmo verificado para o sujeito com audição normal, ou seja, as diferenças verificadas em dB, não foram estatisticamente significativas. No entanto, seguem um padrão diverso do encontrado no sujeito com audição normal: menor diferença para o ponto velar (-26 dB), seguido pelo bilabial (-27 dB) e pelo alveolar (-28 dB).

Cabe-nos observar que, para os dois sujeitos, não foram encontradas diferenças significativas entre as medidas de energia da vogal tônica produzidas por cada um nas dez repetições, nos seis contextos de consoante, fato que indica que a variação da energia relativa decorre da variação na consoante. No gráfico da figura 47, fica evidente a diferenciação no padrão da medida de energia relativa entre as plosivas vozeadas e não-vozeadas, no caso do sujeito com audição normal e a não

diferenciação no caso do sujeito com deficiência auditiva:

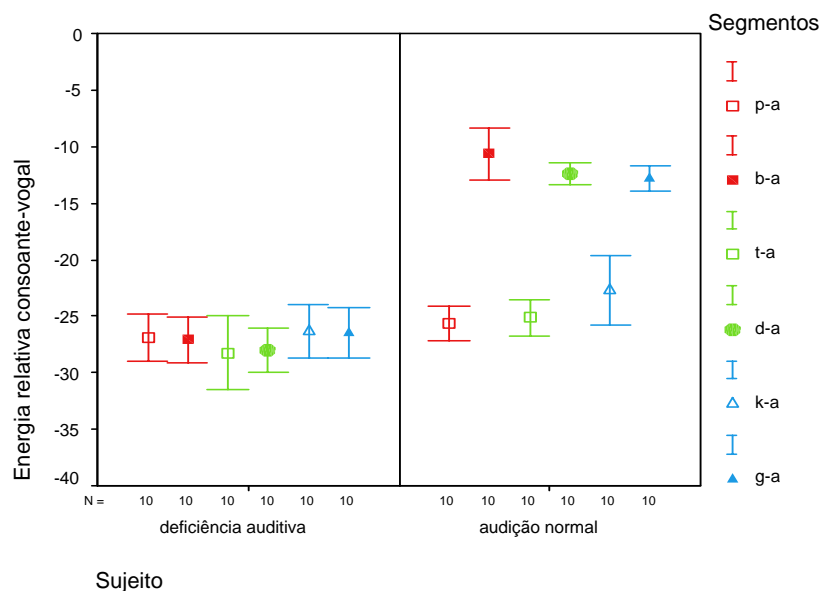


Figura 47- Gráfico representativo da medida de energia relativa da consoante nos seis contextos de plosiva, para os dois sujeitos

4.2.6 Frequência dos formantes das vogais adjacentes

Para obter mais dados sobre aspectos da coarticulação na fala dos dois sujeitos estudados, verificamos a frequência dos primeiro, segundo e terceiro formantes (F1, F2 e F3) das vogais adjacentes às plosivas. Os resultados figuram na tabela 28, a qual inclui médias, desvio padrão e coeficiente de variação para as três posições da vogal: vogal anterior à consoante; vogal tônica e pós-tônica, nos seis contextos de plosivas, para ambos os sujeitos.

Tabela 28 - FREQUÊNCIA DOS FORMANTES DAS VOGAIS ADJACENTES ÀS PLOSIVAS PRODUZIDAS PELOS DOIS SUJEITOS

Sujeito	PATA	Ap F1	Ap F2	Ap F3	pA F1	pA F2	pA F3	PatA F1	PatA F2	PatA F3
AN	média	576	1560	2099	802	1382	2320	545	1348	2574
	(d.p.)	(19)	(58)	(49)	(62)	(51)	(74)	(21)	(60)	(52)
	cf.var.	0,03	0,04	0,02	0,08	0,04	0,03	0,04	0,04	0,02
DA	média	650	1550	2774	1061	1627	2619	497	1551	2867
	(d.p.)	(80)	(78)	(81)	(38)	(84)	(59)	(21)	(90)	(94)
	cf.var.	0,12	0,05	0,03	0,04	0,05	0,02	0,04	0,06	0,03
Sujeito	BATA	Ab F1	Ab F2	Ab F3	bA F1	bA F2	bA F3	batAF1	batAF2	batAF3
AN	média	591	1531	2126	794	1395	2326	537	1395	2599
	(d.p.)	(44)	(74)	(50)	(60)	(42)	(87)	(19)	(76)	(51)
	cf.var.	0,08	0,05	0,02	0,08	0,03	0,04	0,04	0,05	0,02
DA	média	675	1473	2759	1060	1619	2625	488	1573	2858
	(d.p.)	(37)	(59)	(78)	(37)	(67)	(78)	(21)	(68)	(80)
	cf.var.	0,05	0,04	0,03	0,03	0,04	0,03	0,04	0,04	0,03
Sujeito	TATA	At F1	At F2	At F3	tA F1	tA F2	tA F3	tatAF1	tatAF2	tatAF3
AN	média	581	1580	2870	735	1367	2748	540	1322	2582
	(d.p.)	(23)	(97)	(108)	(15)	(64)	(52)	(14)	(22)	(59)
	cf.var.	0,04	0,06	0,04	0,02	0,05	0,02	0,03	0,02	0,02
DA	média	681	1708	2830	1052	1608	2746	503	1522	2768
	(d.p.)	(49)	(54)	(77)	(33)	(32)	(124)	(40)	(89)	(103)
	cf.var.	0,07	0,03	0,03	0,03	0,02	0,05	0,08	0,06	0,04
Sujeito	DATA	Ad F1	Ad F2	Ad F3	dA F1	dA F2	dA F3	datAF1	datAF2	datAF3
AN	média	562	1534	2815	747	1381	2781	545	1359	2600
	(d.p.)	(24)	(52)	(55)	(67)	(64)	(89)	(17)	(61)	(52)
	cf.var.	0,04	0,03	0,02	0,09	0,05	0,03	0,03	0,05	0,02
DA	média	684	1697	2836	1024	1638	2636	506	1593	2818
	(d.p.)	(19)	(34)	(65)	(47)	(36)	(85)	(47)	(60)	(128)
	cf.var.	0,03	0,02	0,02	0,05	0,02	0,03	0,09	0,04	0,05
Sujeito	CATA	Ac F1	Ac F2	Ac F3	cA F1	cA F2	cA F3	catAF1	catAF2	catAF3
AN	média	577	1669	2166	744	1488	2213	541	1339	2518
	(d.p.)	(19)	(65)	(45)	(23)	(20)	(79)	(18)	(71)	(60)
	cf.var.	0,03	0,04	0,02	0,03	0,01	0,04	0,03	0,05	0,02
DA	média	693	1651	2762	1049	1661	2707	539	1602	2830
	(d.p.)	(31)	(70)	(83)	(30)	(52)	(83)	(66)	(69)	(108)
	cf.var.	0,04	0,04	0,03	0,03	0,03	0,03	0,12	0,04	0,04
Sujeito	GATA	Ag F1	Ag F2	Ag F3	gA F1	gA F2	gA F3	gatAF1	gatAF2	gatAF3
AN	média	576	1664	2103	738	1494	2229	539	1348	2546
	(d.p.)	(23)	(50)	(57)	(22)	(58)	(63)	(19)	(75)	(45)
	cf.var.	0,04	0,03	0,03	0,03	0,04	0,03	0,04	0,06	0,02
DA	média	658	1687	2714	1069	1682	2723	555	1582	2774
	(d.p.)	(63)	(63)	(81)	(45)	(83)	(79)	(64)	(86)	(120)
	cf.var.	0,10	0,04	0,03	0,04	0,05	0,03	0,11	0,05	0,04

Legenda

d.p. - desvio padrão

cf.var. - coeficiente de variação

AN - sujeito com audição normal

DA - sujeito com deficiência auditiva

[a] C - vogal pós tônica anterior à plosiva (onde C é uma das consoantes plosiva)

C [a] - vogal tônica

Cat[a] - vogal pós-tônica

Os gráficos, a seguir (figuras 48, 49 e 50), ilustram a variação nas freqüências dos três primeiros formantes da vogal /a/ nas três posições adjacentes às seis plosivas para os dois sujeitos. Os gráficos da figura 48 mostram as freqüências dos formantes da vogal na posição anterior à plosiva. Nos da figura 49 estão dispostos os valores das freqüências dos três primeiros formantes da vogal na posição tônica e, nos gráficos da figura 50, os da vogal na posição pós-tônica.

Em primeiro lugar, descrevemos os resultados para as medidas dos formantes da vogal anterior à plosiva. Com relação ao primeiro formante, F1, não foram encontradas diferenças significativas entre os seis contextos de consoantes plosivas para os dois sujeitos.

Quanto ao segundo formante, F2, verificamos nas produções do sujeito com deficiência auditiva, diferenças significativas [$F_{(5,54)}=24,015$, $p=0,000$], sendo essa diferença entre as bilabiais (com valores mais baixos) e os demais pontos de articulação (entre [p] e [t,d], $p=0,000$; entre [p] e [k] $p=0,028$; entre [p] e [g] $p=0,001$; entre [b] e [t,d,k,g] $p=0,000$). Já para o sujeito com audição normal os valores foram mais baixos para as bilabiais e alveolares, sendo os mais altos para as velares [$F_{(5,54)}=8,407$, $p=0,000$]. Nesse caso as diferenças foram significativas entre [p] e [k] ($p=0,037$); [b] e [k] ($p=0,003$); [b] e [g] ($p=0,005$); [d] e [k] ($p=0,004$); [d] e [g] ($p=0,006$).

Para o terceiro formante, F3, nas produções do sujeito com deficiência auditiva, houve diferenças significativas [$F_{(5,54)}=3,569$, $p=0,007$] mas apenas entre [g], o valor mais baixo e [d], o valor mais alto ($p=0,044$). Nas produções do sujeito com audição normal, os valores das alveolares também foram os mais altos [$F_{(5,54)}=336,030$, $p=0,000$], apresentando diferenças significativas apenas entre essas e as bilabiais ($p=0,000$) e velares ($p=0,000$), sendo que as últimas apresentaram valores iguais.

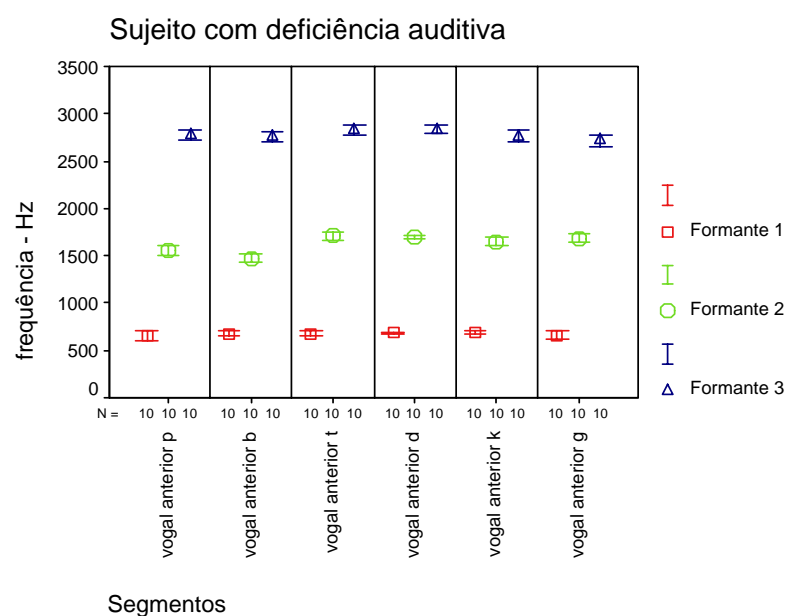
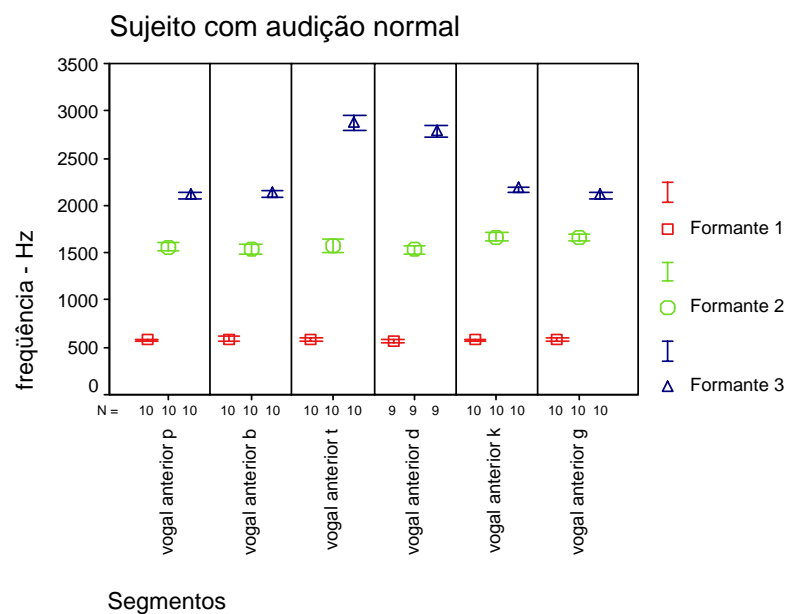


Figura 48 - Gráficos representativos da distribuição dos valores da frequência (Hz) dos Formantes 1, 2 e 3 medidos na porção estacionária da vogal [i] na posição anterior à plosiva para os dois sujeitos

Para as vogais tônicas, não foram encontradas diferenças significativas nos

valores de F1 entre os seis contextos de consoantes para os dois sujeitos. As medidas de F2 das produções do sujeito com deficiência auditiva não apresentaram diferenças significativas na posição tônica. Já para o sujeito com audição normal, a análise estatística mostrou que ocorreram diferenças significativas [$F_{(5,54)} = 12,080$, $p=0,000$] entre as velares e as outras plosivas ([k] e [p] $p=0,003$; [k] e [b] $p=0,013$; [k] e [t] $p=0,003$; [k] e [d] $p=0,003$; [g] e [p] $p=0,001$; [g] e [b] $p=0,007$; [g] e [t] $p=0,000$; [g] e [d] $p=0,001$), sendo que, no contexto das velares, as freqüências de F2 foram as mais altas.

O terceiro formante da vogal [a] na posição tônica apresentou-se da seguinte forma: para o sujeito com deficiência auditiva, não ocorreram diferenças significativas entre os seis contextos de consoantes; para o sujeito com audição normal, os valores de F3 apresentaram diferenças significativas [$F_{(5,54)} = 118,895$, $p=0,000$]. Os valores de freqüência para o F3 da vogal [a] foram mais altos quando ela era subsequente à consoante plosiva alveolar - ([t] > [p, b, k, g] $p=0,000$; [d] > [p, b, k, g] $p=0,000$). A distribuição dos valores das freqüências dos formantes das vogais tônicas nos seis contextos de consoantes para os dois sujeitos está ilustrada no gráfico da figura 49.

Para a vogal [ɐ] em posição pós-tônica, encontramos, tanto nas produções do sujeito com deficiência auditiva quanto nas do sujeito com audição normal, valores de F1 iguais. Para F2, também não foram encontradas diferenças significativas o que ocorreu ainda para F3. Trata-se de um resultado esperado porque o contexto fonético imediato, nesse caso, era idêntico (figura 50).

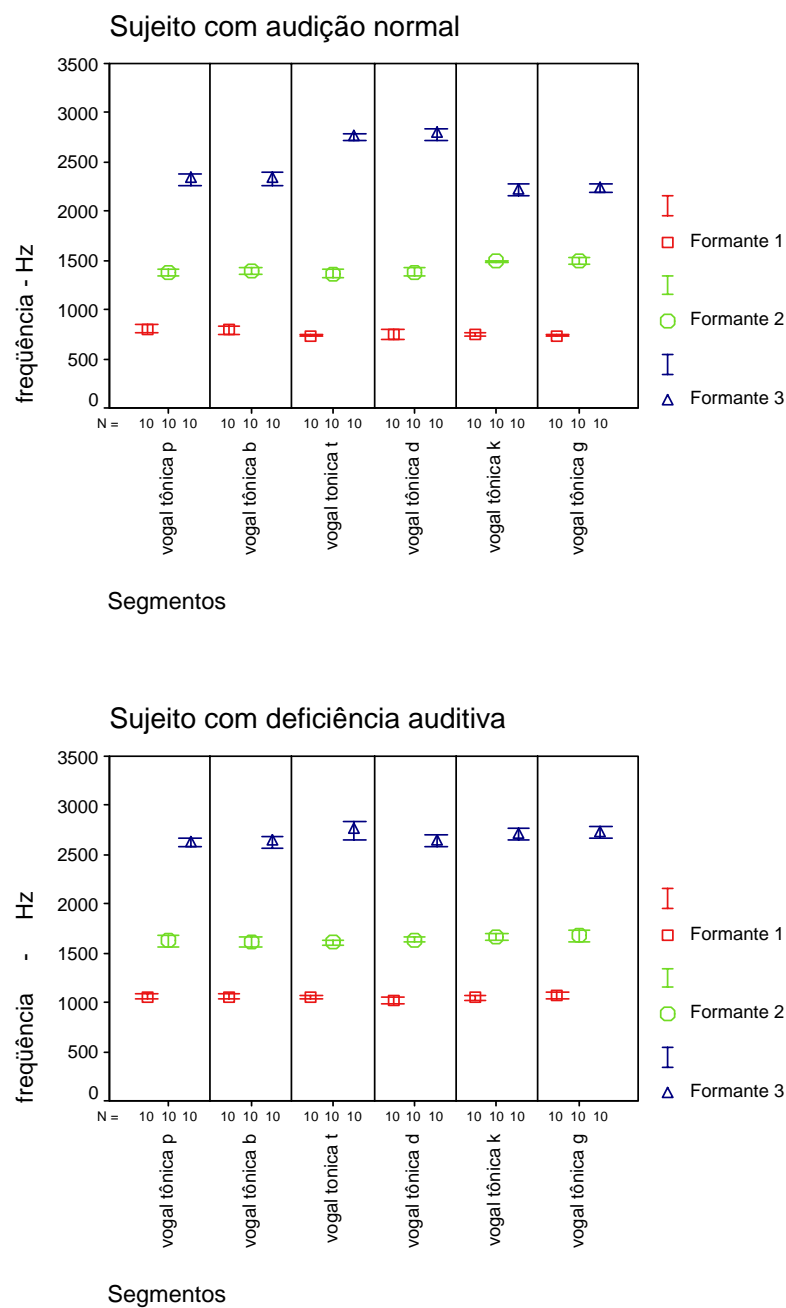


Figura 49 - Gráficos representativos da distribuição dos valores da frequência (Hz) dos Formantes 1, 2 e 3 medidos na porção estacionária da vogal [a] na posição tônica, subsequente à plosiva para os dois sujeitos

