

ADQUIRINDO A FONOLOGIA DE UMA LÍNGUA: PRODUÇÃO, PERCEPÇÃO E REPRESENTAÇÃO FONOLÓGICA

Raquel Santana SANTOS¹

- RESUMO: Este artigo traça um panorama sobre o que sabemos a respeito da aquisição fonológica, em primeira língua – principalmente nos últimos anos, com o advento de novas tecnologias que nos permitem explorar o processo de aquisição antes mesmo de a criança começar a falar. O que significa adquirir a fonologia de uma língua? Que conhecimento é este que as crianças devem adquirir? O que os estudos sobre percepção e produção nos dizem sobre as habilidades infantis? Em que os estudos de aquisição fonológica podem ajudar no desenvolvimento das teorias fonológicas?
- PALAVRAS-CHAVE: Fonologia. Aquisição. Produção. Percepção.

Introdução: O que toda criança deve saber?

Sempre que pensamos em aquisição fonológica, vêm-nos à mente os sons que a criança vai utilizar para veicular o que ela sabe, pensa, quer. No entanto, a fonologia de uma língua não se reduz a sons, e nem a sua aquisição, portanto, deve ser reduzida à percepção do conjunto de fonemas de uma língua.

Então, o que a criança deve “aprender”? Ou o que deve fazer parte de seu conhecimento fonológico?

Obviamente que a criança deve descobrir quais são as unidades que compõem o sistema fonológico de sua língua; mas para isso, em primeiro lugar, ela deve descobrir como mapear o contínuo acústico em unidades discretas. A cultura letrada muitas vezes não nos permite perceber este trabalho. Acostumados a usar letras, parece-nos óbvio (basta perguntar a qualquer leigo) que ouvimos sons separadamente, que inserimos pausas entre palavras. Mas não é este o caso; basta que observemos um espectrograma. O que vemos é uma onda contínua, às vezes com pausas entre palavras, mas que definitivamente não traz fronteiras entre os segmentos, mesmo porque muitas vezes estas fronteiras também não existem: os segmentos podem se intercalar ou se sobrepor. Por exemplo, uma sentença como *a Maria acaba usando ovos na receita* pode ser produzida como [a.ma.ri.a.ka.bu.zã.dwɔ.vuz.na.xe.sej.ta], em que as palavras *maria* e *acaba* sofrem degeminação, *acaba* e *usando* sofrem elisão, e *usando* e *ovos* sofrem ditongação. Todos estes processos unem as palavras de forma que não há uma pausa (silêncio)

¹ USP - Universidade de São Paulo. Faculdade de Filosofia, Letras e Ciências Humanas - Departamento de Linguística. São Paulo - SP - Brasil. 05508-900. - raquelss@usp.br

entre elas. Mesmo assim, as crianças têm que aprender a 'recortar' os segmentos no contínuo sonoro em unidades discretas, pois são sobre estas unidades que elas vão trabalhar.

Além disso, as crianças devem aprender a fazer distinções – tanto na percepção quanto na produção –, descobrindo que diferenças são significativas, contrastivas em sua língua, e quais são insignificantes. Isto é, uma criança adquirindo o português brasileiro deve aprender que uma diferença de sonoridade entre um [b] e um [p] é importante em sua língua, mas que a diferença entre um [u] e um [ʊ] não é relevante, assim como não é relevante a diferença de altura (*pitch*) entre a fala de sua mãe e a de sua babá (não entrarei aqui em considerações sobre diferenças na fala da mãe, por exemplo, que carreguem significados pragmático-discursivos, como uma advertência *versus* uma proibição).

