

# Percepção da fala em bebês no primeiro ano de vida

**Rosana Maria Tristão**

*Universidade de Brasília / Secretaria de Educação do Distrito Federal*

**Maria Angela Guimarães Feitosa**

*Universidade de Brasília*

## Resumo

A fala humana é um som de grande complexidade, cujo processamento perceptual, produção e relações com a linguagem e a cognição necessitam de uma análise integrada, tanto do ponto de vista do conhecimento disponível como também das especificidades metodológicas. Neste artigo faz-se uma breve revisão da literatura sobre as principais aquisições e desenvolvimento da linguagem no primeiro ano de vida de bebês com desenvolvimento normal com enfoque na percepção da fala humana. Busca-se, também, analisar a ocorrência de distúrbios auditivos que podem causar alterações na percepção da fala, com possíveis implicações para o desenvolvimento pré-lingüístico. Atenção especial é dada ao desenvolvimento da habilidade de percepção de fala e de linguagem em bebês com síndrome de Down. É analisada a predisposição, nesta população, a problemas audiológicos, sua relação com alterações no desenvolvimento de linguagem, e a tendência apresentada no primeiro ano de vida para padrões diferenciados de atenção à fala.

Palavras chave: percepção de fala; desenvolvimento; bebês; linguagem; síndrome de down

## Abstract

*Speech perception in infants in their first year of life.* Human speech is a highly complex sound; whose perceptual processing, production and relations to language and cognition require an integrated analysis, not only from the viewpoint of available knowledge but also of its methodological specificities. This article presents a brief review of the literature on the main acquisitions and development of language in the first year of life of normally developing infants, with emphasis on speech perception. One also analyzes the occurrence of auditory disturbances in the first year of life that could jeopardize speech perception, with possible implications for pre-linguistic development. Special attention is given to the development of speech perception and language in Down syndrome infants. The predisposition to audiological problems, its relation to impairment in the development of language, and the tendency presented in the first year of life of differential patterns of attention to speech is analyzed in this population.

Key words: speech perception; development; infants; language; down syndrome

**E**ste trabalho tem o objetivo de sistematizar brevemente o conhecimento disponível na literatura internacional sobre o desenvolvimento da habilidade de percepção de fala em bebês humanos no primeiro ano de vida em condição de desenvolvimento normal, como parte de uma estratégia para avaliar necessidades e prioridades no avanço do conhecimento acerca de bebês em condições diferenciadas de desenvolvimento, em especial o desenvolvimento de percepção de fala naqueles que nascem com a síndrome de Down. Neste período fundamental para o desenvolvimento da linguagem ocorre o desenvolvimento de habilidades pré-lingüísticas. Contudo, a inclusão de bebês do primeiro ano de vida é relativamente recente dentro do grande corpo teórico tanto de estudos sobre a linguagem quanto de estudos na área de psicofísica acústica. São

analisados os principais achados disponíveis na literatura sobre o desenvolvimento da linguagem e da habilidade de percepção de fala no primeiro ano de vida, enfocando a percepção de categorias fonéticas. Espera-se, com isto, chamar a atenção para a necessidade de estudos específicos, necessários para uma compreensão mais satisfatória de características importantes desta síndrome.

## Percepção de fala no primeiro ano de vida

A linguagem se inicia muito precocemente na vida humana com a aquisição de padrões de sons da fala (Pinker, 1995). As principais conquistas lingüísticas durante o primeiro ano de vida são o controle da musculatura da fala, a sensibilidade às distinções fonéticas usadas na fala de seus

pais e a sensibilidade aos marcadores prosódicos que indicam a entonação presente (Jusczyk, Houston, & Goodman, 1998). Os bebês atingem estes marcos antes de produzirem ou compreenderem palavras, sugerindo que esta aprendizagem não depende do estabelecimento da relação entre som e significado. Ou seja, eles não ouvem diferencialmente duas palavras que soam semelhantes porque compreendem seus significados. No modelo proposto por Pinker, eles provavelmente classificam os sons diretamente, de alguma forma sintonizando seu módulo de análise da fala para separar os fonemas<sup>1</sup> usados em sua língua. O módulo funcionaria como linha de frente do sistema que aprende palavras e gramática.

Por volta dos seis meses começam a surgir os balbucios, que se desenvolvem até os nove meses. Nesta época, caso o bebê tenha um *feedback* auditivo adequado do balbucio, há a continuação do caminho evolutivo da fala; caso haja ocorrência de surdez, este balbucio cessa, em uma clara evidência da importância do estabelecimento de uma conexão entre balbucio e audição (MacWhinney, 1998). Esta relação entre balbucio e audição pode ser prejudicada até mesmo quando ocorrem apenas períodos de perda auditiva leve ou moderada em decorrência de otite média. Os bebês já aos seis meses de idade mostram-se atentos às características prosódicas da fala que ouvem e aos oito meses demonstram interesse por padrões sonoros mais longos, como seqüências silábicas (Eimas, 1985).

O sinal da fala é um complexo de unidades acústicas com segmentos breves separados por pausas momentâneas ou picos em intensidade. Os segmentos variam em duração e em composição de freqüências, organização temporal e distribuição de energia dos formantes<sup>2</sup>. A fala inclui também componentes acústicos com características de ruído, correspondendo a aspirações e fricções que são produzidas pela passagem do ar pelo trato vocal. A variação nestes parâmetros acústicos provê a informação relevante para a percepção de fonemas (Eimas, 1985; Kuhl, 1987).

Não se identifica correspondência estreita entre os segmentos acústicos individuais e os fonemas percebidos. Um único segmento acústico pode compreender uma consoante e uma vogal; dois segmentos acústicos distintos podem contribuir para um único som consonantal. Além disso, não é direta a relação entre os segmentos de freqüência e características temporais das unidades acústicas e os fonemas ouvidos. Um ouvinte pode reconhecer estímulos que variam amplamente em número de traços acústicos como exemplos do mesmo fonema; já uma pequena alteração em uma única pista acústica pode, em algumas situações, alterar o fonema percebido (Eimas, 1985).