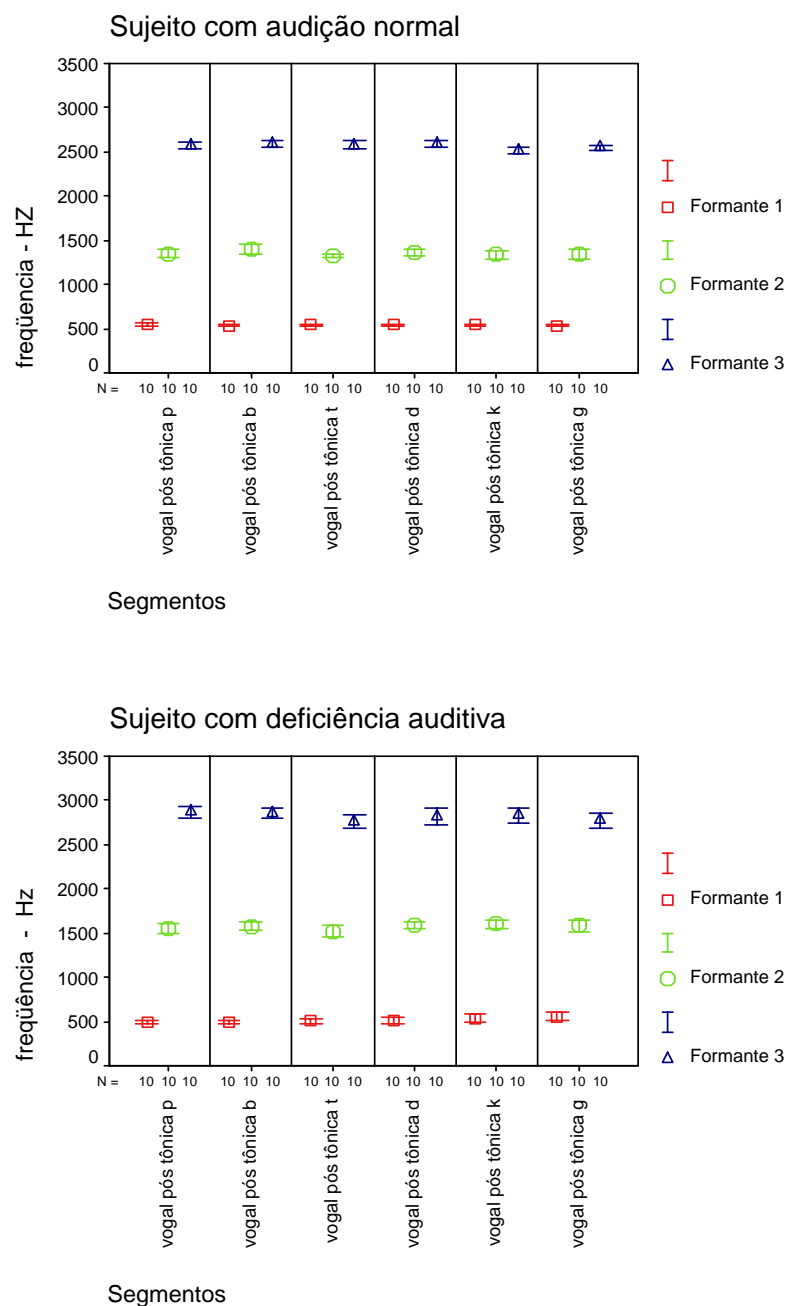


Figura 50 - Gráfico representativo da distribuição dos valores da frequência (Hz) do Formante 1, 2 e 3 medidos na porção estacionária da vogal [ɪ] na posição pós-tônica da palavra para os dois sujeitos

Se compararmos os valores dos formantes nas três posições, para cada sujeito, em cada contexto verificaremos que o primeiro formante eleva-se na posição tônica

para todos os contextos e para os dois sujeitos, o que era esperado visto que F1 está relacionado à abertura da mandíbula (abaixamento da língua). Para o sujeito com deficiência auditiva, encontramos na análise estatística, para as comparações das medidas de F1 nas três posições de vogais, no contexto de [p], os valores [$F_{(2,27)} = 305,262$, $p = 0,000$]; no contexto de [b], [$F_{(2,27)} = 806,881$, $p = 0,000$], no contexto de [t], [$F_{(2,27)} = 467,893$, $p = 0,000$]; no contexto de [d], [$F_{(2,27)} = 439,404$, $p = 0,000$]; no contexto de [k], [$F_{(2,27)} = 332,000$, $p = 0,000$]; para o contexto de [g], [$F_{(2,27)} = 220,466$, $p = 0,000$]. Para o sujeito com audição normal, encontramos, no contexto de [p], os valores [$F_{(2,27)} = 128,282$; $p = 0,000$]; no contexto de [b], [$F_{(2,27)} = 92,791$, $p = 0,000$]; no contexto de [t], [$F_{(2,27)} = 325,293$, $p = 0,000$]; no contexto de [d], [$F_{(2,27)} = 70,297$, $p = 0,000$]; no contexto de [k], [$F_{(2,27)} = 288,083$, $p = 0,000$]; no contexto de [g], [$F_{(2,27)} = 243,557$, $p = 0,000$].

Nos gráficos a seguir, representamos a relação entre F1 e F2 da vogal /a/, nas três posições estudadas, para o contexto das bilabiais (figura 51), alveolares (figura 52) e das velares (figura 53). Neles, evidencia-se a diferença mais marcante entre o primeiro formante na posição tônica e nas pós-tônicas nas produções do sujeito com deficiência auditiva.

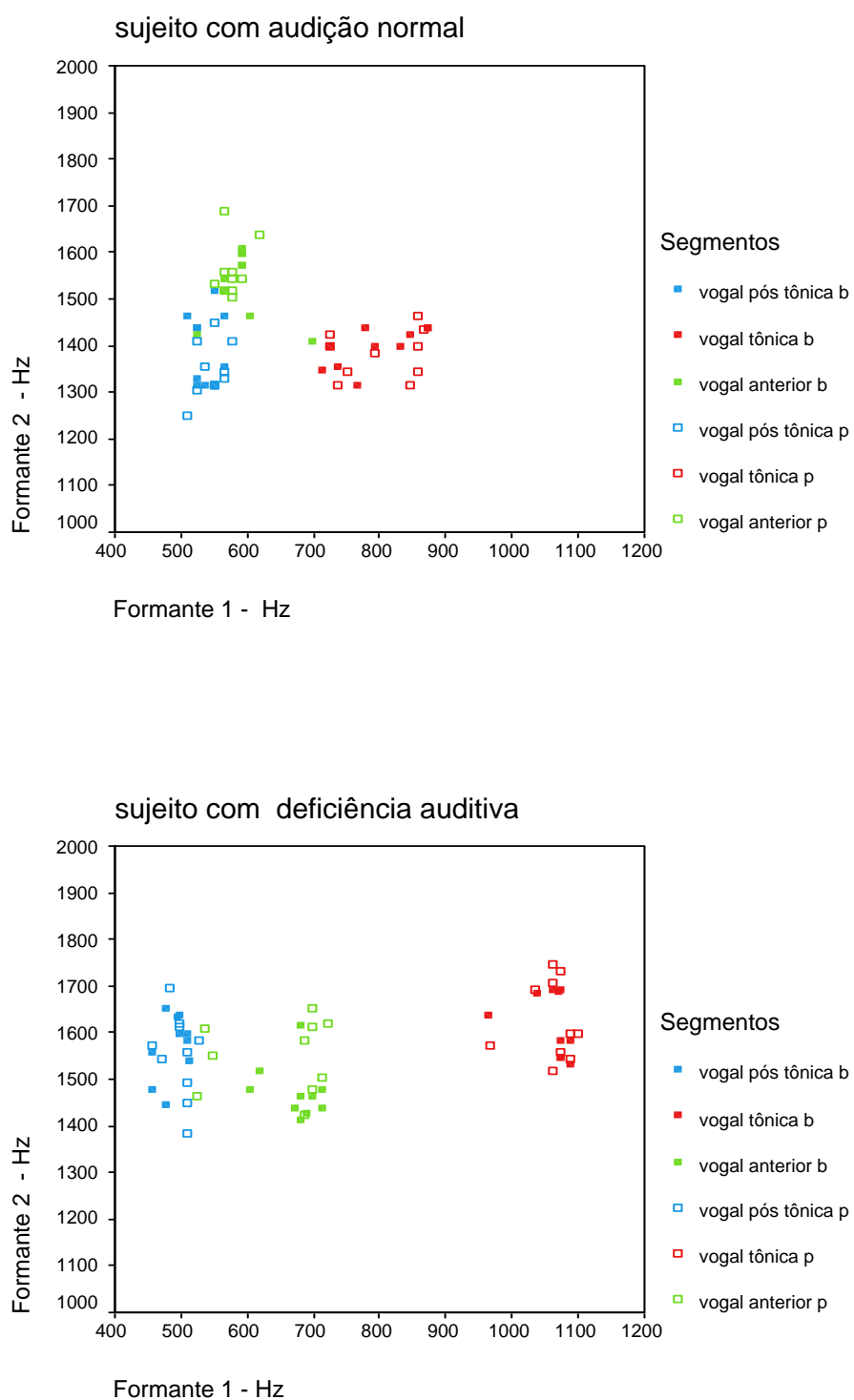


Figura 51 - Gráficos demonstrativos da relação entre as frequências de F1 e F2 medidas na porção estacionária da vogal pós-tônica anterior, tônica e pós-tônica da palavra no contexto bilabial, para os dois sujeitos

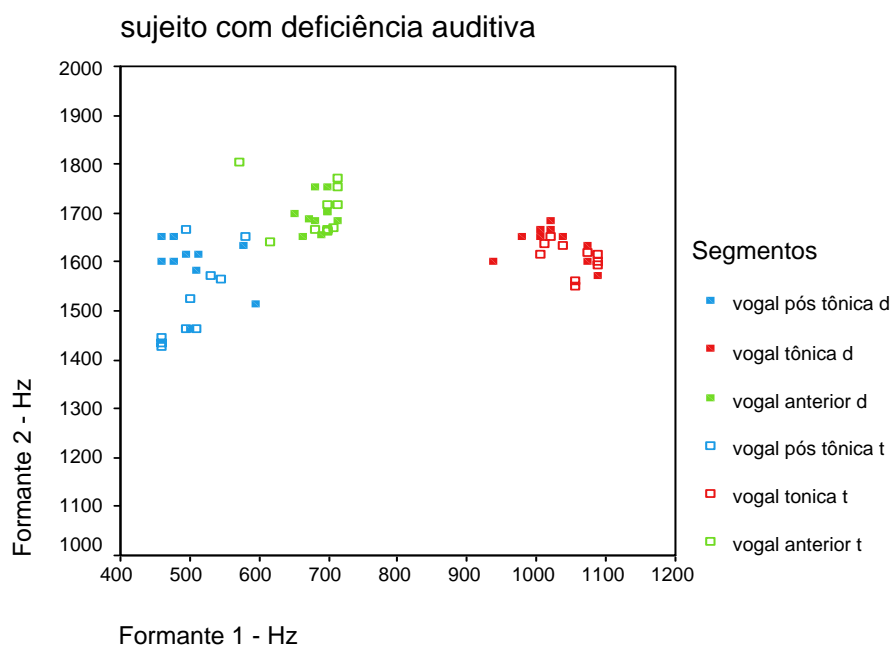
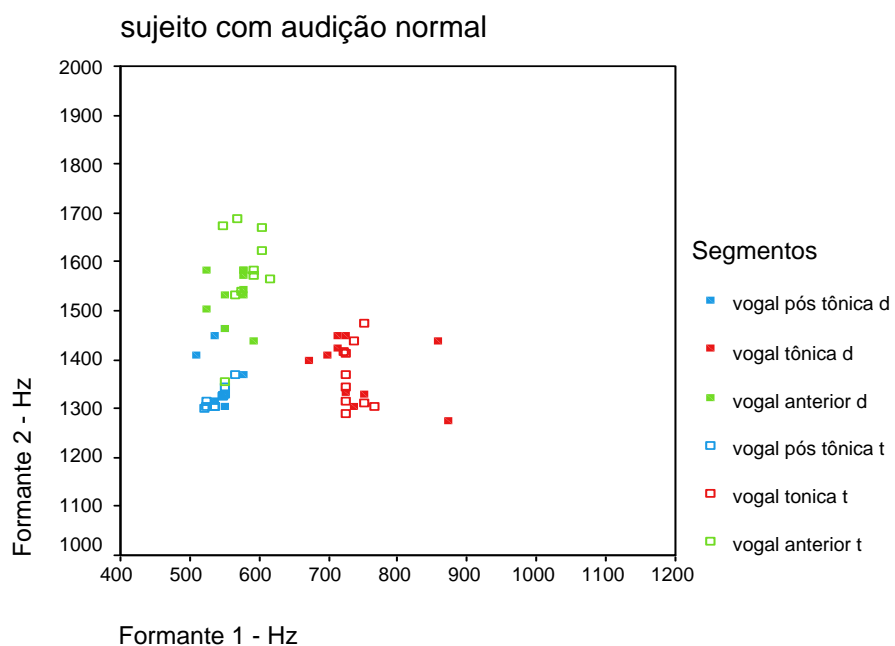


Figura 52 - Gráficos demonstrativos da relação entre as frequências de F1 e F2 medidas na porção estacionária da vogal pós-tônica anterior, tônica e pós-tônica da palavra no contexto alveolar, para os dois sujeitos

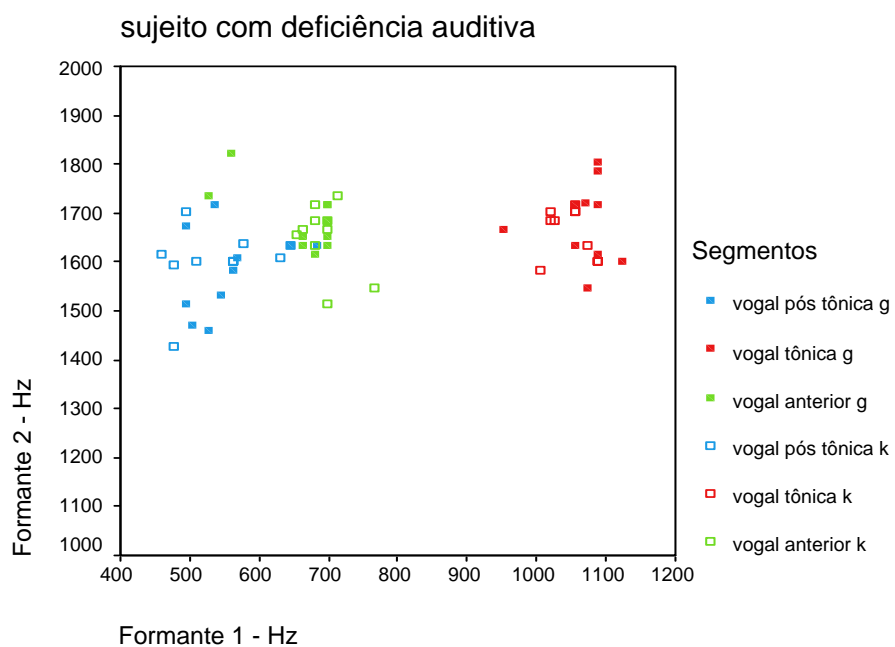
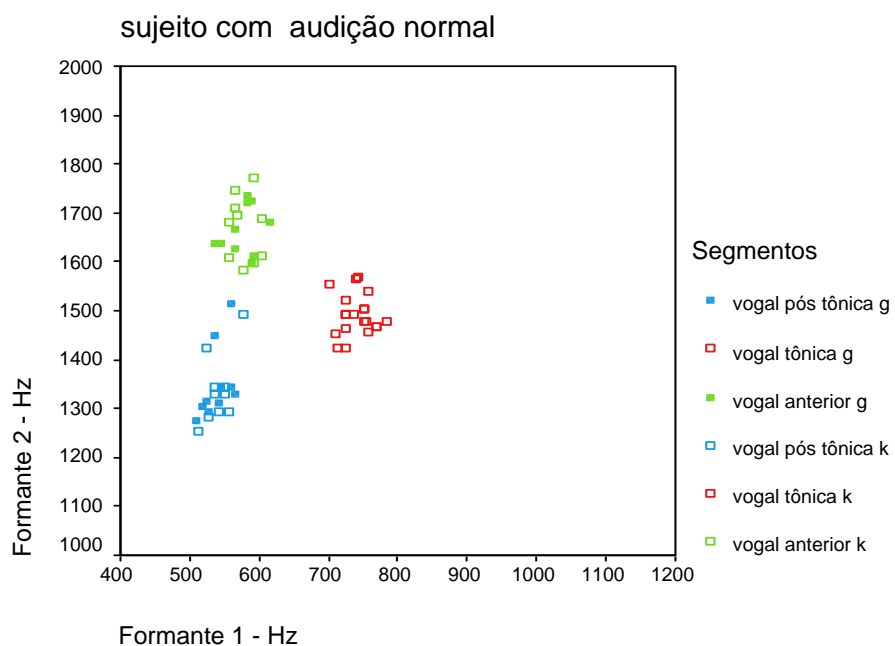


Figura 53 - Gráficos demonstrativos da relação entre as frequências de F1 e F2 medidas na porção estacionária da vogal pós-tônica anterior, tônica e pós-tônica da palavra no contexto velar, para os dois sujeitos

Nos gráficos da figura 54, que demonstram a relação entre F2 e F3 da vogal [a] na posição tônica, observamos que para o sujeito com audição normal ocorre uma separação dos valores relacionadas ao ponto de articulação das plosivas que antecedem a vogal tônica. O mesmo não ocorre para o sujeito com deficiência auditiva, para o qual os pontos não se separam em consequência do ponto de articulação.

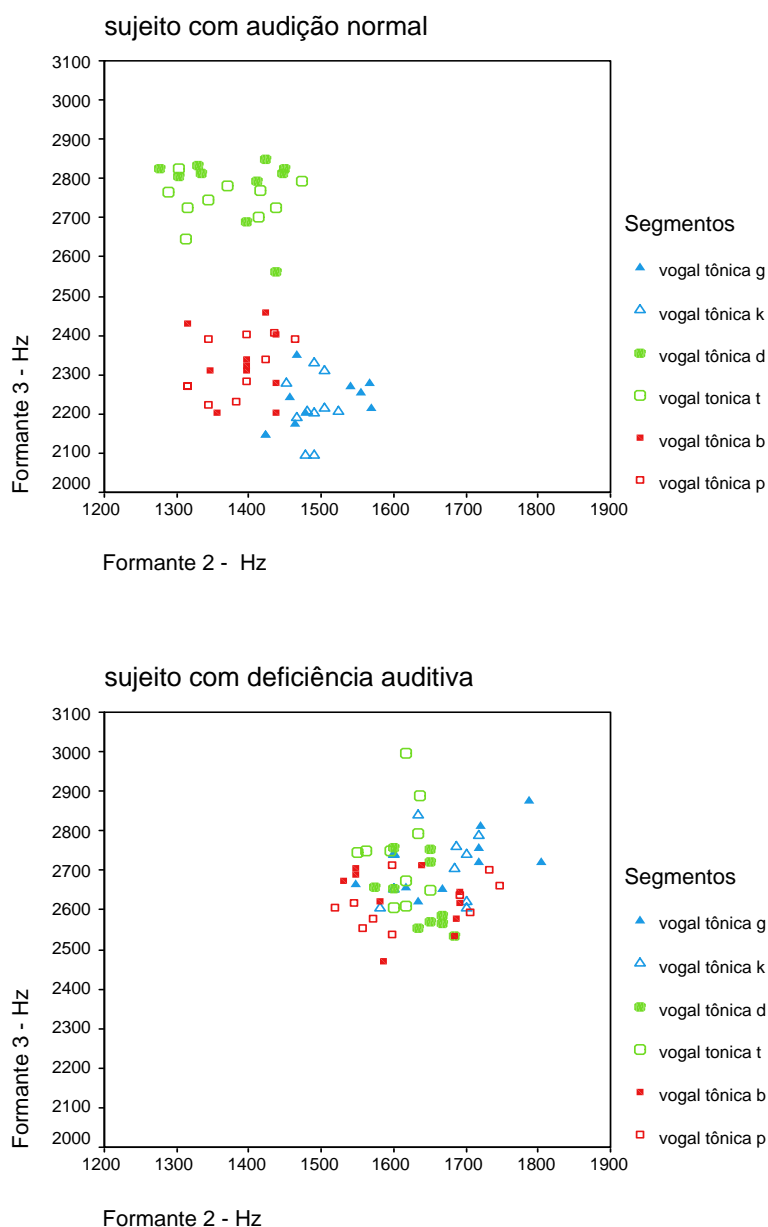


Figura 54 - Gráficos representativos da relação entre as frequências de F2 e F3 medidas na porção estacionária da vogal tônica para os dois sujeitos

Tais dados apontam para a menor ocorrência de efeitos da coarticulação nas produções do sujeito com deficiência auditiva, ou seja, menos informação sobre a consoante está contida na vogal subsequente. Os resultados encontrados na medida da transição dos formantes e da frequência do segundo e terceiro formantes no início da vogal tônica vão confirmar essa tendência.