A criança também deve descobrir as regras fonotáticas e fonológicas que dão forma às palavras de sua língua. Ela deve aprender que uma seqüência [kt] não é uma seqüência bem formada no português brasileiro, por exemplo, e deve também aprender que a distinção de sonoridade muitas vezes deve ser desconsiderada – como nos casos dos segmentos/fonemas /s/ e /z/ que se neutralizam em posição de coda de sílaba (por exemplo, as pessoas dizem [a.'xas.ta] 'arrasta', mas ['xaz.ga] 'rasga'). Mais ainda, ela deve aprender que esta neutralização é condicionada por alguns fatores (em nosso exemplo, a sonoridade do segmento seguinte, [t] e [g], respectivamente). Finalmente, deve descobrir se as regras são opcionais ou obrigatórias. Por exemplo, a neutralização da distinção de sonoridade dos segmentos acima ocorre sempre, enquanto que uma regra como o alçamento de vogal (que nos permite dizer [me'ninu] ou [mi'ninu] 'menino'), é opcional, muitas vezes ocorrendo a depender da formalidade da situação.

Mas a tarefa da criança não termina nos segmentos. Sabemos que a fonologia organiza e atualiza seus segmentos em estruturas. Ilustramos no exemplo acima que a neutralização de um segmento ocorre em uma determinada posição estrutural, a coda da sílaba para a fricativa coronal (e não no ataque da sílaba – e por isso não há neutralização em português brasileiro entre ['ka.za] 'casa' e ['ka.sa] 'caça'). Então, as crianças devem descobrir as regras de combinação dos segmentos em unidades, estruturas maiores, como a sílaba, a palavra, e quais são as combinações proibidas. Por exemplo, a criança deve descobrir que, em português, apenas as consoantes fricativas (coronais), líquidas e nasais podem preencher a posição de coda – como em ['ka.zas] 'casas', [por.'tal] 'portal' e [tãm'bẽ] 'também'. Ou seja, as crianças têm que descobrir, por exemplo, se a língua permite sílaba com coda ou com ataque complexo (e quais os segmentos que podem preenchê-los), se as palavras têm um tamanho mínimo.

Para além dos segmentos, as crianças também precisam descobrir as propriedades prosódicas de sua língua e como elas se organizam (por exemplo:

qual o sistema acentual da língua, como se organiza o ritmo: sílabas pesadas contam? e moras?). Tomemos o sistema acentual do português brasileiro como exemplo. A proposta de Lee (1995) é que a regra geral de acentuação de palavras é dada através da formação de uma unidade binária (duas sílabas), no lado direito da palavra, mas sem contar a sílaba que leva a vogal temática (que marca o gênero). A sílaba acentuada é aquela do lado direito da unidade. Por exemplo, na palavra *menino*, desconsidera-se a sílaba final (que traz o marcador de gênero masculino), o que sobra é [me.ni]. A sílaba acentuada é a da direita, [ni]. No caso de uma palavra como *caracol*, não se desconsidera nenhuma sílaba final (porque não há marcador de gênero). As duas sílabas da unidade mínima são [ra.kɔw], a sílaba a ser acentuada é a direita, [kɔw].

No entanto, para conseguir organizar todo este conhecimento acima, as crianças devem descobrir pistas naquilo que ouvem para as estruturas que organizam sua língua. Voltemos ao nosso exemplo sobre o acento. Santos e Fikkert (2006) chamam a atenção para o fato de que as crianças podem usar diferentes pistas para descobrir que a unidade binária que gera a acentuação de palavras tem a sílaba forte à direita (um iambo). Uma delas é o conjunto de vogais que preenchem as sílabas fracas. Em português brasileiro, no dialeto de São Paulo, por exemplo, as sílabas pré-tônicas podem ser preenchidas por 5 diferentes vogais ([i,e,a,o,u]), enquanto as pós-tônicas são preenchidas por apenas 3 ([i,a,u]). Curiosamente, em holandês (cuja regra geral de acentuação é a criação de uma unidade binária com a sílaba da esquerda acentuada – um troqueu), ocorre o padrão inverso: há mais vogais preenchendo pós-tônicas do que pré-tônicas. Essa é uma possível evidência para as crianças do lado em que está a sílaba acentuada na unidade binária (que gera o acento de palavra). Em português brasileiro, a unidade binária é formada por uma pré-tônica mais uma tônica (portanto, um iambo), enquanto em holandês, é formada por uma tônica mais uma pós-tônica (portanto, um troqueu).