Uma das variáveis que ajudam a definir um fonema e a distinguir palavras faladas é o tempo de surgimento da voz (Cleary e Pisoni, 2001; Eimas, 1985), também identificado na literatura como VOT (*Voice Onset Time*)<sup>3</sup>. Nenhum valor específico de VOT define cada fonema, embora diferenças entre VOTs sejam indícios relevantes para a discriminação entre alguns sons consonantais. Pelo contrário, ouvintes tipicamente percebem uma ampla variedade de valores,

refletindo diferentes falantes, diferentes exemplos de fala e diferenças no ambiente fonêmico circundante, como exemplos do mesmo fonema. Outras variáveis também caracterizam um fonema. Um exemplo é o lugar de articulação, que é o lugar onde ocorre constrição do trato vocal para que o som seja formado. Entre as pistas acústicas que correspondem ao lugar de articulação e permitem ao ouvinte distinguir fonemas estão as freqüências iniciais do segundo e terceiro formantes. No entanto, um conjunto de freqüências pode sinalizar o mesmo lugar de articulação (Eimas, 1985).

Na fala fluente, uma interação entre fatores temporais e espectrais governam a percepção de distinções vocais. Estas propriedades acústicas interagem de forma que podem ser chamadas de relações de intercâmbio: uma alteração no valor de uma propriedade altera o valor de outra na qual os limites perceptuais ocorrem. O sistema perceptual parece ser sintonizado para a relação de compensação, na qual uma alteração de freqüência pode substituir uma alteração na pista temporal. Assim, por exemplo, quando o primeiro formante começa com uma freqüência mais alta, o efeito é o mesmo, como se o VOT tivesse sido prolongado (Eimas, 1985).

O conhecimento acumulado sobre percepção de fala ressalta também o desenvolvimento como importante aspecto nesta competência perceptual (Montgomery & Clarkson, 1997; Ohde, Haley, & McMahan, 1996). Já ao nascimento o bebê possui capacidade de detectar contrastes consonantais, por exemplo, sendo esta habilidade não apenas inata, mas característica do sistema auditivo dos mamíferos. Durante o primeiro ano de vida a acuidade desta habilidade declina para fonemas não pertencentes ao contexto lingüístico da criança, e ela passa a discriminar seletivamente os fonemas da língua a que é exposta, fazendo uma reorganização perceptual, em um claro processo de adaptação à solicitação do meio em que vive (Jenkins, 1992; Kuhl, 1987, 1993). Nesta adaptação está incluído um processo de agrupamento de fonemas inicialmente diferentes, que passam a ser percebidos como sendo equivalentes, um fenômeno conhecido na literatura como percepção categórica.

O entendimento sobre a percepção da fala está ancorado em um corpo teórico diferenciado, embora ainda insatisfatório, que tem sido objeto de revisões sistemáticas (por exemplo, Cleary & Pisoni, 2001). Uma abordagem central, conhecida como a teoria motora, é representada pelo trabalho de A. M. Liberman. Esta teoria foi inicialmente impulsionada pelo desconforto com a variância acústica no sinal da fala, e propõe que o ouvinte faz uso de seus padrões internalizados da atividade motora de articulação para interpretar a fala. A teoria é contemporaneamente criticada pela variância no gesto motor articulatório para explicar satisfatoriamente a percepção de fala. O gesto articulatório, à semelhança do padrão acústico, também é instável demais para explicar como sons da fala são percebidos como iguais ou equivalentes. Mantém-se nas versões modernas da teoria motora a hipótese de que o papel do percebido de fala, como também produtor de fala, deve contribuir com informação relevante, não na forma de atos motores, mas na forma de intenção de gestos motores (Cleary & Pisoni, 2001)

Uma segunda abordagem, conhecida como realista direta, é impulsionada pela preocupação com a desconsideração da função ecológica da fala na exploração das particularidades da variação acústica na fala. Esta abordagem, representada pelo trabalho de C. A. Fowler, pressupõe que os mecanismos que explicam a percepção de fala podem ter um *status* geral, não necessariamente exclusivo do ser humano, que enfatiza o aspecto “público” da percepção, e descarta a necessidade de se incorporar a uma teoria de percepção de fala representações mentais de níveis de processamento (Cleary & Pisoni, 2001).

Uma terceira abordagem, mais recente, é conhecida como modelo de percepção de lógica nebulosa, e pode ser representada pelo trabalho de D. W. Massaro (Cleary & Pisoni, 2001; Nearey, 1997). Esta teoria procura modelar a influência integrada de variáveis visuais e auditivas para a percepção de fala e é criticada por ser geral demais em sua atual formulação.

Apesar destes modelos ainda se mostrarem insatisfatórios e diferirem em vários aspectos, têm sido feitas tentativas de encontrar seus possíveis pontos em comum. Kuhl (1987) identifica concordância na suposição de que existem níveis distintos de processamento envolvidos na transformação de um sinal acústico em mensagem percebida. Os níveis de processamento apontados são: (a) auditivo, (b) fonético, (c) fonológico, e (d) lingüístico de alto-nível. O **processamento auditivo** envolve a transformação de uma onda acústica em um sinal neuralmente codificado. O **processamento fonético** envolve o processamento destes sinais em perceptos específicos da fala. Sua proposição deriva da observação de que não existe uma relação um a um entre sinais acústicos específicos e seus perceptos. No **processamento fonológico** as unidades fonéticas ou segmentos são convertidos em grupos língua-específicos, ou fonemas. É importante enfatizar que fonemas são grupamentos abstratos cujos membros são funcionalmente equivalentes podendo ser confundidos ou separados, conforme a influência cultural (por exemplo, na língua japonesa /r/ e /l/ são ambas derivações de um único fonema). Finalmente, o **processamento lingüístico de alto nível** inclui todos os processamentos de nível superior, como os que envolvem busca léxica, análise sintática, e interpretação semântica.

Um conjunto também diferenciado de metodologias vem sendo empregado para o estudo experimental da percepção de fala no bebê. O desenvolvimento de metodologias apropriadas, em especial as metodologias psicofísicas, tem exigido considerável investimento de pesquisa de distintos laboratórios, não sendo viável incluir sua análise no presente trabalho. O leitor pode encontrar análises metodológicas interessantes em Slater (1998), Tristão (2001) e Werner e Rubel (1992).