4.2.7 Transição dos Formantes

Expomos as frequências dos três primeiros formantes medidas ao longo do início da vogal subsequente (transição dos formantes) para cada sujeito e cada consoante plosiva. As tabelas demonstram a média, o desvio-padrão e coeficiente de variação para cada medida. Em relação ao sujeito com audição normal, apresentamos os resultados das medidas realizadas a partir do LPC (com sobreposição do FFT) gerado com o cursor posicionado no pico dos primeiros 11 ciclos regulares da vogal. Esses 11 ciclos correspondem, em média aos primeiros 56 ms da vogal ou cerca de 40% da duração da vogal tônica (tabelas 29, 30, 31, 32, 33 e 34). Para o sujeito com deficiência, auditiva esses 40% correspondem a cerca de 80 ms ou aproximadamente os primeiros 20 ciclos da vogal e, portanto, a medida dos três primeiros formantes foi realizada posicionando o cursor no pico dos primeiros 20 ciclos da vogal (tabelas 35, 36, 37, 38 , 39 e 40)

Em seguida realizamos uma síntese de como se comportou a transição de cada formante. O fim da transição foi determinado utilizando-se os critérios apresentados por Kewley-Port (1982) – o fim da transição foi considerado quando não se verificava uma variação de mais que 10 Hz por mais de dois pontos medidos em seqüência.

Tabela 29 - TRANSIÇÃO DOS FORMANTES – PATA – SUJEITO COM AUDIÇÃO NORMAL – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 11 CICLOS DA VOGAL SUBSEQUENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11
F1 pa	698	719	729	744	752	763	768	771	779	772	778
d.p. f1	16	26	28	34	37	28	31	31	33	30	40
cf.var.	0,02	0,04	0,04	0,05	0,05	0,04	0,04	0,04	0,04	0,04	0,05
F2 pa	1285	1309	1322	1322	1330	1332	1327	1347	1345	1349	1351
d.p. f2	26	36	40	31	28	25	31	30	28	35	40
cf.var.	0,02	0,03	0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03
F3 pa	2084	2092	2117	2144	2158	2175	2218	2253	2276	2285	2301
d.p. f3	53	47	58	69	72	79	69	75	80	74	74
cf.var.	0,03	0,02	0,03	0,03	0,03	0,04	0,03	0,03	0,04	0,03	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 pa – primeiro, segundo e terceiro formante na transição de [p] para [a]

Tabela 30 - TRANSIÇÃO DOS FORMANTES – BATA – SUJEITO COM AUDIÇÃO NORMAL – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 11 CICLOS DA VOGAL SUBSEQUENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11
F1 ba	665	697	716	734	749	763	777	792	795	796	795
d.p.	34	31	34	29	33	36	40	45	44	42	43
cf.var.	0,05	0,04	0,05	0,04	0,04	0,05	0,05	0,06	0,05	0,05	0,05
F2 bata	1370	1384	1399	1407	1409	1404	1397	1402	1394	1396	1386
d.p.	15	33	39	38	36	41	47	41	40	45	46
cf.var.	0,01	0,02	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03
F3 bata	2156	2160	2176	2196	2247	2238	2255	2288	2302	2315	2331
d.p.	38	43	39	50	81	57	60	62	63	72	70
cf.var.	0,02	0,02	0,02	0,02	0,04	0,03	0,03	0,03	0,03	0,03	0,03

Legenda

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ba – primeiro, segundo e terceiro formante na transição de [b] para [a]

Tabela 31 - TRANSIÇÃO DOS FORMANTES – TATA – SUJEITO COM AUDIÇÃO NORMAL – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 11 CICLOS DA VOGAL SUBSEQUENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11
F1 ta	634	643	641	670	687	695	707	716	716	726	731
d.p.	20	27	44	18	18	17	19	26	24	23	22
cf.var.	0,03	0,04	0,07	0,03	0,03	0,02	0,03	0,04	0,03	0,03	0,03
F2 ta	1530	1507	1501	1474	1446	1410	1385	1369	1357	1353	1350
d.p.	51	58	67	79	76	78	66	54	50	48	47
cf.var.	0,03	0,04	0,04	0,05	0,05	0,06	0,05	0,04	0,04	0,04	0,03
F3 ta	2652	2669	2688	2697	2723	2731	2748	2757	2769	2768	2771
d.p.	22	24	20	23	38	34	28	20	19	22	38
cf.var.	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ta – primeiro, segundo e terceiro formante na transição de [t] para [a]

Tabela 32 - TRANSIÇÃO DOS FORMANTES – DATA – SUJEITO COM AUDIÇÃO NORMAL – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 11 CICLOS DA VOGAL SUBSEQUENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11
F1 da	546	582	609	640	659	679	686	703	713	727	737
d.p.	18	18	27	23	23	29	33	39	40	42	46
cf.var.	0,03	0,03	0,04	0,04	0,03	0,04	0,05	0,05	0,06	0,06	0,06
F2 da	1575	1566	1555	1541	1520	1490	1461	1439	1418	1394	1388
d.p.	42	54	58	62	57	62	57	53	57	47	44
cf.var.	0,03	0,03	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,03	0,03
F3 da	2681	2705	2729	2758	2773	2799	2812	2833	2834	2834	2829
d.p.	23	29	33	37	34	49	40	45	46	50	59
cf.var.	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,02	0,02	0,02	0,02

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 da – primeiro, segundo e terceiro formante na transição de [d] para [a]

Tabela 33 - TRANSIÇÃO DOS FORMANTES – CATA – SUJEITO COM AUDIÇÃO NORMAL – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 11 CICLOS DA VOGAL SUBSEQUENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11
F1 ka	615	643	681	710	715	733	747	759	759	761	764
d.p.	16	21	34	33	32	44	31	36	28	29	26
cf.var.	0,03	0,03	0,05	0,05	0,05	0,06	0,04	0,05	0,04	0,04	0,03
F2 ka	1725	1692	1656	1641	1621	1590	1571	1559	1542	1533	1518
d.p.	46	43	55	59	63	60	64	71	70	66	59
cf.var.	0,03	0,03	0,03	0,04	0,04	0,04	0,04	0,05	0,05	0,04	0,04
F3 ka	2085	2094	2092	2101	2119	2138	2146	2172	2172	2177	2193
d.p.	48	40	35	31	24	37	30	34	38	32	36
cf.var.	0,02	0,02	0,02	0,01	0,01	0,02	0,01	0,02	0,02	0,01	0,02

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ka – primeiro, segundo e terceiro formante na transição de [k] para [a]

Tabela 34 - TRANSIÇÃO DOS FORMANTES – GATA – SUJEITO COM AUDIÇÃO NORMAL – FREQUÊNCIA DE F1, F2 E F3 AO DOS PRIMEIROS 11 CICLOS DA VOGAL SUBSEQUENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11
F1 ga	485	524	553	585	621	652	672	701	703	723	732
d.p.	19	23	33	27	28	32	32	28	26	29	26
cf.var.	0,04	0,04	0,06	0,05	0,04	0,05	0,05	0,04	0,04	0,04	0,04
F2 ga	1824	1789	1760	1735	1711	1687	1656	1631	1609	1600	1566
d.p.	45	34	30	27	30	36	38	37	37	40	48
cf.var.	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03
F3 ga	2102	2105	2108	2108	2122	2127	2139	2142	2157	2168	2173
d.p.	35	46	44	47	35	50	51	55	66	51	56
cf.var.	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03	0,02	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ga – primeiro, segundo e terceiro formante na transição de [g] para [a]

Tabela 35 - TRANSIÇÃO DOS FORMANTES – *PATA* – SUJEITO COM DEFICIÊNCIA AUDITIVA – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 20 CICLOS DA VOGAL SUBSEQÜENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F1 pa	831	846	847	859	828	865	888	896	905	918	942	965	973	986	1009	1014	1037	1044	1046	1038
d.p.	74	75	67	51	119	54	57	62	59	55	76	81	88	92	90	57	30	34	29	53
cf.var.	0,09	0,09	0,08	0,06	0,14	0,06	0,06	0,07	0,07	0,06	0,08	0,08	0,09	0,09	0,09	0,06	0,03	0,03	0,03	0,05
F2 pa	1657	1656	1652	1653	1636	1645	1630	1630	1616	1621	1626	1631	1633	1639	1638	1636	1629	1631	1639	1639
d.p.	48	66	59	62	54	76	71	65	66	77	73	76	77	77	82	69	67	72	66	67
cf.var.	0,03	0,04	0,04	0,04	0,03	0,05	0,04	0,04	0,04	0,05	0,04	0,05	0,05	0,05	0,05	0,04	0,04	0,04	0,04	0,04
F3 pa	2745	2732	2728	2718	2721	2723	2738	2741	2726	2709	2689	2683	2680	2679	2683	2667	2663	2645	2646	2659
d.p.	72	44	68	103	111	109	121	117	114	106	92	100	99	90	81	78	92	99	103	96
cf.var.	0,03	0,02	0,02	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,03	0,04	0,04	0,03	0,03	0,03	0,03	0,04	0,04	0,04

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 pa – primeiro, segundo e terceiro formante na transição de [p] para [a]

Tabela 36 - TRANSIÇÃO DOS FORMANTES – *BATA* – SUJEITO COM DEFICIÊNCIA AUDITIVA – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 20 CICLOS DA VOGAL SUBSEQÜENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F1 ba	748	763	781	803	817	846	852	866	876	896	910	924	948	991	1004	1019	1018	1028	1039	1043
d.p.	59	60	48	49	51	56	55	50	42	47	45	53	62	87	81	77	68	58	53	50
cf.var.	0,08	0,08	0,06	0,06	0,06	0,07	0,06	0,06	0,05	0,05	0,05	0,06	0,07	0,09	0,08	0,08	0,07	0,06	0,05	0,05
F2 ba	1624	1619	1624	1627	1629	1633	1627	1635	1631	1629	1625	1623	1634	1644	1642	1639	1627	1626	1625	1629
d.p.	44	39	40	36	38	31	29	31	39	29	31	30	38	37	38	38	34	39	44	43
cf.var.	0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03
F3 ba	2718	2719	2714	2731	2750	2744	2752	2724	2688	2655	2634	2629	2606	2622	2630	2624	2621	2614	2608	2628
d.p.	45	51	49	48	78	92	88	109	128	125	89	81	85	79	86	90	81	77	70	82
cf.var.	0,02	0,02	0,02	0,02	0,03	0,03	0,03	0,04	0,05	0,05	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ba – primeiro, segundo e terceiro formante na transição de [b] para [a]

Tabela 37 - TRANSIÇÃO DOS FORMANTES – TATA – SUJEITO COM DEFICIÊNCIA AUDITIVA - FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 20 CICLOS DA VOGAL SUBSEQÜENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F1 ta	717	766	817	850	885	906	928	952	972	989	1011	1026	1023	1028	1037	1047	1054	1054	1051	1052
d.p.	37	42	58	58	64	61	65	77	57	59	62	60	57	54	49	39	28	28	32	25
cf.var.	0,05	0,06	0,07	0,07	0,07	0,07	0,07	0,08	0,06	0,06	0,06	0,06	0,06	0,05	0,05	0,04	0,03	0,03	0,03	0,02
F2 ta	1590	1598	1605	1612	1606	1604	1614	1624	1627	1631	1631	1619	1619	1605	1599	1602	1604	1602	1604	1609
d.p.	55	37	28	30	37	38	35	19	24	19	24	27	27	34	33	36	32	30	30	31
cf.var.	0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,01	0,01	0,01	0,01	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02	0,02
F3 ta	2819	2822	2825	2796	2817	2836	2840	2847	2843	2862	2855	2860	2857	2851	2839	2842	2829	2839	2825	2818
d.p.	115	99	98	72	74	73	65	92	99	84	103	116	140	156	147	152	165	179	181	169
cf.var.	0,04	0,04	0,03	0,03	0,03	0,03	0,02	0,03	0,03	0,03	0,04	0,04	0,05	0,05	0,05	0,05	0,06	0,06	0,06	0,06

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ta – primeiro, segundo e terceiro formante na transição de [t] para [a]

Tabela 38 - TRANSIÇÃO DOS FORMANTES – DATA – SUJEITO COM DEFICIÊNCIA AUDITIVA – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 20 CICLOS DA VOGAL SUBSEQÜENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F1 da	697	722	755	802	841	877	913	938	959	969	984	998	1001	1006	1015	1018	1013	1018	1019	1027
d.p.	45	53	57	54	60	72	78	85	83	74	64	72	73	70	49	52	43	40	41	39
cf.var.	0,06	0,07	0,08	0,07	0,07	0,08	0,09	0,09	0,09	0,08	0,07	0,07	0,07	0,07	0,05	0,05	0,04	0,04	0,04	0,04
F2 da	1649	1636	1614	1620	1626	1631	1633	1638	1638	1645	1641	1646	1646	1646	1641	1641	1636	1641	1645	1638
d.p.	80	93	56	52	54	52	57	54	46	49	39	44	38	35	35	37	45	47	45	39
cf.var.	0,05	0,06	0,03	0,03	0,03	0,03	0,04	0,03	0,03	0,03	0,02	0,03	0,02	0,02	0,02	0,02	0,03	0,03	0,03	0,02
F3 da	2855	2838	2848	2848	2852	2822	2817	2834	2864	2879	2904	2901	2894	2906	2905	2918	2893	2901	2871	2868
d.p.	92	84	100	69	63	60	66	102	100	106	120	105	109	116	110	129	145	143	140	145
cf.var.	0,03	0,03	0,04	0,02	0,02	0,02	0,02	0,04	0,03	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,05	0,05	0,05	0,05

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 da – primeiro, segundo e terceiro formante na transição de [d] para [a]

Tabela 39 - TRANSIÇÃO DOS FORMANTES – CATA – SUJEITO COM DEFICIÊNCIA AUDITIVA – FREQUÊNCIA DE F1, F2 E F3 AO LONGO DOS PRIMEIROS 20 CICLOS DA VOGAL SUBSEQÜENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F1 ka	862	875	889	887	892	904	916	927	935	959	978	996	1008	1021	1035	1039	1039	1046	1049	1051
d.p.	54	61	68	64	59	60	61	54	50	49	51	41	43	49	40	34	25	23	20	33
cf.var.	0,06	0,07	0,08	0,07	0,07	0,07	0,07	0,06	0,05	0,05	0,05	0,04	0,04	0,05	0,04	0,03	0,02	0,02	0,02	0,03
F2 ka	1629	1595	1595	1613	1616	1626	1628	1623	1622	1620	1628	1634	1636	1641	1639	1646	1659	1665	1668	1679
d.p.	34	52	53	54	49	44	43	43	41	37	37	38	43	39	41	40	46	50	44	45
cf.var.	0,02	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,02	0,02	0,02	0,03	0,02	0,03	0,02	0,03	0,03	0,03	0,03
F3 ka	2860	2836	2798	2793	2793	2796	2785	2756	2753	2754	2743	2738	2737	2741	2747	2735	2727	2729	2727	2723
d.p.	66	59	52	83	99	76	59	53	79	94	93	89	84	75	69	76	68	68	85	89
cf.var.	0,02	0,02	0,02	0,03	0,04	0,03	0,02	0,02	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,02	0,02	0,03	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ka – primeiro, segundo e terceiro formante na transição de [k] para [a]

Tabela 40 - TRANSIÇÃO DOS FORMANTES – GATA – SUJEITO COM DEFICIÊNCIA AUDITIVA – FREQUÊNCIA DE F1, F2 E F3 DOS PRIMEIROS 20 CICLOS DA VOGAL SUBSEQÜENTE À PLOSIVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
F1 ga	856	884	901	917	923	938	951	957	946	952	967	966	974	996	1013	1017	1024	1037	1038	1046
d.p.	64	57	70	73	56	53	64	64	71	78	76	81	72	73	72	73	76	65	62	68
cf.var.	0,07	0,06	0,08	0,08	0,06	0,06	0,07	0,07	0,07	0,08	0,08	0,08	0,07	0,07	0,07	0,07	0,07	0,06	0,06	0,07
F2 ga	1643	1643	1626	1629	1624	1627	1636	1643	1643	1651	1653	1655	1665	1675	1672	1684	1685	1685	1689	1685
d.p.	61	59	70	58	38	43	43	51	43	49	47	49	54	58	58	60	62	65	64	66
cf.var.	0,04	0,04	0,04	0,04	0,02	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,04	0,04	0,04	0,04	0,04
F3 ga	2791	2799	2790	2796	2809	2797	2775	2765	2751	2734	2715	2712	2703	2697	2713	2713	2715	2722	2731	2729
d.p.	54	71	61	60	67	69	89	78	65	63	46	44	61	60	78	72	74	78	76	74
cf.var.	0,02	0,03	0,02	0,02	0,02	0,02	0,03	0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,03	0,03	0,03	0,03	0,03	0,03

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

F1 F2 F3 ga – primeiro, segundo e terceiro formante na transição de [g] para [a]

Em síntese, para o sujeito com audição normal, verificamos o seguinte comportamento na transição dos formantes, considerando a variação entre os valores médios de frequência encontrados o início da vogal e no ponto considerado como o fim da transição .

Transição de [p] para [a]:

F1 - variação de +70 Hz

F2 - variação de +62 Hz

F3 - variação de +217 Hz

Transição de [b] para [a]:

F1 - variação de +126 Hz

F2 - variação de +37 Hz

F3 - variação de +175 Hz

Transição de [t] para [a]:

F1 - variação de +92 Hz

F2 - variação de - 173 Hz

F3 - variação de +117 Hz

Transição de [d] para [a]:

F1 - variação de +191 Hz

F2 - variação de -187 Hz

F3 - variação de +152 Hz

Transição de [k] para [a]:

F1 - variação de +144 Hz

F2 - variação de -207 Hz

F3 - variação de +87 Hz

Transição de [g] para [a]:

F1 - variação de +247 Hz

F2 - variação de -258 Hz

F3 - variação de +170 Hz

Também sintetizando os dados do sujeito com deficiência auditiva, observamos as seguintes variações entre os valores de frequência no início da transição e no

ponto considerado como fim da transição.