Quando a aquisição fonológica começa? A percepção

Na seção precedente vimos o que toda criança deve saber para chegar à língua adulta, mas não discutimos quando o processo de aquisição começa. Tal processo inicia-se muito antes de a criança emitir os primeiros sons. Os bebês desenvolvem a audição por volta da 25ª semana de gestação (PUJOL, LAVIGNE-REBILLARD; UZIEL, 1991), mas mesmo antes disso (na 16ª semana) já respondem a pulsos sonoros (SHAHIDULLAH; HEPPER, 1992). Mas o que as crianças escutam? Experimentos em que se implantaram microfones no lado externo da parede de úteros de mulheres grávidas mostraram que as crianças escutam sons do ambiente (QUERLEU et al., 1988). Experimentos análogos em ovelhas permitiram

observar que, de dentro do útero, há acesso apenas à estrutura prosódica (rítmica e entoacional), mas não aos segmentos. As paredes uterinas funcionam como um filtro que apaga as frequências mais baixas do contínuo sonoro, responsável pela discriminação dos segmentos e deixam apenas as frequências mais altas, que nos dão a modulação de entoação, por exemplo (HUANG et al., 1997).

Mas a pergunta que se coloca é se esta informação que elas escutam é tratada estruturalmente. Há uma maneira de se testar isso. Se a criança está prestando apenas atenção no som e não na estrutura que o organiza, então não deve importar como o som é apresentado. No entanto, se as crianças estiverem prestando atenção também na estrutura, a forma de apresentar o estímulo deve importar, pois ela pode afetar a estrutura. Elaborou-se um teste que checkou se as crianças percebiam mudança de línguas. Mehler et al. (1988) conduziram dois experimentos com crianças francesas de 4 dias, para observar se elas percebiam a diferença entre o francês e o russo. No primeiro experimento, as crianças ouviam sentenças de russo, e depois de francês. O resultado foi que elas perceberam a diferença. No segundo experimento, as sentenças eram tocadas de frente pra trás. Quando isso ocorre, os segmentos permanecem (os [b], [s] estão todos lá), mas a estrutura em que eles se organizam desaparece (por exemplo, o segmento que estava na coda da sílaba é ouvido antes dos segmentos que estão no núcleo ou no ataque). O resultado foi que as crianças não conseguiram mais distinguir o russo do francês. Em resumo, as crianças, aos 4 dias, estavam prestando atenção à estrutura rítmico-entoacional e não apenas à discriminação de sons de cada língua.

Ainda no âmbito supra-segmental/prosódico, outros experimentos mostram que muito antes de as crianças produzirem os primeiros sons, elas já identificam padrões acentuais em sua língua. Newsome e Jusczyk (1995) e Jusczyk et al. (1999) mostram que as crianças americanas de 7,5 meses acostumadas com *hamlet* e *kingdom* (duas palavras trocaicas – dissílabas com acento na penúltima sílaba) estranhavam quando ouviam, no texto, *ham* e *king*, o que mostra que elas estavam trabalhando com unidades binárias. No entanto, quando acostumadas com *surprise* (uma palavra iâmbica – dissílabas com acento final), elas não estranhavam quando ouviam *price*. Segundo os autores, esta é uma evidência de que as crianças (americanas) estão assumindo que as palavras iniciais são troqueus (dissílabas paroxítonas).²

Os exemplos expostos ilustram a percepção inicial que as crianças têm de estruturas prosódicas. Mas há também experimentos mostrando o trabalho com

² Trabalhos defendendo uma tendência trocaica universal (FIKKERT, 1994; DEMUTH, 1996; ADAM; BAT-EL, 2008) deixaram de lado o fato de que as crianças tinham 7 meses e assumiram esses resultados como evidência do inatismo do troqueu. Infelizmente, não há trabalhos deste tipo para crianças adquirindo línguas iâmbicas (em que o acento é final).

segmentos. Experimentos com crianças com horas de vida (EIMAS et al., 1971) mostram que elas já têm uma percepção categorial e não gradiente do contínuo sonoro. Crianças com 1 mês distinguem /b/ de /p/ com os mesmos valores de *voice onset time* (VOT) que os adultos, isto é, com uma distância de valor 25 ms entre eles.