### Principais achados de pesquisa sobre a percepção de fala por bebês

A despeito de que ainda não se dispõe de uma teoria unificadora, e do fato de que as metodologias disponíveis ainda não dão conta da riqueza da área, o volume de pesquisas

empíricas de qualidade é grande. A revisão dos estudos empíricos contida no presente trabalho procura privilegiar o entendimento do aspecto desenvolvimental.

Bebês muito jovens respondem a estímulos da fala. Evidências para preferência por estímulos de fala são demonstradas mesmo em estágio fetal. O sistema auditivo humano é funcional após 22 a 24 semanas gestacionais, e evolui durante o último trimestre de gestação em um ambiente com presença de sons para os quais o feto responde comportamental, elétrica e neuroquimicamente (Lecanuet, 1998).

### Estudos com fonemas

Em estudos realizados com fonemas, bebês pré-lingüísticos são capazes de detectar diferenças entre sons da fala que definem as classes fonéticas nas línguas de todo o mundo, ou seja, os bebês são percebedores universais (Jenkins, 1992; Kuhl, 1987). Por volta de um ano de idade, eles concluem o processo de especialização de modo a discriminarem apenas os fonemas de sua língua nativa. Nesta visão, os bebês desenvolveriam classes de equivalência que servem para modelar alguns dos limites naturais e suprimir outros que interferem com as equivalências, ou que sejam irrelevantes.

Há três casos relevantes de detecção de equivalência: (a) bebês detectam similaridades entre eventos auditivos discriminavelmente diferentes que sinalizam a mesma unidade fonética (classificação por equivalência), demonstrando a detecção de “emparelhamentos auditivos complexos”, ou classes de equivalência; (b) bebês relacionam informação da fala apresentada ao olho com a apresentada ao ouvido (percepção da fala entre modalidades sensoriais), demonstrando conhecimento de equivalência de classes auditivo-visuais; (c) eles emparelham (comparam) as produções de outros com as suas próprias (imitação vocal), provendo a evidência de conhecimento de equivalência de classes auditivo-motoras (Kuhl, 1987).

Por meio da experiência lingüística, a categorização nos termos abstratos destas classes de equivalências é intensamente praticada e atinge o estado de automaticidade (Jenkins, 1992). É importante apreciar que este aprendizado inicial tipicamente ocorre em condições adversas, o que torna ainda mais surpreendente sua qualidade e velocidade, uma vez que as condições ambientais em cujo contexto ocorre a comunicação oral incluem a participação de muitos outros estímulos, inclusive sonoros. Nozza, Rossman, Bond e Miller (1990) encontraram que a habilidade de bebês de sete a onze meses em discriminar dois sons da fala lingüisticamente relevantes em contexto de ruído é inferior a de adultos, sugerindo que o efeito do ruído seria maior ainda em tarefas mais complexas de reconhecimento de sons com significado, como ocorre na fala fluente.

### Estudos com sílabas e outros componentes de sons complexos

Fetos entre a 36<sup>a</sup> e a 37<sup>a</sup> semana gestacional demonstraram perceber a diferença entre duas notas musicais de frequência

baixa (ré na quarta oitava e dó na quinta oitava) de um piano observando-se a alteração da taxa de batimento cardíaco diante da apresentação de um estímulo novo após a apresentação de um primeiro estímulo (Lecanuet, Graniere-Deferre, Jacquet, & DeCasper, 2000). Estudos com este paradigma têm sugerido que fetos no terceiro trimestre podem se familiarizar com sons recorrentes da fala materna (DeCasper, Lecanuet, Busnel, Graniere-Deferre, & Maugeais, 1994).

A sensibilidade à entonação e prosódia parece estar presente no primeiro ano de vida. Recém-nascidos percebem diferencialmente vogais fracas e reduzidas e vogais fortes e cheias em sílabas, indicando que a sensibilidade para a diferença entre sílabas tônicas e átonas desenvolve-se precocemente (van Ooijen, Bertoncini, Sansavini, & Mehler, 1997). Bebês de três dias de idade preferiam ouvir histórias que suas mães haviam lido em voz alta ao longo de seis semanas antes do nascimento do que histórias que eles nunca tinham ouvido. A explicação desta preferência tão precoce pressupõe a ocorrência de aprendizagem pré-natal de algum aspecto do sinal acústico, provavelmente a prosódia. Os bebês também preferiam ouvir a língua materna a uma língua estrangeira (Lecanuet, 1998). Em um estudo de DeCasper e Fifer (1980), os recém-nascidos demonstraram clara preferência pela voz materna.

Estudos sobre desempenho de bebês em tarefas de discriminação de frequências durante o primeiro ano de vida sugerem uma maturação continuada do sistema auditivo, fazendo com que bebês entre quatro e oito meses sejam considerados bons discriminadores de frequência, idade em que parte dos sujeitos atingiu nível de desempenho comparável ao de adultos. Por exemplo, bebês de sete a oito meses discriminam as frequências de complexos tonais que variavam espectralmente de 20 a 30 Hz, contudo seus desempenhos deterioram com complexos mais fortemente inarmônicos (Clarkson & Clifton, 1995). Estes experimentos com complexos tonais indicam que adultos e bebês podem se basear nas mesmas informações acústicas para ouvir frequências e guiar seu comportamento (Olsho, Schoon, Sakai, Turpin, & Sperduto, 1982). Portanto, modelos de percepção de tonalidade em adultos podem ser adequados para explicar esta capacidade em bebês.

Bebês de nove meses de idade demonstram percepção de estrutura prosódica da fala expressando sensibilidade a limites das frases. Eles preferem ouvir sentenças com pausas artificialmente inseridas entre unidades de frase, do que dentro dessas unidades (Gleitman & Warner, 1982). A mesma sensibilidade não é demonstrada por bebês de 6 meses, mas estes são sensíveis a limites de oração. De modo geral, o uso de estímulos da fala em estudos de percepção auditiva tem produzido resultados mais precisos que o uso de estímulos sonoros que não são de fala. Isto talvez se deva ao uso de estratégias mais consistentes do ouvinte para processar os sons da fala (Doherty & Turner, 1996).