Transição de [p] para [a]:

F1 - variação de +207 Hz

F2 - variação de -18 Hz

F3 - variação de - 82 Hz

Transição de [b] para [a]:

F1 - variação de +295 Hz

F2 - variação de + 5 Hz

F3 - variação de -84 Hz

Transição de [t] para [a]:

F1 - variação de +330 Hz

F2 - variação de +19 Hz

F3 - variação de - 1 Hz

Transição de [d] para [a]:

F1 - variação de +330 Hz

F2 - variação de -11 Hz

F3 - variação de +13 Hz

Transição de [k] para [a]:

F1 - variação de +189 Hz

F2 - variação de +50 Hz

F3 - variação de -104 Hz

Transição de [g] para [a]:

F1 - variação de +181 Hz

F2 - variação de +41 Hz

F3 - variação de -72 hz

No primeiro caso (sujeito com audição normal), a transição dos formantes segue os dados da literatura (Dellatre et al, 1955; Stevens, 1998). O primeiro formante tem transição ascendente em todas as plosivas, sendo mais acentuada nas vozeadas (*onset* de F1 é mais baixo, como já está descrito nos resultados das medidas de F1 no *onset* da vogal subsequente). Para as bilabiais, todos os

formantes apresentaram transição ascendente. Para as alveolares, F2 apresentou transição descendente e F3 ascendente. No caso das velares, F2 teve transição descendente e F3 ascendente, sendo as freqüências no início desses bastante próximas.

No segundo caso (produções do sujeito com deficiência auditiva), F1 também apresentou transição ascendente em todos os contextos de plosivas. Com relação à transição de F2 e F3 observamos que estas em geral são planas e indiferenciadas quanto ao ponto de articulação, e não seguem a direção esperada para cada ponto de articulação, no caso das velares [k] e [g] e da alveolar [t].

A fim de comparar os valores de variação de freqüência para cada sujeito, calculamos o $\Delta F / \Delta T$, ou seja, a relação entre as diferenças de freqüência para cada transição e o intervalo de tempo durante o qual essa variação foi medida. Os valores estão apresentados nas tabelas 41 e 42.

Tabela 41 – TRANSIÇÃO DOS FORMANTES - VARIAÇÃO DE FREQUÊNCIA EM RELAÇÃO AO TEMPO - $\Delta F / \Delta T$ NAS PRODUÇÕES DO SUJEITO COM AUDIÇÃO NORMAL

	Transição [pa]	Transição [ba]	Transição [ta]	Transição [da]	Transição [ka]	Transição [ga]
	F1	F1	F1	F1	F1	F1
média	1,04	2,48	1,70	3,81	2,63	4,52
d.p	0,95	0,71	0,44	0,88	0,59	0,44
	F2	F2	F2	F2	F2	F2
média	1,30	0,41	-3,52	-5,17	-5,50	-6,58
d.p.	0,67	0,68	0,82	0,89	0,93	0,72
	F3	F3	F3	F3	F3	F3
média	4,22	4,40	1,78	2,20	2,29	2,21
d.p.	1,35	1,42	1,14	1,24	0,93	1,34

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

transição [Ca] – $\Delta F / \Delta T$ onde C é uma das plosivas

Tabela 42 - TRANSIÇÃO DOS FORMANTES - VARIAÇÃO DE FREQUÊNCIA EM RELAÇÃO AO TEMPO - $\Delta F/\Delta T$ - NAS PRODUÇÕES DO SUJEITO COM DEFICIÊNCIA AUDITIVA

	Transição [pa]	Transição [ba]	Transição [ta]	Transição [da]	Transição de [ka]	Transição de [ga]
	F1	F1	F1	F1	F1	F1
média	2,59	3,69	4,19	4,12	2,37	2,38
d.p.	0,85	1,01	0,51	0,90	0,86	0,59
	F2	F2	F2	F2	F2	F2
média	-0,22	0,06	0,24	-0,14	0,63	0,52
d.p.	1,21	0,85	0,73	1,07	0,61	1,01
	F3	F3	F3	F3	F3	F3
média	-1,08	-1,12	0,12	0,16	-1,70	-0,77
d.p.	1,46	1,02	2,70	2,33	1,15	1,16

Legenda:

d.p. - desvio padrão

cf.var. - coeficiente de variação

transição [Ca] – $\Delta F/\Delta T$ onde C é uma das plosivas.

Comparamos as variações de F2, que estão relacionadas à identificação do ponto de articulação, e não encontramos diferenças significativas entre os valores das produções do sujeito com deficiência auditiva. Já para o sujeito com audição normal, as variações de F2 apresentaram diferenças significativas entre os 3 pontos de articulação [$F_{(5,54)}=172,751$, $p=0,000$] entre [pa, ta, ka], $p=0,000$; entre [ba] e [da], $p=0,000$; entre [da] e [ga], $p=0,001$; entre [ba] e [ga], $p=0,000$.

A frequência do início da transição dos formantes também diferencia-se de acordo com o ponto de articulação. Para F2, a análise estatística mostra diferenças significativas na frequência de início do formante para os três pontos de articulação [$F_{(5,54)}=260,548$, $p=0,000$], sendo que as bilabiais apresentam as frequências mais baixas, entre 1285 e 1370 Hz, seguidas das alveolares com valores entre 1530 e 1582, e das velares com valores entre 1725 e 1823 Hz.

Para os valores das frequências de F3, as produções do sujeito com audição normal apresentaram sobreposição de valores para as bilabiais e velares, para as quais não foram encontradas diferenças significativas, e variaram de 2083Hz a 2155Hz. Para as frequências de F3, no início da vogal subsequente às consoantes plosivas alveolares, foram encontradas diferenças significativas [$F_{(5,54)}=572,990$;

$p=0,000$], sendo que as alveolares apresentaram valores mais altos que as bilabiais e velares, ($p=0,000$) variando de 2652Hz a 2681Hz.

Para o sujeito com deficiência auditiva, quanto ao valor da frequência de F2 no início da vogal subsequente, não foram encontradas diferenças significativas entre os seis contextos de plosivas, e sua variação foi de 1590 Hz a 1656 Hz. Para F3, observamos a ocorrência de transição descendente para as velares e bilabiais e plana para as alveolares. Além do padrão da transição de F3 não corresponder ao esperado para os três pontos de articulação, a frequência de F3 no início da transição não revelou diferenças significativas sistemáticas entre os pontos de articulação. A análise estatística aponta para diferenças significativas [$F_{(5,54)} = 5,559$, $p=0,000$] que correspondem à diferença verificada apenas em relação à bilabial vozeada [b], para a qual encontramos os valores de frequência do F3 no início da vogal mais baixos e se diferenciaram significativamente dos valores de [d], ($p=0,015$) e [k], ($p=0,011$). Para as outras sílabas, não foram observadas diferenças significativas com relação ao valor da frequência de F3 no início da transição (os valores variaram de 2718 Hz - para [ba] - a 2860 Hz - para [ka]).

O gráfico da figura 55, a seguir, representa a variação dos formantes nas seis sílabas iniciadas pelas seis plosivas do PB, combinadas com a vogal [a]. Nas produções do sujeito com deficiência auditiva, pode-se visualizar claramente a homogeneidade entre as transições, principalmente no que se refere à F2, além de trajetórias de F2 e F3 com padrões diferentes daqueles característicos de cada ponto de articulação:

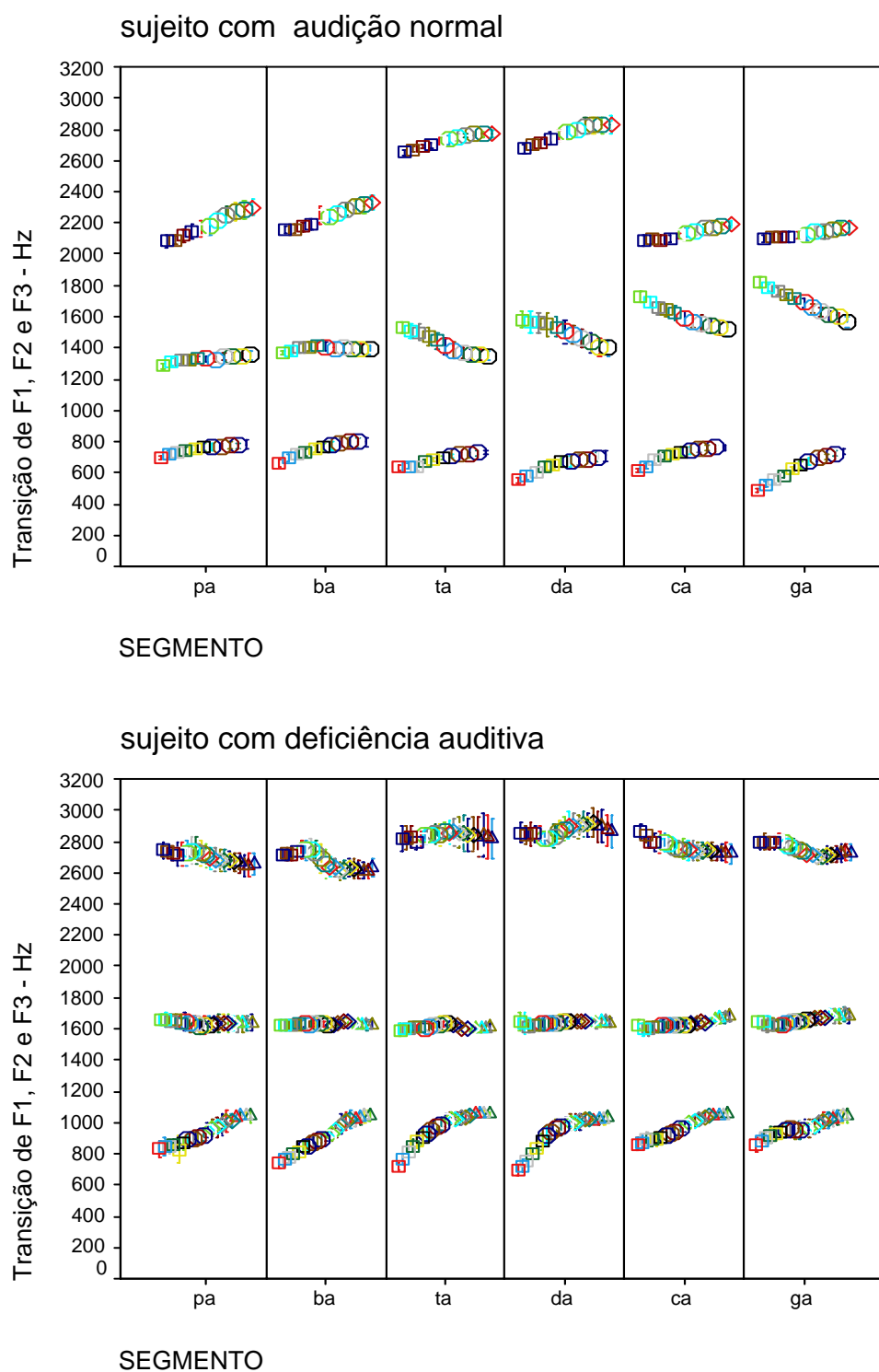


Figura 55 - Gráficos representativos da distribuição dos valores das frequências na transição de F1, F2 e F3 na produção dos dois sujeitos

Considerando os valores da frequência no início da transição, podemos observar na figura 56 a relação entre os valores para F1 e F2:

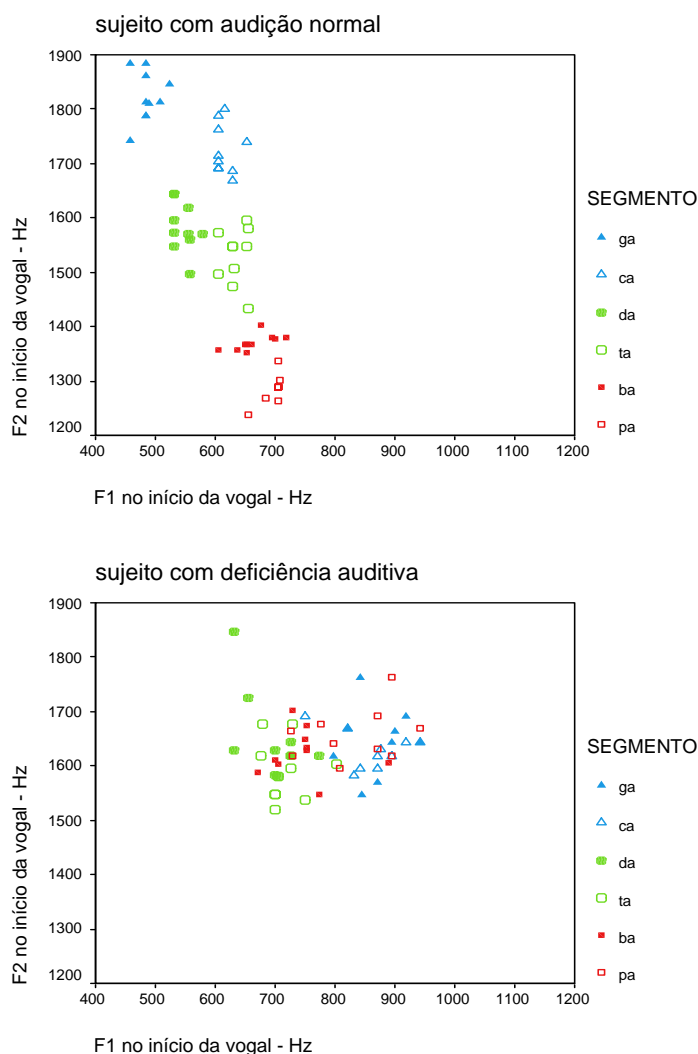


Figura 56 - Gráficos representativos da distribuição dos valores de F1 e F2 no início da transição na produção dos dois sujeitos

Podemos verificar que, para as produções do sujeito com audição normal, ocorre uma diferenciação entre os três pontos de articulação no que se refere à F2 e também entre os pares vozeados e não vozeados no que se refere à F1, o que não se observa nos dados referentes às produções do sujeito com deficiência auditiva.

No gráfico da figura 57, encontram-se plotados os valores de F2 e F3 no início

da transição, encontrados na produção do sujeito com deficiência auditiva e do com audição normal. Nestes também fica clara a diferenciação entre os pontos de articulação nos dados das produções do sujeito com audição normal e a não-diferenciação nos dados do sujeito com deficiência auditiva:

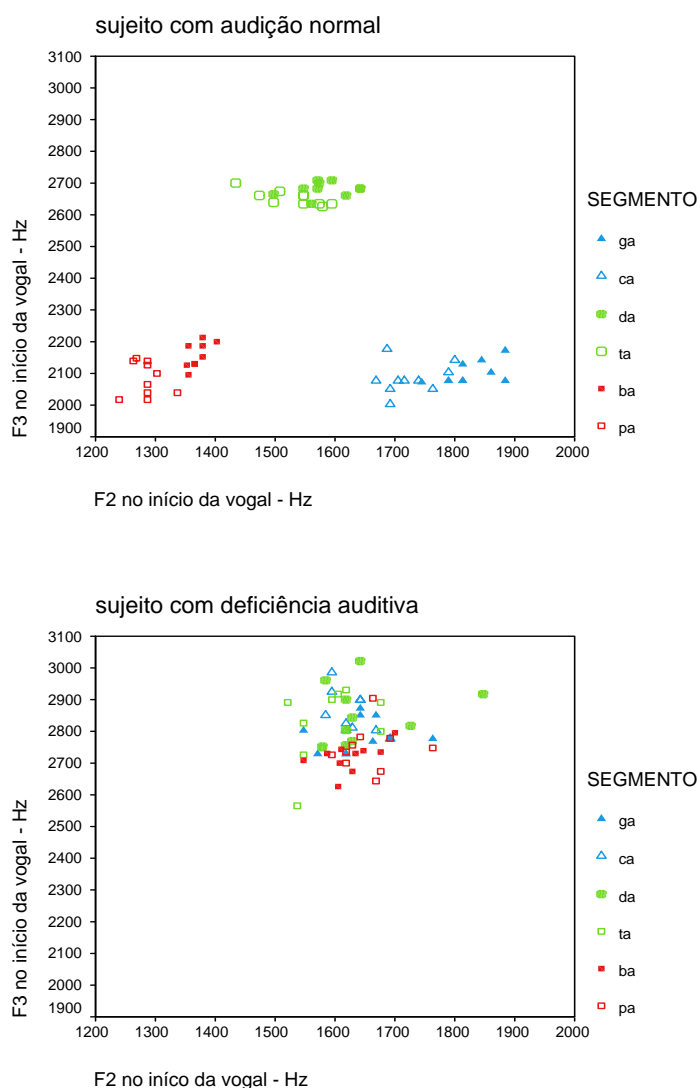
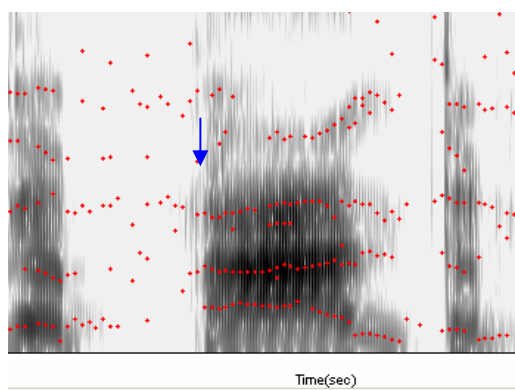


Figura 57 - Gráficos representativos da distribuição dos valores de F2 e F3 no início da transição na produção dos dois sujeitos

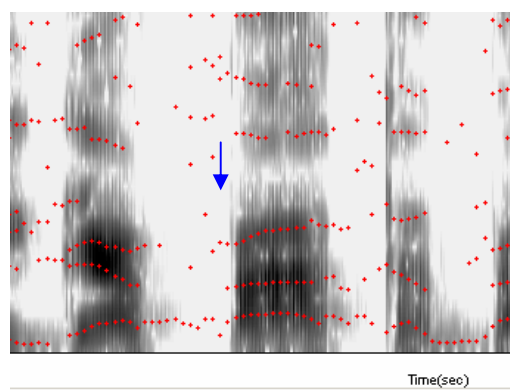
A seguir, apresentamos alguns exemplos de espectrogramas ilustrando a transição das seis consoantes plosivas para a vogal [a] na produção dos dois sujeitos (Figura 58, 59 e 60).