As crianças pequenas distinguem segmentos mesmo que não sejam de sua língua alvo. Werker et al. (1981) testaram americanos de 8 meses que distinguiram os segmentos /t^h/ e /d^h/ do hindi. Além de distinguir categorialmente, as crianças mostram saber bem cedo que distinções são irrelevantes. Aos 2 meses, elas distinguem *bug* de *dug* – e não distinguem erroneamente quando *bug* é produzida por falantes diferentes (JUCSZYK, 1997) – o que indica que as crianças não estão fazendo a comparação levando em conta as características de fala dos falantes. Finalmente, as crianças demonstram perceber as regras de ordenação de segmentos. Moon, Cooper e Fifer (1992) mostraram que crianças com 51 horas de vida percebem a diferença na ordem [pst] vs. [tsp].

Por fim, inúmeros experimentos revelam também que esta capacidade da criança de distinguir segmentos e estruturas que não são de sua língua alvo vai desaparecendo conforme ela fica mais velha. Crianças de 4 anos falham em distinguir os segmentos de hindi do experimento de Werker (WERKER; TESS, 1983). Conforme o tempo passa, elas vão se tornando cada vez mais especializadas na língua alvo, e perdem a habilidade de distinguir diferenças de outras línguas.

Quando a criança começa a falar? A produção

Todos os experimentos acima tornam evidente que muito antes de a criança começar a falar, ela já sabe muito da estrutura fonológica de sua língua. Interessantemente, há um grande descompasso entre o que ela percebe e o que ela produz.

Nas primeiras seis semanas, elas produzem sons vegetativos, como arrotar, engolir, arrulhar, e começam a balbuciar (normalmente repetindo seqüências de consoante-vogal, como *mama*) por volta dos 6 meses. A pergunta que se coloca é se estes sons do balbucio têm alguma relevância para o sistema fonológico. Há análises que vêem neles apenas sons vegetativos, que indicam um trabalho neuromotor por parte das crianças, sem conexão com a linguagem (JAKOBSON, 1941; LENNEBERG, 1967). Outros defendem ser eles já o início do trabalho com o componente fonológico (BOYSSON-BARDIES; SAGARD; DURANT, 1984; PETITTO; MARENTEPTE, 1991, entre outros). Não há uma conclusão propriamente, mas chamo a atenção para os seguintes trabalhos, que lançam luzes sobre esta questão: crianças surdas balbuciam vocalmente (LENNEBERG, 1967; SMITH; BROWN-

SWEENEY; STOEL-GAMMON, 1989). Este é um grande argumento para entender tais segmentos como movimentos reflexos, trabalho motor, mas não lingüístico. No entanto, outros trabalhos mostram haver, já no balbucio, evidências da língua que está sendo adquirida. Em primeiro lugar, o balbucio oral de crianças surdas tem características fonéticas diferentes do balbucio de crianças ouvintes (OLLER et al., 1985). Além disso, Petitto e Marentette (1991) mostram que crianças surdas também balbuciam por sinais (produzindo ou repetindo certas combinações de movimentos ou disposições manuais), e o desenvolvimento deste tipo de produção é paralelo ao balbucio vocal de crianças ouvintes, o que indica sua importância para aquisição de uma língua (quer oral, quer de sinais). No que diz respeito às características fonéticas do balbucio oral, Boysson-Bardies et al. (1984) apontam que adultos conseguem distinguir o balbucio de crianças de 8 meses que estão adquirindo sua língua alvo – o francês – de crianças adquirindo chinês ou árabe. Curiosamente, somente foneticistas conseguiam fazer a mesma distinção para balbucio de crianças de 6 meses. No entanto, os estudos ainda não conseguem dizer o que, no balbucio, pode estar refletindo alguma estrutura da língua alvo.