A habilidade de percepção de fala em bebês difere dos adultos em três aspectos básicos: (1) limiares mais elevados, (2) maior suscetibilidade ao mascaramento de fonemas e (3) processamento temporal mais lento. Isto implica em um padrão

de funcionamento do sistema auditivo inferior ao do adulto, sugerindo que o ambiente acústico deve ser otimizado para que o bebê possa ter o máximo de recepção de linguagem para seu adequado desenvolvimento (Ruben, 1992).

### *Percepção categórica*

É característico dos sistemas perceptuais humanos agrupar estímulos em categorias cognitivamente eficientes, de modo a facilitar o armazenamento e a evocação de informação. O mecanismo complexo de percepção categórica capacita um indivíduo a reconhecer fonemas consistentemente, a despeito de uma grande variação em parâmetros acústicos cruciais. Outros tipos de variabilidade obscurecem a definição do sinal da fala, mesmo entre os estímulos mais diferentes. Duração das sílabas, mudanças na taxa de fala e padrões de ênfases, sexo do falante, idade e estado emocional resultam em amplas variações nas características físicas da fala, tais como a **freqüência fundamental** (o valor mais baixo de freqüência que compõe um som complexo) de vocalização, a **taxa de modulação** (variação no perfil de freqüências ou de intensidade do estímulo sonoro) de fala, e o **espaçamento de freqüências de ressonância**. Contudo, a percepção categórica mantém-se apesar destas variações, em um fenômeno de constância ou normalização perceptual (Eimas, 1985; Jusczyk, Houston, & Goodman, 1998). Este fenômeno está presente já em bebês de dois a três meses, que aparentemente compensam alterações na taxa de variação do espectro de freqüências que compõem o estímulo, neste caso transições iniciais de formantes (Eimas, 1985).

Estudos realizados desde a década de 70 por Eimas (1985), sobre percepção categórica de sons da fala, apóiam a noção de que bebês percebem uma descontinuidade natural para estímulos de um contínuo acústico precisamente no lugar onde as línguas do mundo dividem o contínuo em categorias fonéticas distintas. Esses dados sugerem que os bebês podem ter uma tendência natural para agrupar certos estímulos baseados em critérios que permanecem constantes, a despeito de outros aspectos do estímulo mudarem. Bebês de seis meses de idade reconhecem categorias auditivas que estão em conformidade com categorias vocálicas (/a/ versus /i/) percebidas por adultos. São também capazes de discriminar os sons de vogais com base em contorno de tonalidade (Kuhl, 1979).

Eimas (1985) demonstrou que bebês de língua nativa inglesa com menos de seis meses têm a habilidade de discriminar contrastes fonéticos que pertencem a outras línguas, diferentemente de bebês a partir de seis meses e de adultos, como os falantes de japonês, que têm dificuldade de distinguir entre /l/ e /r/. Contrastos de outras línguas poderiam ser discriminados caso pudessem ser assimilados por contrastes semelhantes da língua nativa (Mehler, Dupoux, Pallier, & Dehaene-Lambertz, 1994). Por outro lado, Pegg e Werker (1997) demonstraram que bebês de 10 e 12 meses e também adultos não discriminam contrastes próprios da língua quando eles não possuem *status* fonológico, ou seja,

percebem apenas variações fonéticas que correspondem a distinções significantes na língua nativa.

Bebês de sete meses demonstraram habilidade de categorização para complexos tonais com diferentes harmônicos baseando-se na frequência fundamental (Clarkson & Clifton, 1995). Efeito semelhante também ocorre em tarefas bem sucedidas de categorização de complexos tonais com fundamental ausente combinados com ruídos de baixa frequência, mas não quando combinados com ruídos de alta frequência, os quais cobriam a faixa dos harmônicos (Montgomery & Clarkson, 1997). Ou seja, muito precocemente bebês já demonstram a percepção da fundamental ausente através da reconstituição perceptual dela a partir dos seus harmônicos, múltiplos da fundamental. Os autores interpretam esses dados como indicativos de que, embora o sistema auditivo destes bebês ainda seja imaturo (seu desempenho da tarefa é pior do que o de adultos), eles recorrem ao processamento central para ouvir o tom da fundamental ausente.

As categorias fonéticas são explicadas pelo modelo prototípico, segundo o qual categorias fonéticas têm um centro representado por um fonema chamado de protótipo, que é entendido por um indivíduo como o melhor exemplar daquela categoria. O processo de categorização fonética é acompanhado pela computação da função de distância destes centros para os novos exemplares que serão então comparados. Uma questão interessante levantada por Samuel (1982) é se todos os falantes de uma língua têm essencialmente o mesmo protótipo para um dado som da fala. Os resultados de seu estudo indicam que os sujeitos variam ligeiramente na localização exata de seus protótipos, e esta variação teria por base a variação no VOT entre ouvintes adultos. Uma questão que se pode levantar é se há diferenças de desenvolvimento nesta construção de protótipos e qual seria o padrão de construção de bebês.

Pode-se afirmar que as categorias da fala são estruturadas internamente, com base em dados de adultos cujas categorizações são uniformemente simétricas em torno de uma localização particular no espaço vocálico. A percepção de excelência na categoria ocorre em níveis, e se altera com fatores como ritmo de produção de fala. Além disso, há um padrão interno de modelo no qual baseia seu julgamento de excelência, ou seja, há um padrão de vogal /i/ ideal, por exemplo. Estes padrões internos parecem basear-se em um bom estímulo ou protótipo da categoria (Kuhl, 1992). Esta construção do protótipo se dá dentro da língua do ouvinte.

Quando um protótipo da categoria serve como referencial, os outros membros da categoria são percebidos como muito similares a ele. O protótipo perceptualmente assimila seus vizinhos na categoria, efetivamente reduzindo a distância perceptual entre ele e os outros membros da categoria, funcionando como um ímã perceptual (Kuhl, 1991, 1992, 1993). Em outras palavras, o espaço perceptual subjacente às categorias da fala é modificado, de modo tal que a distância perceptual em torno do protótipo é reduzida em reação ao espaço perceptual em torno do não protótipo. Membros distantes da categoria são então perceptualmente

puxados para o centro da categoria. Esta deturpação do espaço perceptual subjacente às vogais pode ser devida a mecanismos de atenção e Kuhl denomina este fenômeno de “efeito de ímã”.