Sujeito com deficiência auditiva

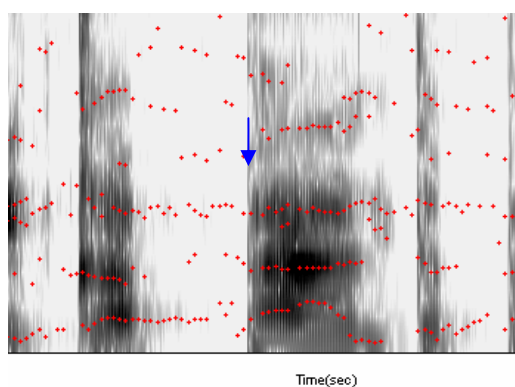


diga PAta baixinho

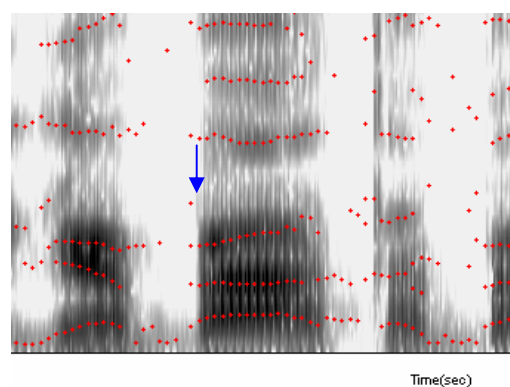
Sujeito com audição normal



diga PAta baixinho



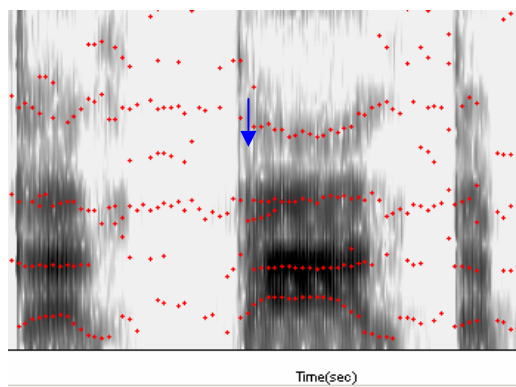
diga BAta baixinho



diga BAta baixinho

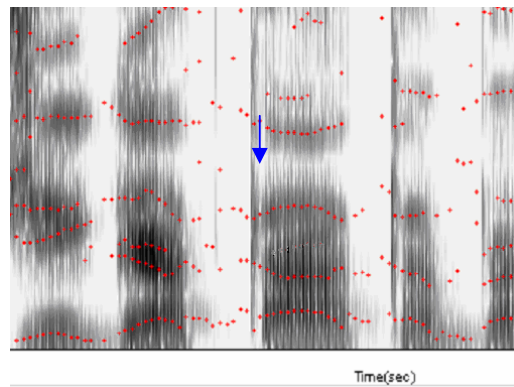
Figura 58 - Espectrogramas demonstrando a transição dos formantes da consoante plosiva bilabial (vozeada e não-vozeada) para a vogal [a] nas produções dos dois sujeitos

Sujeito com deficiência auditiva

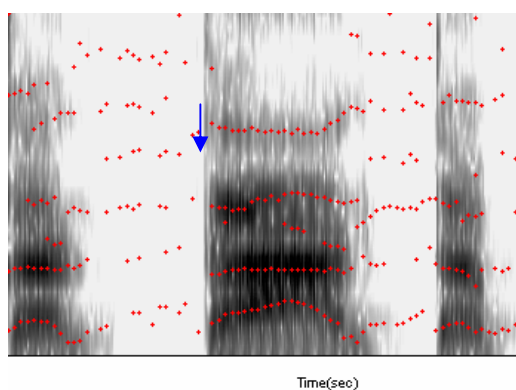


diga TAta baixinho

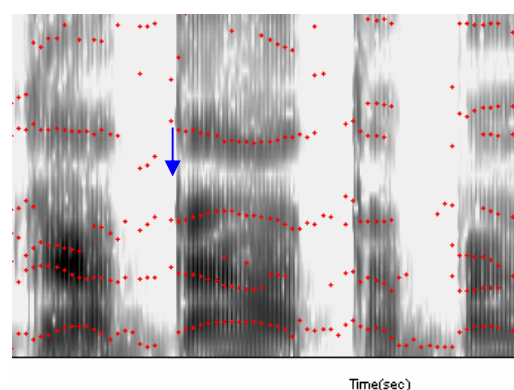
Sujeito com audição normal



diga TAta baixinho



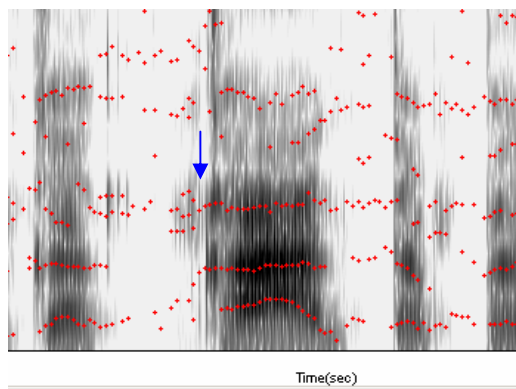
diga DAta baixinho



diga DAta baixinho

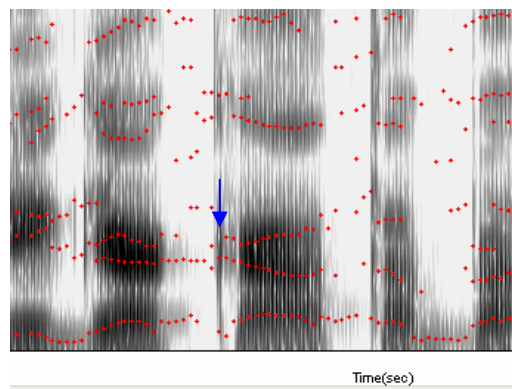
Figura 59 - Espectrogramas demonstrando a transição dos formantes da consoante plosiva alveolar (vozeada e não-vozeada) para a vogal [a] nas produções dos dois sujeitos

Sujeito com deficiência auditiva

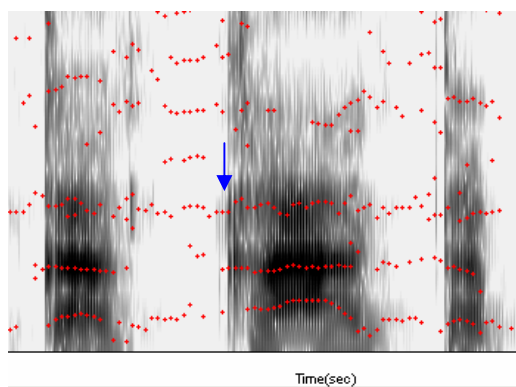


diga CAta baixinho

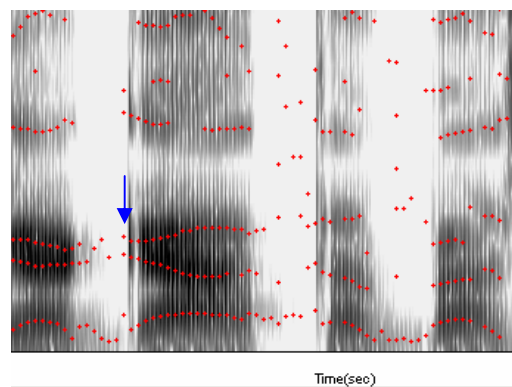
Sujeito com audição normal



diga CAta baixinho



diga GAta baixinho



diga GAta baixinho

Figura 60 - Espectrogramas demonstrando a transição dos formantes da consoante plosiva velar (vozeada e não-vozeada) para a vogal [a] nas produções dos dois sujeitos

Com relação às transições de formantes e aos valores de frequência no início da vogal subsequente, ou seja, início da transição, os resultados da análise acústica das produções do sujeito com deficiência auditiva sugerem a menor coarticulação, ao menos no que se refere ao contexto fonético estudado – sílaba do tipo CV (onde C é uma das seis plosivas do PB e V a vogal [a], em posição tônica, em início de palavra).

Os resultados apresentados serão discutidos no próximo capítulo.

5 DISCUSSÃO

A condução de estudos sobre a produção e percepção da fala com base na análise espectrográfica do sinal acústico, independentemente da corrente lingüística que se adote, é extremamente valiosa, pois permite que se façam inferências sobre as ações dos articuladores, impossíveis de serem realizadas somente em simples inspeção auditiva e, portanto, ela fornece subsídios para a melhor compreensão dos fenômenos de fala.

Neste estudo, pudemos verificar a coerência entre os dados das tarefas de percepção das plosivas e os dados da produção. Se considerarmos os resultados da tarefa na qual as produções do sujeito com deficiência auditiva foram submetidas à avaliação de sujeitos com audição normal (tarefa 1), verificamos, em primeiro lugar, que o modo de articulação foi bem identificado, uma vez que as plosivas foram confundidas com outra classe de consoantes apenas 8 vezes em 360 apresentações. Isso provavelmente, se deve ao fato de o sujeito com deficiência auditiva implementar efetivamente o fechamento do trato vocal, o que resulta em um intervalo de silêncio, que é uma pista acústica necessária para a percepção de uma plosiva (Kent, Read, 1992).

No que se refere ao contraste de vozeamento, entretanto, os resultados das tarefas de identificação apontam para uma dificuldade dos sujeitos com audição normal em identificar plosivas vozeadas nas produções do sujeito com deficiência auditiva. As plosivas não-vozeadas foram identificadas como tais (desconsiderando-se os erros por ponto de articulação) em 89% das vezes, enquanto que as plosivas vozeadas foram identificadas corretamente apenas em 8% das vezes. Por outro lado, verificamos que o sujeito com deficiência auditiva também não identificou auditivamente o contraste de vozeamento, nem nas produções do sujeito com audição normal, nem em suas próprias produções. Os resultados para as tarefas de discriminação das plosivas realizados por esse sujeito mostraram que a discriminação do contraste de vozeamento não passa do nível de acerto casual, sendo de 0,2% para as produções do sujeito com audição normal e de 0,1% para as

próprias produções. Por outro lado, verificamos que as produções do sujeito com audição normal foram todas identificadas corretamente por sujeitos com audição normal. Esses dados de percepção vão ao encontro dos resultados encontrados na análise acústica das produções do sujeito com deficiência auditiva e com audição normal.

Com relação aos dados de produção, um dos aspectos mais evidentes na análise acústica das produções do sujeito com deficiência auditiva refere-se à dificuldade na implementação do contraste vozeado/não-vozeado. Para todos os parâmetros acústicos associados ao vozeamento que foram investigados, não foram encontradas, nas produções do sujeito com deficiência auditiva, evidências de diferenciação entre o que seriam os pares vozeados e não-vozeados, ou seja, não foram observadas diferenças entre as medidas de *VOT*, frequência de F1 no início da vogal subsequente, f0 no início da vogal subsequente, duração da consoante e na relação de energia entre a consoante e a vogal subsequente. A ausência de pistas acústicas vinculadas ao contraste de vozeamento nas produções do sujeito com deficiência auditiva explica a não-identificação desses sons pelos sujeitos com audição normal. Por outro lado, todos esses parâmetros apresentaram-se diferenciados nas produções do sujeito com audição normal, o que nos dá indícios de que o contraste de vozeamento está marcado em suas produções a partir de várias pistas acústicas. Entretanto, as pistas acústicas relacionadas ao vozeamento presentes na fala do sujeito com audição normal, de forma suficiente para que todas as suas produções fossem corretamente identificadas, não foram percebidas pelo sujeito com deficiência auditiva, em virtude de limites perceptivo-auditivos impostos pela perda auditiva.

É importante lembrar que as habilidades em perceber contrastes de fala são bastante variadas entre sujeitos com perda auditiva e nem sempre apresentam uma relação direta com o grau de perda auditiva (Sammeth et al, 1996; Dorman, Hannley, 1985) embora, em geral, estejam mais prejudicadas para sujeitos com maior grau de perda auditiva (Boothroyd, 1982; Longone, Borges, 1996; Delgado, Bevilacqua, 1999). Se observarmos os resultados apresentados no capítulo 3 (pág. 81, figura 16), o sujeito número 9, que é o mesmo sujeito cuja produção de fala estamos aqui considerando, apresenta resultados diversos da maior parte do grupo. Esse sujeito, e

também o de número 10, apresentaram evidências de maior dificuldade para discriminar o contraste de vozeamento em relação ao contraste de ponto de articulação. Esses resultados, contrariam a literatura, que aponta para a maior preservação na identificação do contraste de vozeamento, e maior dificuldade em perceber contrastes de ponto de articulação (Miller, Nicely, 1955; Pickett, Danaher, 1975; Boothroyd, 1982) e, explicitam a importância do estudo de cada caso.

No que se refere ao ponto de articulação, também encontramos concordância entre os dados de percepção e de produção. Os resultados da análise acústica revelaram a diferenciação dos três pontos de articulação nas produções do sujeito com audição normal, que foram totalmente identificadas pelos sujeitos com audição normal, tanto em relação à transição dos formantes quanto aos valores das frequências no início de F2 e F3.

Já na identificação do contraste de ponto de articulação nas produções do sujeito com deficiência auditiva por sujeitos com audição normal, observamos que ele foi mais identificado que o contraste de vozeamento, principalmente no que tange às bilabiais e velares. Esse também foi o contraste mais bem discriminado pelo sujeito com deficiência auditiva, que apresentou a maior porcentagem de acerto quando a oposição envolvia os pontos de articulação bilabial e velar. Entretanto, com relação à identificação do ponto alveolar, pudemos observar não só a maior dificuldade de identificação por parte dos sujeitos com audição normal como também a maior dificuldade por parte do sujeito com deficiência auditiva na discriminação envolvendo o ponto alveolar. Esses resultados, se confrontados com os dados de produção, podem ser explicados, especialmente, pelos padrões de transição de formantes encontrados nas produções do sujeito com deficiência auditiva. Além das transições dos formantes não apresentarem as trajetórias esperadas para cada ponto de articulação, os valores das frequências no início da transição não se diferenciaram entre eles. As transições de F2, que evidenciam aspectos da coarticulação entre as consoantes e vogais, são planas e indiferenciadas, sugerindo que a co-produção dos movimentos dos articuladores para as produções de consoantes e vogais não ocorreu na fala do sujeito com deficiência auditiva do mesmo modo que nas produções do sujeito com audição normal, tanto neste estudo como nos trabalhos considerados na revisão da literatura (Waldstein, Baum, 1991; Baum, Waldstein,

1991; Okalidou, Harris, 1999) .

A maior dificuldade em relação à identificação/produção das alveolares na fala do sujeito com deficiência auditiva pode ter ocorrido em virtude de essa produção envolver duas variáveis do trato – ponta da língua e corpo da língua – associadas a um único articulador. Além disso, a produção da plosiva alveolar requer maior precisão no movimento articulatorio, por envolver a oclusão com a ponta da língua, o que exige maior controle motor na realização do gesto, (Madureira et al, 2002).

Se tomarmos como referência o estudo de Smits et al (1996), podemos compreender o porquê das velares terem sido bem identificadas nas produções do sujeito com deficiência auditiva. Segundo esses autores, a identificação dos pontos de articulação das plosivas depende de pistas acústicas diferentes de acordo com o ponto de articulação. A identificação das velares parece que se dá mais a partir das pistas do *burst* e menos a partir da transição. Já para as bilabiais e alveolares, a transição dos formantes é a pista acústica mais robusta. Portanto, mesmo sem a presença da transição característica da plosiva velar em direção à vogal [a] subsequente, os sujeitos com audição normal puderam identificar as produções do sujeito com deficiência auditiva, e provavelmente apoiaram-se nas pistas acústicas contidas no *burst*. As plosivas bilabiais puderam ser identificadas, porque o padrão de formantes encontrado nas produções do sujeito com deficiência auditiva é semelhante ao padrão encontrado na produção das bilabiais do sujeito com audição normal, principalmente no que se refere ao F2. Para as alveolares, contudo, a informação de formantes não correspondeu ao padrão esperado. Se essa é a pista mais importante para a identificação do ponto alveolar, a maior dificuldade dos sujeitos com audição normal em identificar as plosivas alveolares produzidas pelo sujeito com deficiência auditiva pode ser explicada.

No caso de considerarmos os dados de Kewley-Port (1982) e Egstrand et al (2000), que apontam para a frequência do início da transição de F2 e F3 como um parâmetro acústico importante para a identificação do ponto de articulação das plosivas, também verificamos que as produções do sujeito com deficiência auditiva não fornecem essa pista para o ouvinte.

Gostaríamos, assim de ressaltar que esses dados de produção e percepção

encontram-se altamente congruentes. Há coerência entre os resultados da percepção e os dados da produção do sujeito com deficiência auditiva: o contraste bilabial *versus* velar foi o que apresentou maior porcentagem de acerto na tarefa de discriminação e, o mais bem produzido, considerando-se sua identificação pelos sujeitos com audição normal. Entretanto, os resultados da tarefa de discriminação (tarefa nº 4), na qual o sujeito com deficiência auditiva julgou suas próprias produções, divergem dos anteriores. Nessa tarefa, a porcentagem de discriminação do contraste de ponto de articulação ficou bem abaixo do que o observado nos resultados da tarefa de discriminação das produções do sujeito com audição normal pelo sujeito com deficiência auditiva (tarefa nº 3). Surpreendente foi o fato de que o sujeito com deficiência auditiva identificou corretamente apenas 4 dos 20 itens correspondentes ao ponto velar apresentado em oposição ao bilabial.

Como o sujeito com deficiência auditiva produz o ponto de articulação velar de uma forma que ele é bem percebido e, por outro lado, é o ponto de articulação mais bem discriminado, principalmente quando em oposição ao bilabial, poderíamos esperar que ele também percebesse esse contraste em suas próprias produções, o que não ocorreu. Uma possível explicação para esse resultado seria que, para o sujeito com deficiência auditiva, a transição dos formantes e/ou informação de frequência de F2 e F3 no início da vogal são as pistas mais perceptíveis. Além disso, a informação do *burst*, que provavelmente foi a pista usada pelos sujeitos com audição normal para perceberem o ponto velar em sua fala, não teria sido percebida pelo sujeito com deficiência auditiva. Assim, podemos supor que a falta da pista da transição dos formantes na fala do sujeito com deficiência auditiva, sugestiva de menor coarticulação, seja um fator que contribuiu para que ele tenha encontrado maior dificuldade em sua percepção.

As evidências de menor coarticulação na fala do sujeito com deficiência auditiva não ficaram restritas a aspectos espectrais no início da vogal. Os dados relacionados às medidas de duração revelam, na produção do sujeito com deficiência auditiva, o alongamento exagerado de determinados segmentos, entre eles a consoante plosiva. Embora a maior duração da vogal em posição tônica e, por outro lado, a menor duração da vogal átona seja o padrão esperado para o PB, a diferença entre tais valores encontra-se bastante aumentada na fala do sujeito com deficiência auditiva.

O alongamento da vogal em posição tônica nos dados de produção desse sujeito coincidiu com a elevação também exacerbada do F1 nessa vogal, que apresenta frequências bastante elevadas em relação às vogais nas posições átonas, sinalizando para a posição da língua muito mais abaixada na posição tônica. Esse dado sugere um movimento de abertura da mandíbula muito maior na realização da tônica em comparação às realizações nas vogais em posição átona. O aumento da magnitude do movimento do articulador resulta em valores de duração mais elevados.

Covém ressaltar que, no caso desse sujeito que apresenta deficiência auditiva de grau severo desde os 18 meses de idade, não podemos deixar de considerar o aspecto das pistas visuais que ele utiliza e utilizou para perceber e produzir a fala. O aumento da amplitude dos ajustes que envolvem os articuladores mais visíveis (abaixamento da mandíbula) e a movimentação mais restrita ou não sincronizada de estruturas menos visíveis ou totalmente invisíveis (como é o caso da língua e pregas vocais, respectivamente) podem ser alterações plausíveis para um sujeito que se utiliza da leitura oro-facial como um canal importante de percepção de fala.