As primeiras palavras surgirão apenas por volta dos 12 meses. Embora estejamos discutindo aquisição de fonologia, o surgimento das palavras é fundamental para esta discussão, dado que só podemos discutir se as crianças adquiriram um determinado segmento na língua alvo quando elas o utilizam corretamente em uma palavra (isto é: o procedimento do investigador é observar se a criança já está usando o fonema – unidade mínima contrastiva). Todos os trabalhos sobre aquisição segmental vão então dizer quando as crianças adquiriram um determinado segmento olhando para palavras.

Os estudos de Lamprecht e seu grupo (LAMPRECHT et al., 2004) mostram um caminho na aquisição segmental do português brasileiro: plosivas e nasais ([‘papa] ‘papai’, [‘mãmã] ‘mamãe’) >> fricativas ([si’zẽw] ‘feijão’) >> líquidas ([‘olu] ‘olho’).

Esses trabalhos tornam evidente também que tais segmentos surgem em momentos diferentes, conforme a posição estrutural: primeiro, em ataque; depois, em coda (por exemplo, no período em que ela diz [sor’veɥfɪ] ‘sorvete’, ainda diz [‘kaka] ‘casca’). São primeiro estabilizados em sílabas tônicas, depois em átonas (ela produz primeiro [kẽta] ‘canta’ enquanto ainda produz [ko’mo] ‘comprou’). Os segmentos em coda final são estabilizados antes que os de coda medial (ela diz [a’xos] ‘arroz’ quando ainda produz [‘kaka] ‘casca’).

Estes resultados revelam mais uma vez a interação entre a aquisição segmental e a estrutura. Enquanto as estruturas e os segmentos não são adquiridos (estabilizados – e voltarei mais tarde a esta questão), a criança usa de inúmeras estratégias: omissão/ apagamento de sílabas e segmentos: [‘na.na] ‘banana’, [‘pe.to] ‘preto’; assimilação: [pa.‘pẽw] ‘chapéu’; metátese: [‘per.to] ‘preto’;

reduplicação: [ma.ma.'mãw] 'mão'; epêntese: [ba'ra.bo] 'bravo'; substituição: ['pli.pis] 'príncipe'.

Surge então a questão de por que parece haver uma ordem na aquisição dos segmentos e das estruturas (silábica, por exemplo). Os resultados de Lamprecht et al. (2004) para o português corroboram a proposta de Jakobson (1941) de que a aquisição é pré-ordenada. Segundo o autor, isto seria devido a leis/universais implicacionais: se um contraste mais baixo aparece, então um mais alto deverá existir. Se temos CVC, então temos CV. Se temos fricativas, então temos plosivas.

O trabalho de Mota (1997) sugere como se dá o caminho de aquisição dos segmentos do português (o trabalho foi inicialmente proposto para lidar com os dados de aquisição com desvio, mas posteriormente aplicado com sucesso para dados de aquisição normal). Não só os segmentos são formados por conjuntos de traços hierarquizados (como propõem Clements e Humes (1995)), como também há uma relação implicacional entre estes traços que leva em conta a noção de marcação e complexidade dos mesmos (no sentido de Calabrese (1995)). É esta relação implicacional que geraria a ordem. Por exemplo, para uma criança produzir /g/, ela deve marcar os traços de ponto de dorsal e de vozeado. Por isto, antes de produzir /g/, ela produzirá /k/, que é dorsal, mas não vozeado. Esta proposta poderia ser vista simplesmente como um desenvolvimento da proposta de Jakobson (1941), mas é mais que isso. A proposta de Jakobson não permite variações, e elas são encontradas na fala infantil. Algumas crianças começam produzindo distinções de sonoridade (/p/ *versus* /b/) antes de ponto de dorsal (/p/ *versus* /k/); outras fazem o contrário. A proposta de Mota (1997) mostra como é possível, levando em conta as relações entre os traços, haver diferentes caminhos para as crianças, no processo de aquisição do inventário fonológico. Voltando ao exemplo, algumas crianças seguirão o seguinte percurso: /p,t,m,n/ >> /p,b,t,d,m,n/ >> /p,b,t,d,k,m,n/ >> /p,b,t,d,k,g,m,n/ (marcando primeiro a sonoridade, depois o ponto de dorsal), enquanto que outras seguirão um caminho diferente: /p,t,m,n/ >> /p,t,k,m,n/ >> /p,b,t,d,k,m,n/ >> /p,b,t,d,k,g,m,n/ (marcando primeiro o ponto de dorsal depois a sonoridade). Levelt e Veijer (2004) discutem esta mesma possibilidade de variação na aquisição de estruturas silábicas em holandês: algumas crianças seguem o caminho CV >> ... >> CVCC, VCC >> CCV, CCVC, enquanto outras seguem CV >> ... >> CCV, CCVC >> CVCC, VCC. Os autores argumentam que isto é devido à diferente ordem de demerção das restrições.