A presença de protótipos específicos à linguagem nativa em bebês de seis meses sugere que a fala é, pelo menos em uma forma primitiva, mentalmente representada na memória de longa duração já em bebês de idade muito precoce. Nesta faixa etária Kuhl (1992, 1993) identificou efeito de ímã maior para sons da língua nativa do que para sons de língua não nativa, pelo menos para sons de vogais. Kuhl acha possível que o efeito de ímã para vogais esteja presente ao nascimento e que estas tendências inatas sejam alteradas pela experiência lingüística pós-natal. É discutido atualmente se o efeito de ímã pode ser atribuído a fatores sensoriais apenas, ou a uma associação destes a outros fatores como memória de longa duração.

#### *A unidade do estímulo da fala: fonema versus sílaba*

Como anteriormente indicado, a fala é um estímulo de caracterização difícil. A linguagem falada tem uma natureza contínua. Além disso, quando descontinuidades ocorrem na forma de pequenos períodos de pausa ou alterações abruptas na natureza da energia acústica, essas descontinuidades tendem a não ter relação consistente com os perceptos, sejam eles palavras ou unidades pré-léxicas (por exemplo, segmentos fonéticos e sílabas). O sinal da fala é também marcado pela carência de invariância espectral, de informação temporal para a identificação de contrastes fonéticos específicos e combinações silábicas, dificultando a explicação dos processos de aquisição e reconhecimento de palavras (Kuhl, 1987).

Fonemas e sílabas são frequentemente propostos como unidades significativas da fala. Os fonemas provêm o menor inventário de unidades e, considerando o pequeno número de fonemas em qualquer língua, a aquisição e o reconhecimento de palavras por meio de um procedimento fonema-a-fonema podem ser relativamente eficientes. Contudo, a carência de invariância que surge do efeito de coarticulação (quando as propriedades acústicas de um fonema se ajustam às propriedades acústicas do fonema vizinho, descaracterizando a identidade de ambos), entre outros fatores contextuais, transformam em desafios a concepção física e a representação do fonema. Apesar do repertório de sílabas em qualquer língua ser considerado maior, especialmente em línguas com estruturas complexas, os efeitos coarticulatórios podem representar um problema menor. Além disso, deve-se considerar que se a busca de sentido léxico de uma palavra começar por qualquer fonema identificado ou qualquer sílaba, uma economia maior será alcançada se a busca começar por estruturas silábicas, à medida que a razão de buscas bem sucedidas será consideravelmente mais alta com buscas silábicas (Eimas, 1999).

Em uma revisão de trabalhos que investigaram a habilidade de bebês perceberem formações com representação tipicamente de segmentos fonêmicos ou sílabas, Eimas (1999) verificou que bebês de quatro dias de vida discriminavam a

reversão de consoantes iniciais e finais de padrões com três segmentos sintéticos quando as consoantes estavam na posição inicial e final (CVC), uma estrutura possível dentro da língua inglesa falada, mas não quando somente consoantes compunham a expressão (CCC), uma estrutura não usual. Além disso, a reversão da estrutura da primeira e terceira consoantes como em VCCCV, uma estrutura permissível para a língua inglesa, foram igualmente discrimináveis. Esses resultados são forte evidência para a idéia de que a sílaba é uma unidade natural de segmentação e processamento. Eimas também relata que bebês de quatro dias de vida distinguiram uma lista composta de expressões dissilábicas de outra trissilábica. Bebês da mesma população não discriminaram expressões de duas sílabas que tinham quatro ou seis fonemas, a despeito do fato de que a diferença na duração entre essas expressões era aproximadamente a mesma das expressões di e trissilábicas.

### *Percepção de fala e desenvolvimento de linguagem em pessoas com síndrome de Down*

Estudos direcionados à construção de metodologia de investigação adequada a bebês permitem a geração de conhecimentos esclarecedores acerca dos processos subjacentes ao desenvolvimento das habilidades de percepção de fala e de linguagem. A solidificação desta base de informação certamente permitirá compreender melhor o desenvolvimento em condições diferenciadas como no caso de síndromes genéticas, em especial a síndrome de Down, que é de grande incidência na população. A síndrome de Down (SD) é a manifestação de uma alteração cromossômica que tem inúmeras implicações para o desenvolvimento humano e que é diagnosticada em aproximadamente um em cada 800 nascidos vivos (Rondal, 1993). Está presente desde a formação do bebê, causando diferenças estruturais significativas quando comparadas ao desenvolvimento normal. As pessoas com essa síndrome apresentam peculiaridades fisiológicas e comportamentais e, dependendo dos contextos social, educacional e familiar em que estejam inseridas, podem vir a ter maior ou menor facilidade em se adaptarem a seu meio e, conseqüentemente, de usufruírem de seus direitos à qualidade de vida possível ao seu grupo social. Conhecer cientificamente os processos psicológicos específicos ao desenvolvimento deste grupo é relevante, pois esses conhecimentos podem subsidiar o planejamento de programas de intervenção, especialmente os de intervenção precoce.

Dentre os processos psicológicos básicos, os que estão mais prejudicados pela presença de um terceiro cromossomo 21 nas células somáticas são os de percepção auditiva, atenção, cognição, motivação e linguagem (Tristão & Feitosa, 1998, 2000). Estudos sobre desenvolvimento da linguagem em pessoas com SD evidenciam que os atrasos maiores parecem ocorrer principalmente na área da comunicação (Fischer, 1988; Miller, 1995), e de forma especial na expressão verbal. Por outro lado, a compreensão verbal associada à expressão gestual pode ser considerada área em vantagem no desenvolvimento de crianças com SD (Caselli et al., 1998).

Anomalias de cabeça e pescoço que podem afetar a função da trompa de Eustáquio, presentes na síndrome de Down, são condições propiciadoras de risco maior para perda auditiva temporária ou permanente devido à otite média (infecção de ouvido médio) com efusão crônica (Daly, Hunter, & Giebink, 1999; Nozza, 1988). A literatura tem descrito alterações em todas as partes do sistema auditivo periférico e central em pessoas com SD (Balkany, 1980; Becker, Mito, Takashima, & Onodera, 1991; Bilgin, Kasemsuwan, Schachern, Paparella, & Le, 1996; Dahle & Baldwin, 1992; Strome & Strome, 1992). No estudo de Tristão e Feitosa (2000) foi mostrada a relação entre distúrbios encontrados no aparelho auditivo, replicando os dados da literatura, e um comprometimento maior no desenvolvimento da linguagem.