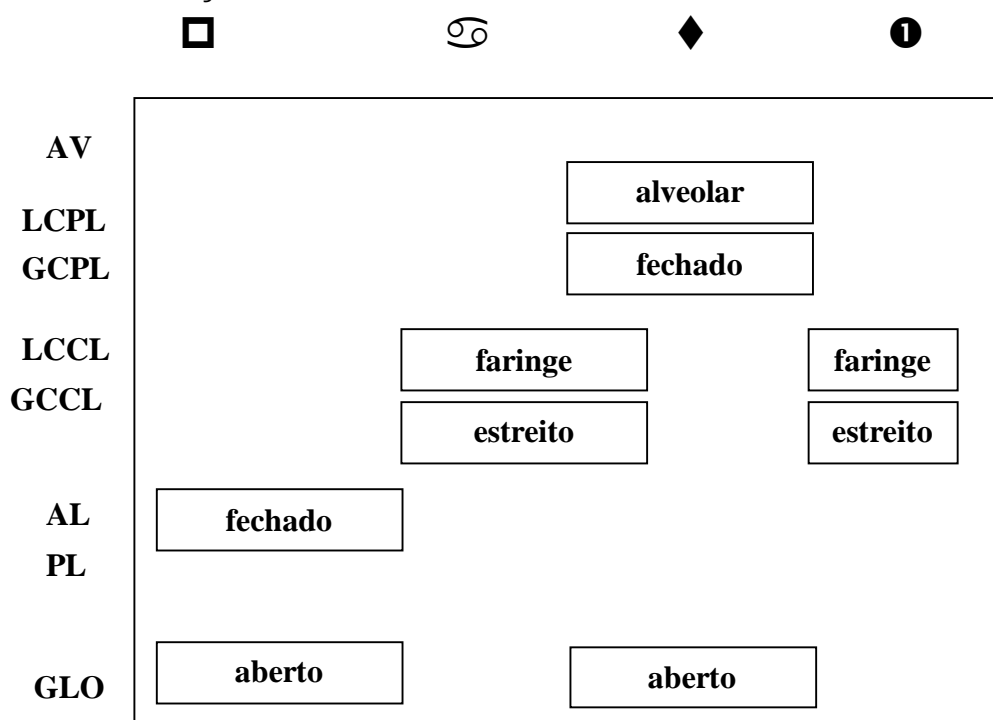
Usando como ponto de partida a interpretação dos resultados das medidas acústicas realizadas, idealizamos pautas gestuais correspondentes às produções do sujeito com deficiência auditiva e do sujeitos com audição normal, as quais discutiremos considerando os gestos articulatórios envolvidos na produção das palavras.

A evidente não-implementação do contraste de vozeamento nas produções do sujeito com deficiência auditiva foi representada nas pautas gestuais (figuras 61, 62, 63, 64, 65 e 66) das palavras que envolvem as plosivas vozeadas, com a introdução do gesto de abertura da glote, quando ele não deveria ocorrer. Se observarmos as pautas gestuais das produções do sujeito com deficiência auditiva, verificaremos que tanto para as produções das plosivas vozeadas como para as não-vozeadas encontramos o bloco relativo à variável do trato glote (Glo) indicando o gesto de abertura da glote²⁶. Tal gesto aparece em sincronia com o gesto de fechamento de

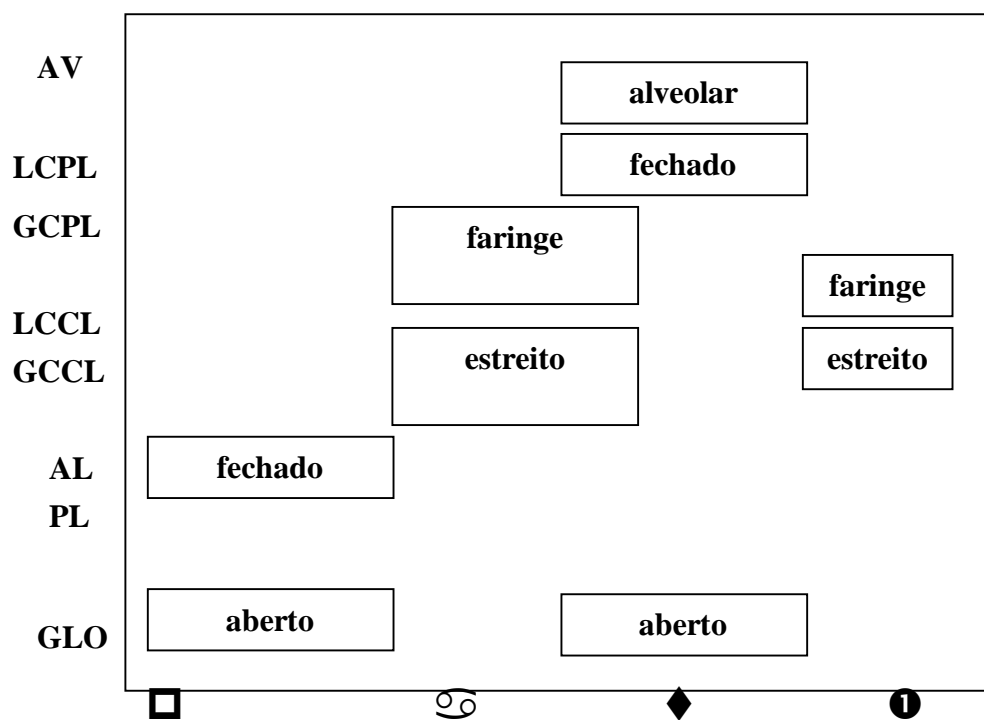
²⁶ Para a FAR, o modo de vibração das pregas vocais é considerado o *default* e na pauta gestual só é marcado o gesto de abertura da glote.

lábios no caso das bilabiais, de ponta da língua no caso das alveolares e do corpo da língua no caso das velares. A introdução do referido gesto resulta na produção de um som cujo percepto corresponde a uma plosiva não-vozeada.

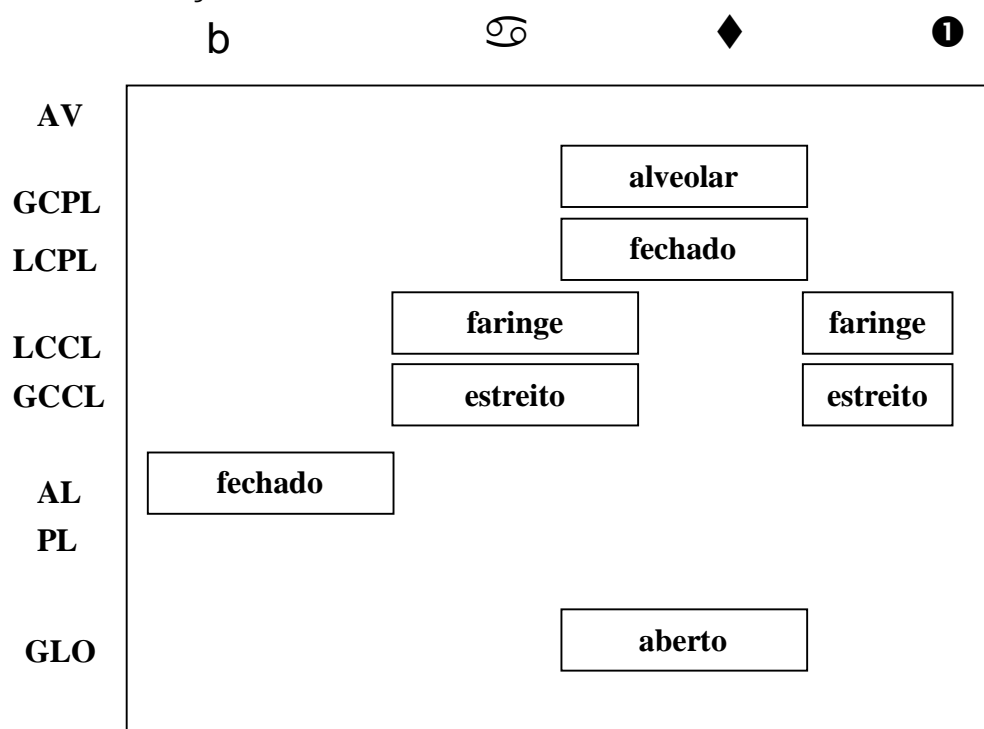
Sujeito com audição normal



Sujeito com deficiência auditiva



Sujeito com audição normal



Sujeito com deficiência auditiva

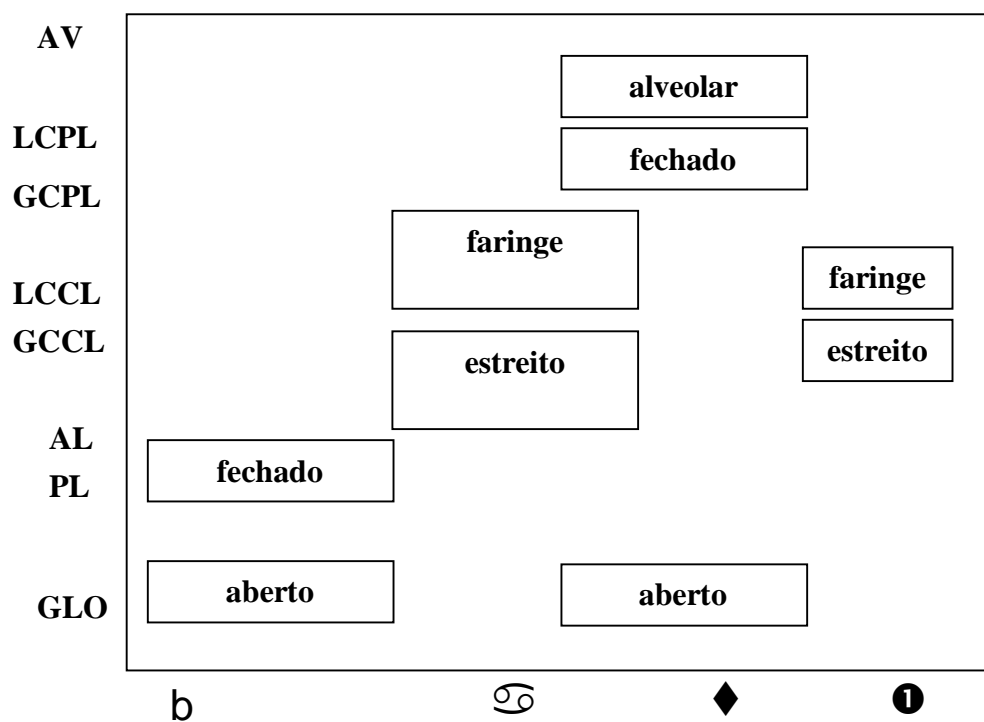
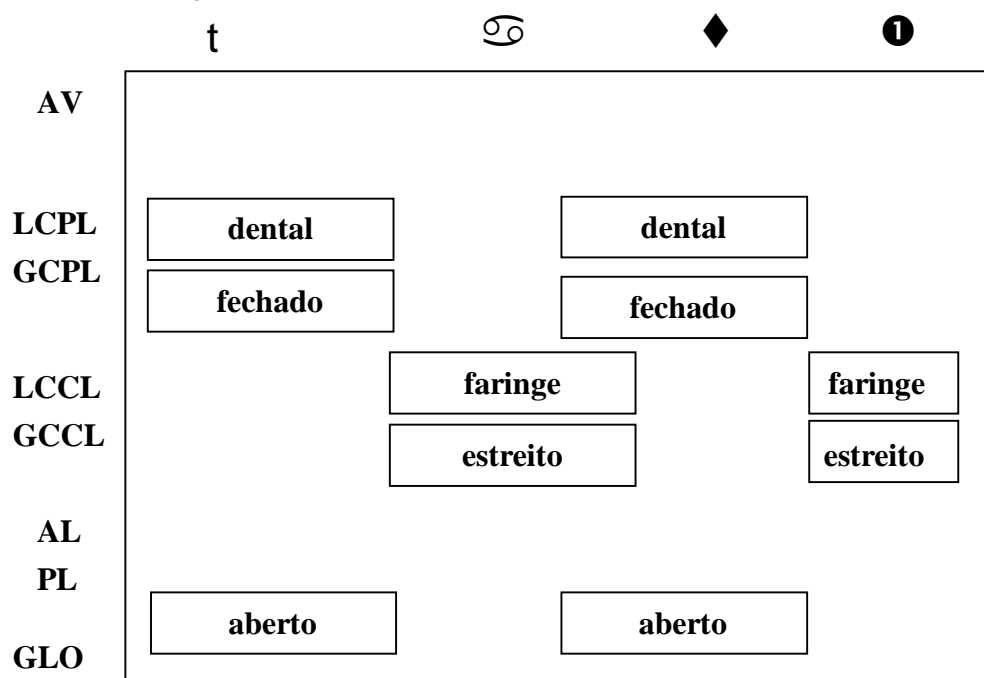


Figura 62 - Pautas gestuais referentes à palavra *bata* para os dois sujeitos

Sujeito com audição normal



Sujeito com deficiência auditiva

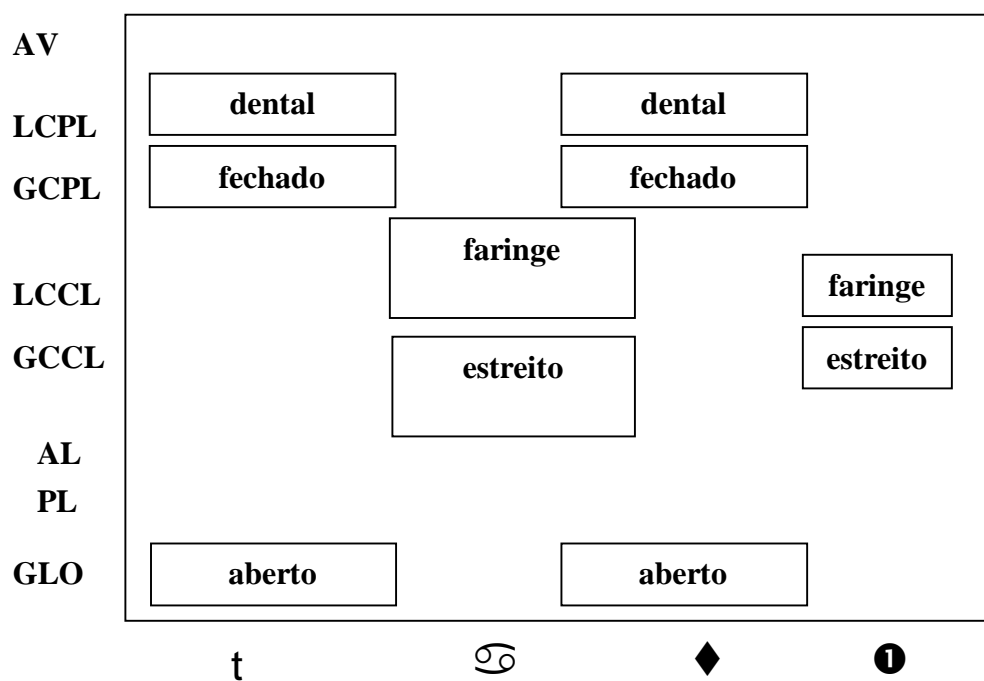
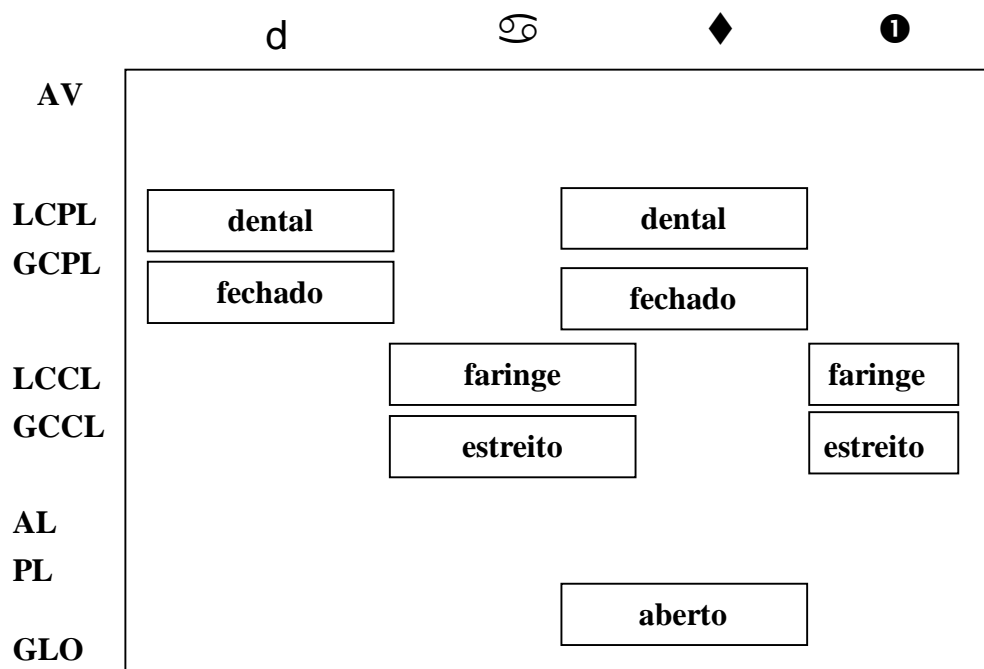