Os trabalhos em aquisição pouco têm abordado a aquisição de regras fonológicas da língua, elemento essencial na elaboração da linguagem pela criança. No caso do português, as crianças devem aprender, por exemplo, que há uma neutralização entre /s/ e /z/ em posição de coda. Alguns pesquisadores (SMITH, 1973; STEMBERGER, 1989; NEWTON; WELLS, 1999, 2002) consideram

que há processos naturais que são aplicados automaticamente, e que as crianças devem aprender a suprimi-los. O processo de sonorização (vozeamento) de fricativa parece ser um excelente candidato para um processo natural, dado que pode ser facilmente explicado como a sobreposição de um traço de um segmento (no caso, o traço [vozeado]) sobre outro segmento. No entanto, os dados mostram que o processo não é tão simples assim, como aponta a análise de Silva (2008).

Em primeiro lugar, o processo de vozeamento não é igual em todas as línguas. O português apresenta vozeamento regressivo (casa[s], casa[z] amarelas), enquanto outras línguas apresentam progressivo (em grego, sin+pléko si[m]léko 'eu tricoto', ton+patéra to[m]b]aterá 'o pai' (NESPOR; VOGEL, 1986). Mais ainda: muitas vezes esta regra é condicionada morfológicamente (em inglês, o plural sofre assimilação progressiva bug[z], cat[s], enquanto palavras terminadas em consoantes sonoras que são parte do radical sofrem desvozeamento quando há uma palavra iniciada por obstruinte desvozeada, como em [faiv] *five* 'cinco' e [faif tons] *five tons* 'cinco toneladas' – Roca e Johnson (1999)). Logo, não estamos à frente de um processo natural, mas de uma regra, com propriedades a serem adquiridas (Quais segmentos sofrem a assimilação? Qual a direção do processo? Há influência morfológica? Silábica?).

Vejam, agora, o que a literatura do português brasileiro diz sobre a aquisição deste processo. No que se refere à estrutura silábica, a estrutura CVC, necessária para a produção da coda fricativa, é adquirida por volta dos 2;0 anos (SANTOS, 1998, 2001; MEZOMMO; MENEZES, 2001). No que diz respeito à aquisição destes dois fonemas, em posição de ataque medial o /z/ é adquirido aos 2;2 anos, e o /s/ aos 2;5 (SILVA, 2008). Ora, uma vez que o fonema sonoro é adquirido antes do fonema surdo, o esperado seria que as crianças errassem na seguinte direção: os segmentos sonoros seriam produzidos corretamente, enquanto os segmentos surdos seriam produzidos com "erros". Não é isto que ocorre: nos 3.067 dados analisados, o surgimento e uso dos segmentos surdos na posição de coda é correto (a partir dos 1;8 anos), enquanto as crianças erram no uso dos segmentos sonoros (ou apagando, ou produzindo-os como surdos, entre outras estratégias) até os 3;2 anos. Logo, podemos concluir que o processo de vozeamento se dá por regra, e não ocorre automaticamente.

Mas quando ocorre a aquisição?