Distúrbios auditivos parecem exercer um forte papel nas dificuldades de linguagem citadas, e há inúmeros trabalhos apontando alterações na condição auditiva das pessoas com SD (Tristão & Feitosa, 1998; Werner, Mancl, & Folsom, 1996). Pueschel e Sustrova (1996) revisaram estudos sobre percepção auditiva em bebês com SD no primeiro ano de vida e alguns achados são de especial importância para uma compreensão mais clara das características dessa categoria de habilidade perceptual neste grupo: (1) bebês com SD, da mesma forma que bebês sem esta alteração cromossômica, demonstram preferência por músicas com rimas cantadas por vozes femininas, em relação a músicas instrumentais; (2) demonstram respostas com durações mais longas para estímulos auditivos complexos do que bebês de grupo controle; (3) são mais suscetíveis em testes laboratoriais à distração auditiva e a olhar em outra direção em tarefas visuais; (4) demonstram memória auditiva com desempenhos inferiores à média em tarefas de reconhecimento; (5) também evidenciam lateralização invertida do processamento auditivo para estímulos verbais; (6) em testes de potencial evocado os tempos de reação indicam processamento auditivo mais lento que os grupos controle emparelhados por idade cronológica e por idade mental.

Estudos voltados especificamente para a descrição do desenvolvimento e do padrão de especialização cerebral na síndrome de Down têm mostrado especificidades em crianças e adultos. Vários desses estudos convergem para um modelo de dissociação da base biológica para a percepção de fala (de hemisfério direito) e para a produção de fala (de hemisfério esquerdo), proposto em 1987 por Elliott, Weeks e Elliott (Heath & Elliott, 1999). Esta organização é distinta da tipicamente encontrada em controles sem deficiência mental, ou com deficiência mental indiferenciada, nos quais se encontra vantagem do hemisfério esquerdo para percepção de fala. Heath e Elliott propõem que a dissociação hemisférica entre os sistemas responsáveis pela percepção de fala e por seu controle de execução motora poderia explicar dificuldades de processamento de informação encontrados nesta síndrome. Além disso, há evidências de que o cerebelo exerce importante papel no controle motor da produção de fala (Ackermann & Hertrich, 2000), e de que o perfil de microgirria encontrado em crianças com síndrome de Down inclui volume significativamente menor não apenas do cérebro como um

todo, mas também do cerebelo em particular (Pinter, Eliez, Schmitt, Capone & Reiss, 2001). Por outro lado, não foram encontradas alterações volumétricas no giro temporal superior, que pudessem ajudar a explicar as dificuldades de percepção de fala nesta síndrome.

Existe evidência de que o desenvolvimento pré-lingüístico e o uso funcional da linguagem em bebês com SD sejam diferentes de bebês normais. O surgimento de contato visual e o estabelecimento de níveis elevados deste comportamento parecem estar alterados na fase pré-lingüística (Spiker, 1990), implicando em uso funcional diferenciado de contato visual em situações interativas. O desenvolvimento fonológico é lento em pessoas com SD, embora a seqüência geral pareça acompanhar o desenvolvimento de crianças normais. A inteligibilidade da fala permanece baixa na maioria das pessoas com esta síndrome; verifica-se maior variabilidade na frequência fundamental e inadequação no controle da taxa de fala e na entonação da sentença (Chapman & Hesketh, 2000). Os erros cometidos são semelhantes aos observados no desenvolvimento fonológico de crianças normais, incluindo principalmente reduções nos agrupamentos e assimilações (Smith, 1987). Na fase adulta há uma incidência maior de tartamudez (gagueira) e hipernasalidade (Chapman & Hesketh, 2000). Dificuldades articulatórias parecem intensificar estas dificuldades fonéticas e fonológicas (Dodd, 1975; Rondal, 1988).

### Perspectivas para novos estudos

O estudo de populações específicas pode ser um recurso valioso para compreensão da ontogênese da percepção de fala. Devido a uma condição estrutural e fisiológica diferenciada, os bebês com SD apresentam um desenvolvimento diferente do apresentado em condições normais. A compreensão acurada de seu fenótipo pode embasar a proposição de alternativas de intervenção comportamental que favoreçam uma aquisição mais próxima possível do padrão natural de desenvolvimento.

Aspectos supraliminares da percepção auditiva devem ser mais pesquisados no bebê com síndrome de Down, principalmente aqueles relacionados à comunicação oral e à percepção do espaço, capacitações básicas para a convivência funcional com o meio físico e social. De forma mais específica, há necessidade de mais estudos sobre percepção categórica, competência aparentemente fundamental para a percepção de fala. Os resultados do BERA (*Brain Stem Evoked Response Audiometry* ou, em português, Audiometria de Resposta Evocada do Tronco Cerebral) neste grupo sugerem alterações no tronco cerebral indicando possíveis alterações neurais também em estruturas mais próximas ao córtex auditivo. O quanto isso teria a ver com alterações nas habilidades inatas de percepção de fala, e sua relação com alteração em representação fonética nestes indivíduos (Cossu, Rossini, & Marshall, 1993), são de interesse para a pesquisa na área da audição e desenvolvimento da fala. O quanto perdas leves podem

dificultar a distinção entre as unidades fonéticas da fala é também um tema ainda não pesquisado nesta população. Desconhece-se, ainda, o quanto fatores como atenção auditiva estão relacionados com as alterações de linguagem, que podem estar enraizadas no estabelecimento de relações alteradas entre som e perceptos fonéticos. Sabe-se por enquanto que bebês com SD apresentam um padrão diferenciado de atenção e habituação a estímulos de fala no primeiro ano de vida (Tristão, 2001; Tristão & Feitosa, 2002), que precisam ser mais bem investigados junto com uma análise da relação deste padrão com a condição auditiva neste período. Além da investigação desta relação, seria oportuno seu acompanhamento através de estudo longitudinal do desenvolvimento da linguagem desses bebês.

### Agradecimentos

A realização deste trabalho contou com apoio financeiro do CNPq/PIBIC para a primeira autora, e de uma bolsa de pesquisa do CNPq para a segunda.