Figura 63 - Pautas gestuais referentes à palavra *tata* para os dois sujeitos

Sujeito com audição normal



Sujeito com deficiência auditiva

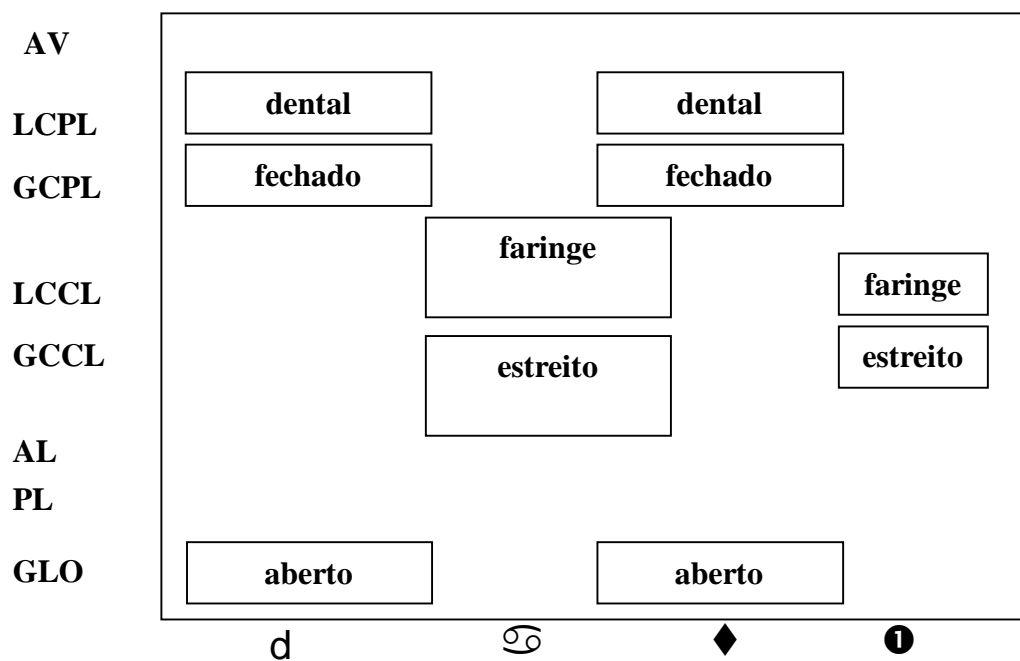
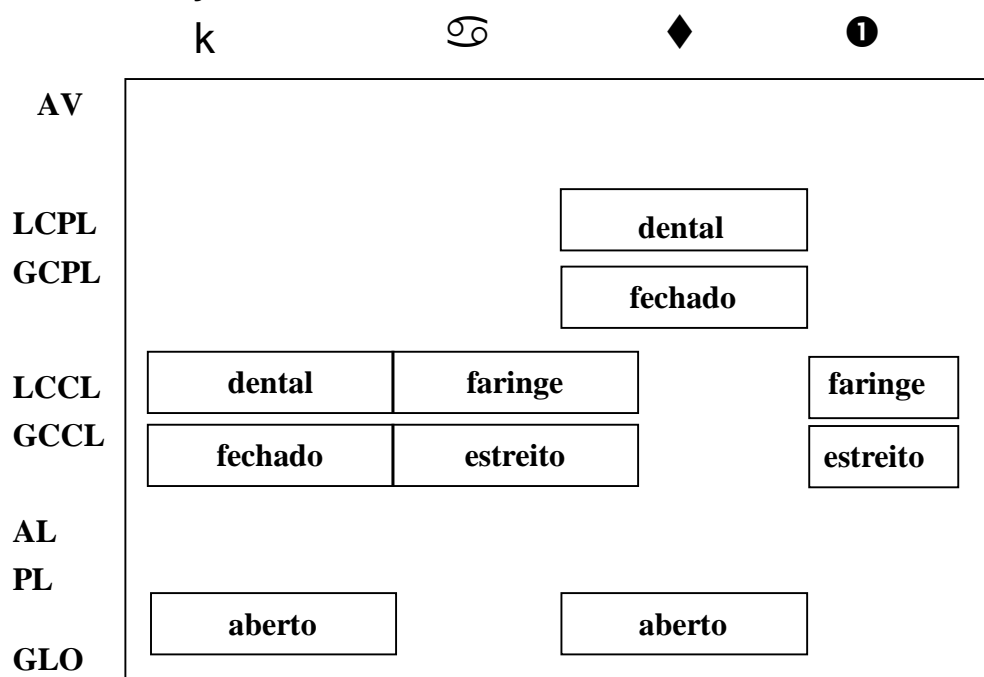


Figura 64 - Pautas gestuais referentes à palavra *data* para os dois sujeitos

Sujeito com audição normal



Sujeito com deficiência auditiva

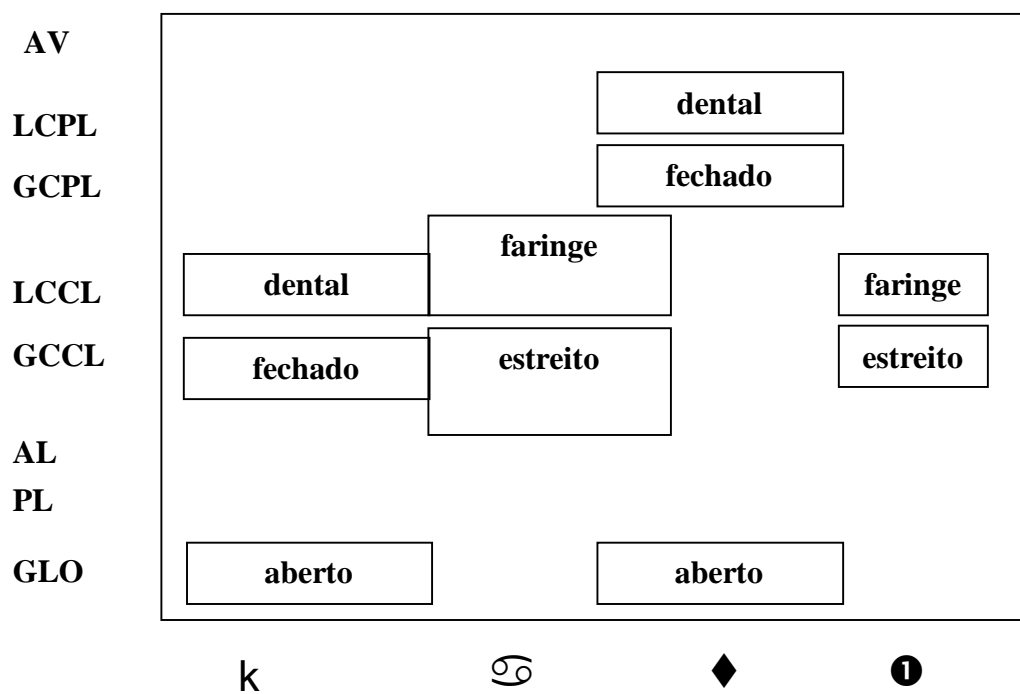
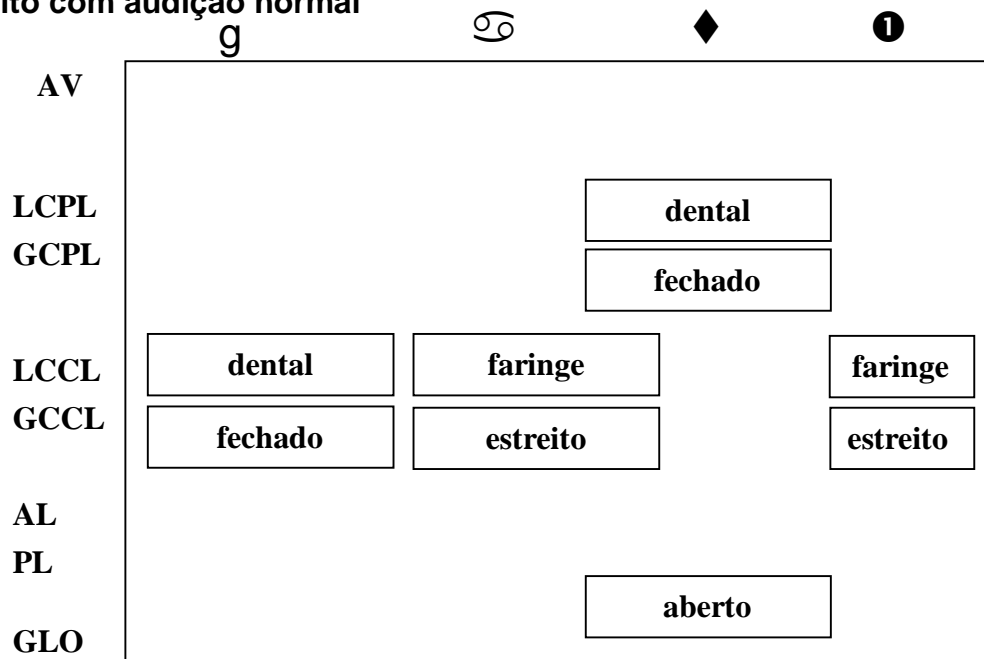


Figura 65 - Pautas gestuais referentes à palavra *cata* para os dois sujeitos

Sujeito com audição normal



Sujeito com deficiência auditiva

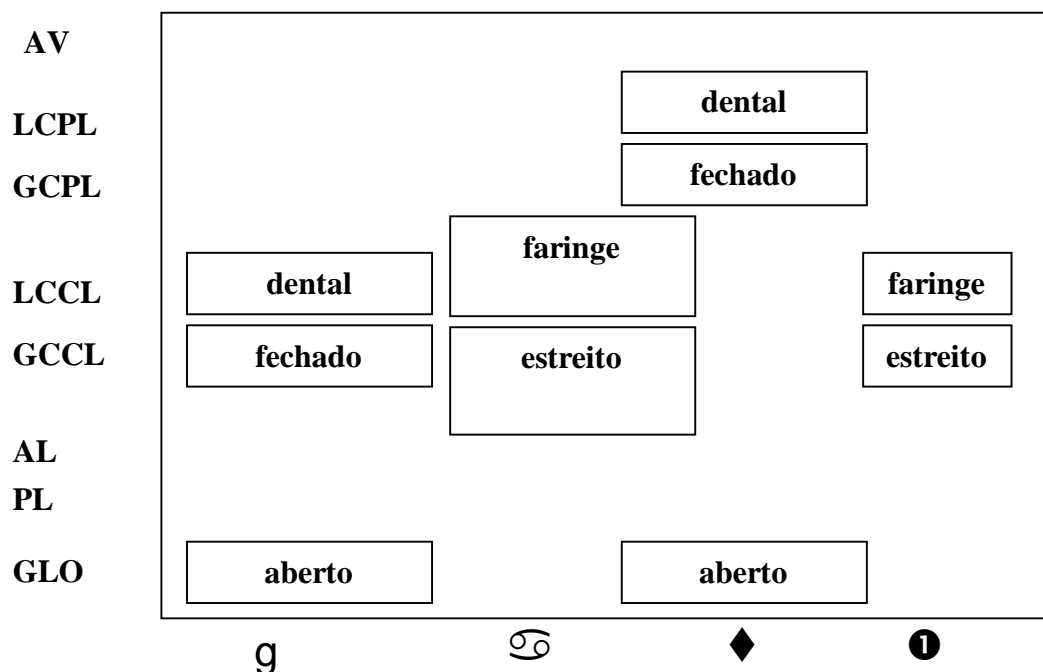


Figura 66 - Pautas gestuais referentes à palavra *gata* para os dois sujeitos

A pauta gestual no modelo da FAR refere-se ao *input*, ou seja, àquilo que corresponde à representação mental dos gestos articulatórios e que é determinado pelo componente lingüístico. A questão que emerge diz respeito à possibilidade de a ausência da informação acústica, a partir dos 18 meses de idade, justificar uma composição diversa na pauta gestual para esse sujeito com deficiência auditiva. Com base nos dados de produção e percepção anteriormente discutidos, e considerando a importância do elo acústico-articulatório apontado por Albano (2001), consideramos que essa é uma hipótese razoável, embora deva ser tomada apenas como hipótese inicial, que deverá, pois, ser objeto de outras pesquisas. Weismer *et al.* (1995) enfatizam que o fato de a FAR estar baseada na premissa de que o sistema fonológico pode apoiar-se diretamente na articulação torna o modelo bastante interessante para a explicação das alterações de fala decorrentes de patologias. Kent (1997a) sugere que as alterações encontradas na fala de pacientes com alterações neurológicas podem ser explicadas pelas alterações nas pautas gestuais.

Além do não-vozeamento das plosivas vozeadas nas produções do sujeito com deficiência auditiva, outro aspecto representado nas pautas gestuais construídas para representar suas produções - que merece destaque - é o aumento da magnitude dos gestos vocálicos na sílaba tônica que, por sua vez, resulta no aumento da duração no *output*²⁷. Não só o alongamento da duração encontrado na análise acústica das produções desse sujeito é sistemático, como também o primeiro formante está bastante elevado na vogal em posição tônica, se comparado com as posições não acentuadas. Assim, consideramos que ambos os fatos sugerem um movimento de mandíbula mais amplo, que é coerente com a representação do aumento da magnitude do gesto.

Refletir acerca dos resultados da análise acústica e das tarefas de percepção auditiva, considerando os pressupostos de um modelo fonológico que permite lidar com aspectos dinâmicos da fala, possibilitou-nos alcançar um poder explanatório mais amplo sobre a questão das alterações presentes na fala do sujeito com deficiência auditiva. Desse modo, não nos limitamos apenas a descrever as produções do sujeito com deficiência auditiva como caracterizadas pela troca de

²⁷ Barbosa PA. (2003) Comunicação pessoal.

sons vozeados por não-vozeados ou pela ausência do traço de sonoridade e nem a fazer referência a distorções na produção da plosiva alveolar. Fomos além, e passamos a considerar aspectos da dinâmica da fala. Desse modo, por meio do uso do modelo da FAR, uma explicação mais fortemente ancorada nos dados de produção, inferidos pela observação do dado acústico, foi possível.

Considerando a pauta gestual que propusemos, podemos supor que:

- a ocorrência do fechamento do trato nas produções do sujeito com deficiência auditiva pode ser correlacionada aos bons resultados em termos da percepção do modo de articulação, uma vez que a pista do silêncio é necessária para a percepção dos sons plosivos;
- o fato de a maior magnitude do gesto de dorso da língua poder ser representada na pauta gestual amplia a possibilidade explanatória em relação à especificidade da dinâmica de produção da fala implementada pelo sujeito com deficiência auditiva para a produção de uma plosiva seguida de vogal, pois um gesto (vocálico) especificado com maior magnitude precedido de gestos que envolvem fechamento completo como no caso das plosivas implica maior tensão tanto no nível laríngeo quanto no supra-laríngeo, o que dificulta a vibração das pregas vocais. Isso está de acordo com os dados de produção do sujeito com deficiência auditiva que, nas posições de ataque de sílaba e início de palavra, sistematicamente não implementou o vozeamento nas plosivas sonoras. A interação prosódia-segmento pode ser captada pela especificação de diferenças na magnitude do gesto;
- gestos que envolvem duas variáveis do trato associadas a um mesmo articulador trazem mais dificuldades de produção para esse sujeito com deficiência auditiva e provocam, conseqüentemente, mais dificuldades em nível perceptivo.

6 CONSIDERAÇÕES FINAIS

Com base nos dados da análise acústica e nos resultados da avaliação da percepção auditiva dos sons estudados, tomando como fundamentação uma abordagem dinâmica de produção de fala, foi-nos possível, nesta pesquisa, refletir sobre as alterações encontradas na produção das plosivas pelo sujeito com deficiência auditiva que foi selecionado para o estudo. Ainda que o material aqui apresentado e discutido não implique um *insight* imediato sobre possíveis estratégias terapêuticas que venham a modificar aspectos da produção do sujeito com deficiência auditiva, ele constitui, de alguma maneira, um subsídio para melhor compreender como esses sons de fala estão sendo produzidos.

A análise acústica utilizada como ferramenta para auxiliar a compreensão das alterações de fala de sujeitos com deficiência auditiva pode vir a contribuir para o avanço do processo terapêutico fonoaudiológico de pacientes com deficiência auditiva, propiciando a possibilidade de buscarmos novas alternativas terapêuticas. O conhecimento mais amplo dos processos de produção de fala, acrescido de dados a respeito de sua percepção, sem dúvida, é capaz ainda de oferecer subsídios para a seleção e o ajuste de recursos tecnológicos que visem a aumentar as pistas auditivas (e mesmo visuais ou táteis) que favoreçam a percepção da fala.

Nesse sentido, alguns dos parâmetros acústicos que aqui se mostraram relevantes para o conhecimento da produção das plosivas poderão ser utilizados para a avaliação e o acompanhamento da evolução de pacientes com deficiência auditiva bem como para a avaliação de procedimentos terapêuticos. Assim como as medidas de *VOT*, duração dos segmentos consonantais e vocálicos, transição de F2 e a frequência de F2 e F3 no início da vogal subsequente revelaram dados importantes para a compreensão dos processos de produção no caso estudado, seu uso tanto para realização de novas pesquisas como para acompanhamento terapêutico pode vir a ser de extrema valia, especialmente se considerarmos estudos de outros casos, envolvendo sujeitos com diferentes tipos de deficiência auditiva.

Também é preciso enfatizar que a análise acústica das produções do sujeito

com audição normal a que procedemos neste estudo, além de, evidentemente, termos fornecido parâmetros para avaliar os resultados encontrados, foi importante porque nos possibilitou conhecer aspectos acústicos dos sons de fala que foram utilizados no procedimento de avaliação da discriminação das plosivas do PB. Embora não tenhamos esgotado todas as medidas acústicas que poderiam ter sido realizadas, particularmente as referentes ao espectro do *burst*, aquelas que foram por nós utilizadas indicaram que, nas produções do sujeito com audição normal, há pistas acústicas suficientes para a identificação desses sons, o que pôde ser confirmado não apenas pela semelhança entre os dados acústicos revelados nesta pesquisa e os dados constantes na literatura específica, mas também pelos altos índices de acerto na tarefa de identificação.

É importante ainda ressaltar que a relação entre produção e percepção de fala configura-se como bastante relevante para a clínica fonoaudiológica, notadamente no que se refere ao atendimento de pacientes com deficiência auditiva, ainda que seja apenas um entre os inúmeros pontos a serem considerados no atendimento terapêutico dessas pessoas. Para os indivíduos que apresentam perdas de audição congênitas ou adquiridas na primeira infância – caso inclusive do sujeito que estudamos –, o processo terapêutico fonoaudiológico visa ao desenvolvimento da linguagem, que vai favorecer a participação dessa pessoa na sociedade: a possibilidade de falar de forma mais inteligível e compreender melhor a fala do outro, mesmo não sendo determinante, contribui sobremaneira para esse propósito.

Analisar os resultados da percepção dos sons plosivos do PB com dados de produção forneceu-nos subsídios para a compreensão do modo como a produção é afetada pela percepção para esse sujeito com deficiência auditiva, considerando as particularidades de nossa língua.

7 REFERÊNCIAS

Albano EC. O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro. Campinas (SP): Mercado das Letras; 2001.

Albano EC. Da fala à linguagem tocando de ouvido. São Paulo: Martins Fontes; 1990.

Albano EC et al. Segment frequency and word structure in Brazilian Portuguese. In: XIII International Congress of Phonetic Sciences; 1995; Stockholm. Proceedings. Stockholm: ICPhS; 1995. p.346-9.

Baken RS, Daniloff RG. Readings in clinical spectrography of speech. San Diego (US): Singular; 1991.

Balieiro CR, Ficker LB. Reabilitação aural: a clínica fonoaudiológica e o deficiente auditivo. In: Lopes Filho O, editor. Tratado de fonoaudiologia. São Paulo: Roca; 1997. p.311-25.

Barbosa PA. At least two macrorhythmic units are necessary for modeling brazilian portuguese duration: emphasis on automatic segmental duration generation. Cad Est Ling 1996; 31: 33-53.

Barbosa PA. Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração de ciência e tecnologia de fala. In: Scarpa E, organizador. Estudos de prosódia. Campinas (SP): Ed. Unicamp; 1999. p. 21-52.

Barbosa PA. Syllable-timing in Brazilian Portuguese: uma crítica a Roy Major. DELTA 2000; 16(2): 369-402.

Baum SR, Waldstein RS. Perseveratory coarticulation in the speech of profoundly hearing-impaired and normally hearing children. J Speech Hear Res 1991; 34:1286-92.

Behlau MS. Análise do tempo de início da sonorização na discriminação de sons plosivos do português [tese]. São Paulo: Escola Paulista de Medicina, UNIFESP; 1986.

Benki JR. Place of articulation and first formant transition pattern both affect perception of voicing in english. J Phonetics 2001; 29(1):1-22.

Bevilacqua MC, Formigoni GMP. Audiologia educacional: uma opção terapêutica para a criança deficiente auditiva. Carapicuíba (SP): Pró-Fono; 1997.

Bevilacqua MC, Tech EA. Elaboração de um procedimento de avaliação de percepção de fala em crianças deficientes auditivas profundas a partir de cinco anos de idade. In: Marchesan IQ, Zorzi JL, Gomes ICD, editores. Tópicos em fonoaudiologia. São Paulo: Lovise; 1996. p.411-33.