Uma questão para a qual gostaria de voltar é a da estabilização de um segmento, ou, dito de outra maneira, quando um segmento é adquirido. Sempre que estamos falando em aquisição, estamos pensando em produção. A maior parte dos estudos defende que as crianças tenham que produzir corretamente entre

80% e 100% das vezes – às vezes por 2 ou 3 sessões seguidas – para considerar um segmento estabilizado.

No entanto, vimos que a percepção infantil está muito à frente. Com horas de vida, a criança já consegue distinguir categorialmente os segmentos. Por volta dos 8 meses, a capacidade infantil de distinguir segmentos/estruturas que não são de sua língua alvo decresce brutalmente, mas não há mudanças significativas na discriminação dos segmentos/estruturas de sua língua alvo. Aos 12 meses elas só retêm o que é específico da língua alvo. Focalizando a época de estabilização das fricativas (2;0 para /z/), temos então um *lapso* de 15 meses para explicar. O que temos nestes 12 meses de intervalo?

A resposta é a interação com outros componentes da gramática, e esta inserção afetará tanto a discriminação dos segmentos, quanto a aplicação de processos.

Os experimentos feitos com crianças recém-nascidas e de até aproximadamente 7 meses são todos desenhados para testar as capacidades infantis sem precisar levar em conta o significado. Os principais métodos são de *high amplitude sucking* (que mede a sucção das crianças em uma chupeta ligada ao computador) e uma modificação do *preferential looking*. Há duas versões deste último experimento:

- a criança é condicionada a detectar a modificação visual sempre após uma modificação sonora. Assim, ao ouvir um som diferente, ela vira-se para ver o que acontece.
- ela aprende a comandar o computador. Se virar para um lado, ouvirá um tipo de estímulo; se virar para outro lado, ouvirá outro tipo de estímulo. A criança, então, escolhe o segmento que quer ouvir.

Os métodos, a partir dos 8 meses, já envolvem significado. No *intermodal split paradigm*, por exemplo, a criança vê na tela uma pata e uma gata e ouve uma das duas palavras. A criança tende a fixar mais o olhar na figura que tem o nome pronunciado.

Interessantemente, crianças de 6 meses mostram excelentes resultados em tarefas discriminatórias do tipo *high amplitude sucking* (que, repito, não envolvem significado; elas têm apenas que distinguir [t] *versus* [d], por exemplo) (EIMAS et al., 1971). No entanto, quando estes mesmos segmentos são apresentados em palavras, o desempenho da criança piora consideravelmente – tanto na produção (as crianças holandesas trocam os segmentos até os 6 anos – Zamuner, Kerkhoff e Fikkert (2006)) – quanto na percepção (também até por volta dos 6 anos – KUIJPERS, 1996; FEEEST, 2007). O fenômeno *puzzle-puddle* também ilustra muito bem a não estabilização de um segmento (SMITH, 1973). A criança adquirindo inglês produz *puddle* como [p^gl], o que indicaria uma dificuldade em produzir

coronais, mas ela produz *puzzle* como [p[^]dl], o que indica que o problema não é na produção, e sim nos movimentos articulatórios que ela deve fazer e combinar. Em português brasileiro, Lamprecht et al. (2004) apontam que as crianças produzem *sala* como [tala], mas *chave* como [savi].

Uma resposta rápida e simples para estes problemas é que a entrada do significado traria uma maior carga para o processamento infantil, e por isso o erro. No entanto, essa não necessariamente é a melhor resposta. Um teste com crianças holandesas de 17 meses mostra que crianças têm um comportamento que não é de simples substituição de segmentos (FIKKERT, 2008, experimento do segundo estágio). As crianças foram acostumadas a ouvir *din* como nome para um boneco até então desconhecido. As crianças não percebiam diferença quando o boneco era chamado de *bin*; no entanto, quando foram acostumadas que o nome do boneco era *bin*, elas percebiam a diferença se no teste ele era chamado de *din*. Esses resultados levam-nos a nos perguntar por que as crianças aceitavam uma determinada modificação, mas não outra.