### Referências

- Ackerman, H., & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics*, 13, 95-116.
- Balkany, T. J. (1980). Otologic aspects of Down syndrome. *Seminars in Speech, Language and Hearing*, 1, 39-48.
- Becker, L., Mito, T., Takashima, S., & Onodera, K. (1991). Growth and development of the brain in Down syndrome. *Progress in Clinical Biological Research*, 373, 133-152.
- Bilgin, H., Kasemsuwan, L., Schachern, P. A., Paparella, M. M., & Le, C. T. (1996). Temporal bone study of Down's syndrome. *Archives of Otolaryngology - Head & Neck Surgery*, 122, 271-275.
- Caselli, M. C., Vicari, S., Lombardi, E., Lami, L., Pizzoli, C., & Stella, G. (1998). Gestures and words in early development of children with Down syndrome. *Journal of Speech, Language and Hearing Research*, 41, 1125-1135.
- Chapman, R., & Hesketh, L. J. (2000). Behavioral phenotype of individuals with Down. *Mental Retardation and Developmental Disabilities: Research Reviews*, 6, 84-95.
- Clarkson, M. G., & Clifton, R. K. (1995). Infant's pitch perception: inharmonic tonal complexes. *Journal of the Acoustical Society of America*, 98, 1372-1379.
- Cleary, M., & Pisoni, D. B. (2001). Speech perception and spoken word recognition: research and theory. In E. B. Goldstein (Org.), *Blackwell handbook of perception* (pp. 499-534). Malden: Blackwell.
- Cossu, G., Rossini, F., & Marshall, J. C. (1993). When reading is acquired but phonetic awareness is not: a study of literacy in Down's syndrome. *Cognition*, 46, 129-138.
- Dahle, A. J., & Baldwin, R. L. (1992). Audiologic and otolaryngologic concerns. In S. Pueschel & J. Pueschel (Orgs.), *Biomedical concerns in persons with Down syndrome* (pp. 69-80). Baltimore: Paul Brookes.
- Daly, K. A., Hunter, L. L., & Giebink, G. S. (1999). Chronic otitis media with effusion. *Pediatrics in Review*, 20, 85-93.
- DeCasper A. J., & Fifer, W. P. (1980). Of human bonding - Newborns prefer their mothers voices. *Science*, 208, 1174-1176.
- DeCasper A. J., Lecanuet, J. P., Busnel, M. C., Graniere-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior & Development*, 17, 159-164.