Blumstein SE, Stevens KN. Acoustics invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. J Acoustic Soc Am 1979; 66(4):1001-16.

Boothroyd A. Auditory perception of speech contrasts by subjects with sensorineural hearing loss. J Speech Hear Res 1984; 27: 134-44.

Boothroyd A. Evaluation of speech production of the hearing-impaired: some benefits of forced-choice testing. J Speech Hear Res 1985; 28: 185-96.

Boothroyd A. Hearing impairments in young children. Englewoods Cliffs (NJ): Prentice Hall; 1982.

Boothroyd A, Hanin L, Eran O. Speech perception and production in children with hearing impairment. In: Bess FH, Gravel JS, Tharpe AM, editors. Amplification for children with auditory deficits. Nashville (TN): Bill Wilkerson Center ; 1996.

Brito AT. Análise fonético-fonológica do estudo do vozeamento, desvozeamento e não vozeamento de obstruintes na fala infantil normal e com desvios [tese]. Belo Horizonte: Universidade Federal de Minas Gerais; 2000.

Browman CP, Goldstein LM. Articulatory gestures as phonological units. In: Phonology Yearbook, 6; 1989. p.201-51.

Browman CP, Goldstein LM. Articulatory phonology: an overview. Phonetica 1992; 155-80.

Browman CP, Goldstein LM. Tiers in articulatory phonology. In: Kingston J, Beckman M, editors. Between the grammar and physics of speech. Cambridge (UK): Cambridge University; 1990. [Papers in laboratory phonology; 1]

Browman CP, Goldstein LM. Towards an articulatory phonology. In: Phonology Yearbook, 3; 1986. p 219-52.

Camargo ZA. Análise da qualidade vocal de um grupo de indivíduos disfônicos: uma abordagem interpretativa e integrada de dados de natureza acústica, perceptiva e eletroglotográfica [tese]. São Paulo: Pontifícia Universidade Católica de São Paulo; 2002.

D'Angelis WR. Traços de modo e modos de traçar geometrias: línguas Macro-Jê & teoria fonológica [tese]. São Paulo: Universidade Estadual de Campinas; 1998.

Delgado EMC, Bevilacqua MC. Lista de palavras como procedimento de avaliação da percepção dos sons da fala para crianças deficientes auditivas. *Pró-fono* 1999;11(1): 59-64.

Dellatre P, Liberman AM, Cooper FS. Acoustic Loci and transitional cues for consonants. *J. Acoustic Soc Am* 1955; 27: 769-73.

Dorman MF, Hannley MT. Identification of speech and speechlike signals by hearing impaired listeners. In: Danilooff RG, editor. *Speech science: recent advances*. San Diego (CA): College Hill; 1985.

Dubno JR, Dirks DD, Ellison DE. Stop-consonant recognition for normal-hearing listeners and with high-frequency hearing loss: I – the contribution of selected frequency regions. *J Acoustic Soc Am* 1989; 85(1): 347-64.

Engstrand O, Krull D, Lindblom B. Sorting stops by place in acoustic space. In: XIIIth Swedish Phonetics Conference [online]; 2000; Sköde, Sweden. *Proceedings*. Sköde, Sweden: PERILUS, 2000: 53-6. [citado 2003 Jan 11]. Disponível em: http://www.ling.su.se/fon/perilus/2000_05.pdf (Fonetik 2000).

Erber NP. Speech perception by hearing-impaired children. In: Bess FH, Freeman BA, Sinclair JS, editors. *Amplification in education*. Washington (DC): Alexander Graham Bell Assoc Deaf; 1981. p.69-88.

Estabrooks W. *Auditory-verbal therapy for parents and professionals*. Washington (DC): Alexander Graham Bell Assoc Deaf; 1994

Ewing A, Ewing EC. *Educational guidance and the deaf child*. Manchester (US): University Press; 1971.

Fant G. *Acoustics theory of speech production*. Mouton: Hague; 1960.

Fant G. Stops in CV-syllables: speech analyses. Spectrographic characteristics of normal speech. San Diego (CA) : Singular; 1991. [Readings in clinical spectrography of speech]

Farnetani E, Recasens D. Coarticulation models in recent speech productions theories. In Hardcastle, Hewlett, editors. *Coarticulation: theory, data and techniques*. Cambridge (UK): Cambridge University; 2000.

Fowler C. Listeners do hear sounds, not tongues. *J Acoustic Soc Am* 1996; 99(3):1730-41.

Fowler C. The perception of phonetic gestures. In: Mattingly JG, Studdert-Kennedy M, editors. *Modularity and the motor theory of speech perception*. Hillsdale: Ed. Lawrence Erlbaum; 1991.

Freyman R, Nerbonne GP. The importance of consonant-vowel intensity ratio in the intelligibility of voiceless consonants. *J Speech Hear Res* 1989; 32: 524-35.

Gama MR. *Percepção da fala: uma proposta de avaliação qualitativa*. São Paulo: Pancast; c1994.

Gama-Rossi AJ. *Relações entre desenvolvimento lingüístico e neuromotor: a aquisição da duração no português brasileiro [tese]*. Campinas (SP): Universidade Estadual de Campinas; 1999.

Garcia LV, Pereira LD, Fukuda Y. Teste de contrastes: discriminação do traço de sonoridade com e sem ruído: proposta para avaliação audiológica. *Acta Awho* 1999; 18(3): 119-23.

Hall JW, Wood EJ. Stimulus duration and frequency discrimination for normal-hearing and hearing-impaired subjects. *J Speech Hear Res* 1984; 27: 252-59.

Halle M, Hughes GW, Radley JPA. Acoustic properties of stop consonants. San Diego (CA): Singular; 1991. p.216-25. [Readings in clinical spectography of speech, 1991].

Haskins Laboratories. Introduction to articulatory phonology and gestural computational model. [online]. New Haven (US): Haskins Lab; [1992?]. [citado 2003 jan 10]. Available from: <http://www.haskins.yale.edu/Haskins/MISC/RESEARCH/GesturalModel.html>

Hayward K. *Experimental phonetics*. Edinburgh (Eng): Pearson Education; 2000.

Holt LL, Lotto A, Kluender KR. Influence of fundamental frequency on stop-consonant voicing perception: a case of learned covariation or auditory enhancement. *J Acoustic Soc Am* 2001; 109(2): 764-74.

Johnson D, Whaley P, Droman MF. Processing of cues for stop voicing by young hearing-impaired listeners. *J Speech Hear Res* 1984; 27: 112-18.

Kent RD, Read C. *The acoustics analyses of speech*. San Diego (CA): Singular; 1992.

Kent RD. Gestural phonology: basic concepts and application in speech-language pathology, the new phonologies, developments. In: Ball M, Kent RD, editors. *Clinical linguistics*. San Diego (CA): Singular; 1997a

Kent RD. The speech sciences. San Diego (CA): Singular; 1997b.

Kewley-Port D. Measurement of formant transition in naturally produced stop consonant-vowel syllables. *J Acoustic Soc Am* 1982; 72(2): 379-89.

Kewley-Port D. Time varying feature as correlates of place of articulation in stop consonants. *J. Acoustic Soc* 1983; 73(1): 322-35.

Kewley-Port D, Pisoni DB, Studdert-Kennedy M. Perception of static and dynamic cues to place of articulation in initial stops consonants. *J Acoustic Soc Am* 1983; 73(5):1779-83.

Kluender KR. Effects of first formant onset properties on voicing judgments results from processes not specific to humans. *J Acoustic Soc Am* 1991; 90(1): 83-96.

Ladefoged P, Maddieson I. Stops in the sounds of the world's language. Oxford (UK): Blackwell; 1996. p.47-103.

Lieberman AM, Dellatre P, Cooper FS. The role of selected stimulus-variables in the perception of the unvoiced stop consonants. In: Liberman AM, editor. *Speech: a special code*. Cambridge (Mass): MIT Press; 1996. p.61-74.

Lieberman AM, Mattingly IG. The motor theory of speech perception revised. In: Liberman AM, editor. *Speech: a special code*. Cambridge (Mass): MIT Press; 1996. p.237-65.

Lindblom B. The status of phonetic gesture. Stockholm: PERILUS; 1990. p.21-39. [Perilus,11]

Ling D. The foundation of spoken language for hearing-impaired children. Washington (DC): Alexander Graham Bell Assoc Deaf; 1989.

Ling D, Ling AH. Aural habilitation: the foundation of verbal learning. Washington (DC): Alexander Graham Bell Assoc Deaf; 1978

Lisker L, Abramsom A. A cross language study of voicing in initial stops: acoustical measurements. *Word J Linguistic Circle NY* 1964; (20)3: 384-422.

Longone E, Borges ACC. Teste de reconhecimento de fala em indivíduos portadores de perda auditiva neurossensorial. *Pró-Fono* 1997; 9(1): 3-9.

Longone E, Borges ACC. Principais trocas articulatórias envolvidas na obtenção do índice percentual de reconhecimento de fala em indivíduos portadores de perda auditiva neurossensorial. *Acta Awho* 1998; 17(4): 186-92.

Madureira S, Barzaghi L, Mendes B. Voicing contrasts and the deaf: production and perception issues. In: Windsor F, Kelly ML, Hewlet N, editors. Themes in clinical phonetics and linguistics. London: Ed. Lawrence Erlbaum; 2002.

Massini G. A duração no estudo do acento e do ritmo do português brasileiro [tese]. Campinas, SP: Universidade Estadual de Campinas; 1991.

McGarr N, Löfqvist A. Laryngeal kinematics in voiceless obstruents produced by hearing-impaired speakers. J Speech Hear Res 1988; 31: 234-39.

McGarr N, Löfqvist A. Obstruent production by hearing-impaired speakers: interarticulator timing and acoustics. J Acoustic Soc Am 1982; 72(1): 34-42.

Mello ME, Novaes BCAC. Caderno de experiências no processo terapêutico de uma criança portadora de deficiência auditiva. Pró-Fono 2001; 13: 242-48.

Mencher GT. Early identification of hearing loss. Basel: S. Karger; 1976.

Mendel LL, Danhauer JL. Audiological evaluation and management and speech perception assessment. San Diego (CA): Singular; 1997.

Miller GA, Nicely PE. Na analyses of perceptual confusion among some english consonants. J Acoustic Soc Am 1955; 27: 338-46

Monsen RB. The production of English stop consonants in the speech of deaf children. J Phonetics 1976; 4: 29-41.

Ochs MT, Humes LE, Ohde R, Grantham DW. Frequency discrimination ability and stop-consonant identification in normally hearing and hearing-impaired subjects. J Speech Hear Res 1989; 32: 133-42.

Ohala JJ. Phonetic explanation for sound patterns: implication for grammars of competence. In: XIII International Congress of Phonetic Sciences; 1995; Stockholm. Proceedings. Stockholm: ICPhS; 1995. p.52-9.

Ohala JJ. Speech perception is hearing sounds, not tongues. J Acoust Soc Am 1996; 99 (3): 1718-25.

Ohde RN, Stevens KN. Effect of burst amplitude on the perception of stop consonant place of articulation. J Acoustic Soc Am 1983; 74: 706-14.

Öhman SEG. Coarticulation in VCV utterances: spectrographic measurements. J Acoustic Soc Am 1966; 39:151-8.

Okalidou A, Harris KS. A comparison of intergestural patterns in deaf and hearing adult speakers: implications from an acoustic analysis of disyllables. *J Acoustic Soc Am* 1999; 106(1): 349-410.

Orlandi ACL. Procedimento de avaliação de percepção da fala em deficiência auditiva profunda nos primeiros anos de vida [tese]. São Paulo: Pontifícia Universidade Católica de São Paulo; 1996.

Osberger MJ, Levitt H. The effect of timing errors on the intelligibility of deaf children's speech. *J Acoustic Soc Am* 1979; 66(5):1316-24.

O'Shaughnessy D. Speech perception: acoustic or articulatory. *J Acoustic Soc Am* 1996; 99(3): 1726-9

Parady S, Dorman MF, Whaley P, Raphael LJ. Identification and discrimination of a synthesized voicing contrast by normal and sensorineural hearing-impaired children. *J Acoustic Soc Am* 1981; 69(3): 783-90.

Penrod JP. Speech threshold and word recognition/discrimination testing. In: Katz J, editor. *Handbook of clinical audiology*. Baltimore (Maryland): Williams and Wilkins; 1994

Peterson GE, Lehiste I. Duration of syllable nuclei in english. In: Baken RJ, Daniloff RG, editors. *San Diego (CA): Singular; 1991. p.134-44. [Readings in clinical spectrograph of speech]*

Picheny MA, Durlach NI, Braida LO. Speaking clearly for the hard of hearing ii: acoustic characteristics of clear and conversational speech. *J Speech Hear Res* 1986; 29(4): 434-45.

Pickett JM, Danaher EM. On discrimination of formant transitions by persons with sever sensorineural hearing loss In: Fant G, Tatham MAA., editors. *Auditory analysis and perception of speech*. London: Academic Press; 1975.

Pickett JM, Revoile SG, Holden LD A speech production approach to speech perception by deaf persons. In: Hochber, Levitt, Osberger, editors. *Speech of the hearing impaired: research, training and personnel preparation*. Baltimore (Maryland):University Park Press; 1983

Pollack D. *Educational audiology for the limited hearing infant and preschooler*. 2nd ed. Illinois (US): Charles C Thomas; 1995.

Pupo AC. Alguns aspectos do processo de discriminação auditiva de sons da fala em crianças [tese]. São Paulo: Pontifícia Universidade Católica de São Paulo; 1981.

Revoile SG. Hearing loss and the audibility of phoneme cues. In: Pickett, editor. The acoustics of speech communication. Boston (US): Allyn and Bacon.;1999

Revoile SG, Pickett JM, Holden LD. Acoustic cues to final stop voicing for impaired- and normal-hearing listeners. J Acoustic Soc Am 1982; 72(4): 1146-54.

Revoile SG, Pickett JM, Kosma-Spytek L. Spectral Analyses of /d/, /n/,/l/ by normal and hearing impaired listeners. J Acoustic Soc Am 1991; 9(2): 787-98.

Rodrigues N. Discriminação auditiva: normas para avaliação de crianças de 5 a 9 anos. São Paulo: Cortez; 1981.

Saltzman E L, Munhall KG. A dynamical approach to gestural patterning in speech production. Echol Psychol 1989; 1(4): 333-82.

Saltzman E L, Kelson JAS. Skilled actions: a task dynamic approach. Psychol Rev 1987; 94:84-106.

Sammeth CA, Dorman MF, Stearns CG. The role of consonant-vowel amplitude ratio in the recognition of voiceless stop consonants by listeners with hearing impairment. J Speech Hear Res 1999; 42: 42-55.

Sammeth CA, Tetzli M, Ochs MT. Consonant recognition performance of hearing impaired listeners using one linear and three non linear hearing aids. J Am Acad Audiol 1996; 7: 240-50.

Sanders DA. Auditory perception of speech: an introduction to principals and problems. Englewoods Cliffs (NJ): Prentice Hall; 1977.

Shimizu K. A cross-language study of voicing contrasts of stops consonants in Asian Languages. Tokyo: Seibido; 1996.

Silman S, Silverman CA. Auditory diagnosis: principles and applications. San Diego (CA): Singular; 1997

Smits R, Bosch L, Collier R. Evaluation of various sets of acoustics cues for the perception of prevocalic stop consonants: perception experiment I. J Am Acad Audiol 1996; 100(6): 3852-64.

Stevens KN. Acoustic phonetics. Cambridge (Mass): MIT Press, 1998

Stevens KN. On the quantal nature of speech. J Phonetics 1989;17: 3-45.

Stevens KN. On the quantal nature of speech: evidences from articulatory-acoustic data. In: David Jr EE, Denes PB, editors. Human communication: a unified view. New York: McGraw Hill; 1972. 51-65.

Stevens KN. The speech signal. In: Kavanagh J, Strange W, editors. Speech and language in the laboratory, school and clinic. Cambridge (Mass): MIT Press; 1978

Stevens KN, Blumstein SE. Invariant cues for place of articulation in stop consonants. J Acoustic Soc Am 1978; 64(5): 1358-68.

Stevens KN. Articulatory–acoustic–auditory relationships. In: Hardcastle WJ, Laver J, editor. The handbook of phonetic science. Oxford (UK): Blackwell; 1997. p.462-506.

Sussman HM, Mccaffrey HA, Matthews SA. Na investigation of locus equations as a source of relational invariance for stop consonants. J Acoustic Soc Am 1991; 90(3): 1309-25.

Sussman HM. Neural coding of relational invariance in speech: human language analogs to barn owl. Psychol Rev 1989; 96(4): 631-42.

Syrdal AK, Gopal HS. A perceptual model of vowel recognition based on the auditory representation of american english vowel. J Acoustic Soc Am 1986; 79(4):1086-100.

Tatham M. Introduction to articulatory phonology and a computational gestural model. In: 1st ESCA Tutorial and Research Workshop on Speech Production Modeling: from Control Strategies to Acoustics, and 4th Speech Production Seminar: models and data [online]; 1996 May 20-24; Autrans. Proceedings. Grenoble: Institut de la Communication Parlée, Université Stendhal; 1996. [citado 2003 fev 07]. Disponível em: <http://www.essex.ac.uk/speech/pubs/presents/aut-96/aut-96.html>

Tsui IY, Ciocca V. Perception of aspiration and place of articulation of Cantonese initial stops by normal and sensorineural hearing-impaired listeners. Int J Lang Commun Disord 2000; 35(4): 507-25.

Turner CW, Brus S. Providing low-and mid-frequency speech information to listeners with sensorineural hearing loss. J Acoustic Soc Am 2001; 109(6): 2999-3006

Tye-Murray N. Effects of vowel contest on the articulatory closure postures of deaf speakers. J Speech Hear Res 1987; 30: 99-104.

Tye-Murray N. The establishment of open articulatory postures by deaf and hearing talkers. J Speech Hear Res 1991; 34: 453-9.

Tyler RS, Moore BC. Consonant recognition by some of the better cochlear implant patients. J Acoust Soc Am 1992; 92(6): 3068-77.

Van Tansell DG, Hagen LT, Koblas LL, Penner SG. Perception of short-term spectral cues for stop consonant place by normal and hearing-impaired subjects. *J Acoust Soc Am* 1982; 72(6): 1771-80.

Veloso J. The role of consonantal duration and tenseness in the perception of voicing distinctions of Portuguese stops. In: XIII International Congress of Phonetic Sciences; 1995; Stockholm. Proceedings. Stockholm: ICPHS; 1995. p.266-69.

Waldstein RS, Baum SR. Anticipatory coarticulation in the speech of profoundly hearing-impaired and normally hearing children. *J Speech Hear Res* 1991; 34:1276-85.

Weismer G, Tjaden K, Kent RD. Can articulatory speech disorders be accounted for by theories of normal speech production?. *J Phonetics* 1995; 23:149-64.

Werner L, Marean GC. Human auditory development. Dubuque (IA): Brown and Benchmarkers; 1996. [Developmental psychology series]

Whalen DH. Three lines of evidence for direct links between production and perception in speech. In: XIV. International Congress of Phonetic Sciences; 1999; San Francisco (US). Proceedings. San Francisco (US): ICPHS; 1999.

Wieringen A, Pols LCW. Frequency and duration discrimination of short-first formant speech-like transitions. *J Acoust Soc Am* 1994; 95(1): 502-11.