- Dodd, B. (1975). Recognition and reproduction of words by Down's syndrome and non-Down's syndrome retarded children. *American Journal of Mental Deficiency, 80*, 306-311.
- Doherty, K. A., & Turner, C. W. (1996). Use of a correlational method to estimate a listener's weighting function for speech. *Journal of the Acoustical Society of America, 100*, 3769-3773.
- Eimas, P. D. (1985). The perception of speech in early infancy. In W. S-Y Wang (Org.), *The emergence of language – development and evolution* (pp. 117-127). Nova York: W. H. Freeman.
- Eimas, P. D. (1999). Segmental and syllabic representations in the perception of speech by young infants. *Journal of the Acoustical Society of America, 105*, 1901-1911.
- Fischer, M. A. (1988). The relationship between child initiations and maternal responses in preschool-age children with Down syndrome. In K. Marfo (Org.), *Parent-child interaction and developmental disabilities* (pp. 126-145). Nova York: Praeger.
- Gleitman, L. R., & Warner, E. (1982). Language acquisition: the state of the state of the art. In E. Warner & L. R. Gleitman (Orgs.), *Language acquisition: the state of the art* (pp. 3-48). Nova York: Cambridge University Press.
- Heath, M., & Elliott, D. (1999). Cerebral specialization for speech production in persons with Down syndrome. *Brain and Language, 69*, 193-211.
- Jenkins, J. J. (1992). The organization and reorganization of categories: the case of speech perception. In H. L. Pick Jr., P. V. D. Brook & D. C. Knill (Orgs.), *Cognition, conceptual and methodological issues* (pp. 11-31). Washington: American Psychological Association.
- Jusczyk, P. W., Houston, D., & Goodman, M. (1998). Speech perception during the first year. In A. Slater (Org.), *Perceptual development – visual, auditory and speech perception in infancy* (pp. 357-387). Hove, Reino Unido: Psychology Press.
- Kuhl, P. K. (1979). Speech perception in early infancy: perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America, 66*, 1668-1679.
- Kuhl, P. K. (1987). Perception of speech and sound in early infancy. In P. Salapatek & L. Cohen (Orgs.), *Handbook of infant perception* (pp. 275-382). Orlando, Florida: Academic.
- Kuhl, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50*, 93-107.
- Kuhl, P. K. (1992). Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes. In L. A. Werner & E. W. Rubel (Orgs.), *Developmental psychoacoustics* (pp. 293-331). Washington: American Psychological Association.
- Kuhl, P. K. (1993). Developmental speech perception: implications for models of language impairment. *Annals of the New York Academy of Sciences, 682*, 248-263.
- Lecanuet, J. (1998). Foetal responses to auditory and speech stimuli. In A. Slater (Org.), *Perceptual development – visual, auditory and speech perception in infancy* (pp. 317-355). Hove, Reino Unido: Psychology Press.
- Lecanuet J. P., Graniere-Deferre C., Jacquet, A. Y., & DeCasper, A. J. (2000). Fetal discrimination of low-pitched musical notes. *Developmental Psychobiology, 36*, 29-39.
- MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology, 49*, 199-227.
- Mehler, J., Dupoux, E., Pallier, C., & Dehaene-Lambertz, G. (1994). Cross-linguistic approaches to speech processing. *Current Opinion in Neurobiology, 4*, 171-176.
- Miller, J. F. (1995). Individual differences in vocabulary acquisition in children with Down syndrome. *Progress in Clinical and Biological Research, 393*, 93-103.
- Montgomery, C. R., & Clarkson, M. G. (1997). Infants' pitch perception: masking by low- and high-frequency noises. *Journal of the Acoustical Society of America, 102*, 3665-3672.
- Nearey, T. M. (1997). Speech perception as pattern recognition. *Journal of the Acoustical Society of America, 101*, 3241-3254.
- Nozza, R. J. (1988). Recent advances in audiologic assessment. *Pediatric Infection Disease, 7*, S169-S173.
- Nozza, R. J., Rossman, R. N. F., Bond, L. C., & Miller, S. L. (1990). Infant speech-sound discrimination in noise. *Journal of the Acoustical Society of America, 87*, 339-350.
- Ohde, R. N., Haley, K. L., & McMahon, C. W. (1996). A developmental study of vowel perception from brief synthetic consonant-vowel syllables. *Journal of the Acoustical Society of America, 100*, 3813-3824.
- Olsho, L. W., Schoon, C., Sakai, R., Turpin, R., & Sperduto, V. (1982). Auditory frequency discrimination in infancy. *Developmental Psychology, 18*, 721-726.
- Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America, 102*, 3742-3753.
- Pinker, S. (1995). Language acquisition. In L. R. Gleitman & M. Liberman (Orgs.), *Language* (2<sup>a</sup> ed., pp. 135-182). Cambridge: MIT.
- Pinter, J. D., Eliez, S., Schmitt, J. E., Capone, G. T., & Reiss, A. L. (2001). Neuroanatomy of Down's syndrome: a high-resolution MRI study. *The American Journal of Psychiatry, 158*, 1659-1665.
- Pueschel, S. M., & Sustrova, M. (1996). Visual and auditory perception in children with Down's syndrome. In J. Rondal, J. Perera, L. Nadel & A. Comblain (Orgs.), *Down's syndrome: psychological, psychobiological, and socio-educational perspectives* (pp. 53-64). California: Singular Publishing.
- Rondal, J. A. (1988). Parent-child interaction and the process of language acquisition in severe mental retardation: beyond the obvious. In K. Marfo (Org.), *Parent-child interaction and developmental disabilities* (pp. 114-125). Nova York: Praeger.
- Rondal, J. A. (1993). Down's syndrome. In D. Bishop & K. Mogford (Orgs.), *Language development in exceptional circumstances* (pp. 165-176). Hillsdale: Lawrence Erlbaum.
- Ruben, R. J. (1992). The ontogeny of human hearing. *Acta Otolaryngologica (Stockholm), 112*, 192-196.
- Samuel, A. G. (1982). Phonetic prototypes. *Perception & Psychophysics, 31*, 307-314.
- Slater, A. (1998) (Org.). *Perceptual development – visual, auditory and speech perception in infancy*. Hove, Reino Unido: Psychology Press.
- Smith, L. (1987). The structure of “dialogue” in early language development: longitudinal case studies of Down syndrome and nonretarded toddlers. In H. Rauh & H. C. Steinhausen (Orgs.), *Psychobiology and early development* (pp. 201-213). North Holland: Elsevier.
- Spiker, D. (1990). Intervention strategies from a developmental perspective. In D. Cicchetti & M. Beeghly (Orgs.), *Down syndrome: the developmental perspective*. Nova York: Cambridge University Press.
- Strome, S. E., & Strome, M. (1992). Down syndrome: an otolaryngologic perspective. *The Journal of Otolaryngology, 21*, 394-397.
- Tristão, R. M. (2001). *Percepção de fala em crianças com síndrome de Down no primeiro ano de vida*. Tese de doutorado não-publicada, Universidade de Brasília, Brasília.
- Tristão, R. M., & Feitosa, M. A. G. (1998). Linguagem na síndrome de Down. *Psicologia: Teoria e Pesquisa, 14*, 127-137.
- Tristão, R. M., & Feitosa, M. A. G. (2000). Percepção auditiva e implicações para o desenvolvimento global e de linguagem em crianças com síndrome de Down. *Arquivos Brasileiros de Psicologia, 52*(2), 118-142.
- Tristão, R. M., & Feitosa, M. A. G. (2002). Use of the visual habituation paradigm to investigate speech perception in Down syndrome infants.

- Proceedings of the Eighteenth Annual Meeting of the International Society for Psychophysics – Fechner Day 2002, Vol 18* (pp. 552-557). Rio de Janeiro: International Society for Psychophysics.
- van Ooijen, B., Bertoni, J., Sansavini, A., & Mehler, J. (1997). Do weak syllables count for newborns? *Journal of the Acoustical Society of America*, 102, 3735-3741.
- Werner, L. A., Mancl, L. R., & Folsom (1996). Preliminary observations on the development of auditory sensitivity in infants with Down syndrome. *Ear & Hearing*, 17, 455-468.
- Werner, L. A., & Rubel, E. W. (1992). (Orgs.) *Developmental psychoacoustics*. Washington: American Psychological Association.

---

## Notas

1. “Um fonema é uma entidade linguística abstrata que se refere à menor unidade de som que pode distinguir uma palavra com sentido de outra em uma língua específica. O que constitui um fonema difere de língua para língua. O termo fone é algumas vezes usado para referir a uma possível instância de um fonema” (tradução livre de Cleary & Pisoni, 2001, p. 526).
2. “Formantes são regiões concentradas de energia acústica, criadas com a intensidade acentuada de certas harmônicas da frequência fundamental e a atenuação de outras harmônicas da fundamental, devidas às características de ressonância do trato vocal.” (Cleary & Pisoni, 2001, p. 505).
3. VOT é “a diferença de tempo entre o jorro de ruído causado pela liberação abrupta de articuladores fechados (como os lábios) até a primeira vibração da corda vocal que ocorrer a seguir” (tradução livre de Cleary & Pisoni, 2001, pg. 503).

---

Rosana Maria Tristão, doutora em Psicologia (Processos comportamentais) pela Universidade de Brasília, atua profissionalmente na Secretaria de Educação do Distrito Federal e Universidade de Brasília.  
 Maria Angela Guimarães Feitosa, doutora em Psicologia (Psicobiologia) pela Universidade de Michigan, Ann Arbor (EUA), é professora no Instituto de Psicologia, Universidade de Brasília (UnB).  
 Endereço para correspondência: Universidade de Brasília, Instituto de Psicologia, Departamento de Processos Psicológicos Básicos, Laboratório de Psicobiologia; 70910-900 – Brasília, DF. Tel.: (61) 307-2625. Fax: (61) 273-0203. E-mail: rmtt@unb.br