É necessário levar em conta dois fatos para entender o que aconteceu. Em primeiro lugar, as crianças são excelentes na tarefa de comparação e discriminação de sons, levando em conta todos os “detalhes” dos segmentos ouvidos (como vimos nos experimentos descritos acima). Em segundo, muitas pesquisas têm apontado para o fato de que as crianças, quando aprendem uma palavra, não guardam todas as informações fonológicas sobre elas (voltaremos a este assunto mais abaixo). Ora, as pesquisas sobre línguas naturais (falo aqui de sistemas alvo, fala adulta) apontam para o valor não-marcado dos segmentos coronais (no sentido de JAKOBSON, 1941): estes segmentos estão presentes nos diversos inventários fonológicos e processos fonológicos tendem freqüentemente a tornar outros segmentos coronais. Se o coronal é um segmento não-marcado (e assumindo que seja, então, subespecificado quanto ao ponto), ele deve estar assim, também, na fala infantil. *Din* é formado por coronais. Como as coronais são subespecificadas, é como se a forma guardada fosse 0-0-0 (quanto ao ponto de articulação). Quando, por *din*, diz-se *bin*, troca-se a coronal pela labial. Como a forma guardada era subespecificada, não há incompatibilidade (*mismatch*) entre os pontos de articulação (a criança compara zero com labial). Quando, por outro lado, troca-se a labial (*bin*) pela coronal (*din*), agora temos uma incompatibilidade: a forma guardada era labial-0-0 e a criança escuta coronal-coronal-coronal (lembro que a criança tem acesso a todas as informações nas tarefas de discriminação). O segmento utilizado, agora, tem um traço diferente do segmento que a criança tem na memória. Chamo a atenção para o fato de que, se fosse uma questão de processamento, as duas formas alternantes deveriam ser igualmente rejeitadas pelas crianças.

Esta questão aponta, então, para direções de pesquisas atuais em aquisição de fonologia que apenas começaram: a informação fonológica do léxico é armazenada? Há várias hipóteses. A primeira é que a criança registra a forma fonética. Neste caso, *gosta* é guardado como [ˈgɔsta] e *gosma* como [ˈgɔzma]. Se a criança guarda este tipo de informação lexical, então ela deve procurar outras pistas, que não as segmentais, para saber se em um ambiente como [ewkõprejũ kabunifõrmi] para *eu comprei um cabo uniforme* se aplica uma regra fonológica como a elisão ou degeminação (entre [ʋʋ]).

Uma segunda hipótese é que a criança guarda/registra a forma fonológica das palavras, da mesma forma que os adultos. Neste caso, *gosta* é guardado como /gɔsta/ e *gosma* como /gɔsma/. Se a criança desde o início guarda este tipo de informação, então ela tem que descobrir quais são as neutralizações possíveis, de forma a não guardar *casa* e *caça* como homônimas: */kaSa/.

Finalmente, a terceira hipótese é que a criança guarda/registra a forma fonológica, inicialmente subespecificada: neste caso, no início do processo de aquisição, palavras como *din* seriam guardadas como: Consoante(coronal subespecificada)+Vogal(coronal subespecificada)+Consoante(coronal subespecificada). Neste caso, devemos então investigar como e quando a representação lexical na criança atinge a forma adulta.

Aquisição e Propostas Fonológicas

Dados de aquisição como o discutido acima sobre o vozeamento servem não apenas para ajudar a explicar como as crianças adquirem regras (ou ao menos para mostrar que as regras são/devem ser aprendidas), mas também ajudam a tomar algumas decisões em relação às análises das línguas adultas. Explicome: o processo de dessonorização é comum em diversas línguas. Em holandês, por exemplo, assume-se que o segmento final é subjacentemente sonoro, e por regra é dessonorizado (por exemplo /pad/ é produzido [pat] – compare com /lat/, produzido [lat] – GRIJZENHOUT; KRÄMER, 2000). Dado que se trata de um processo tão comum, é legítimo levantar a questão de se o que temos em português brasileiro é um segmento sonoro que dessonoriza (/z/ -> [s]) ou um segmento surdo que sonoriza (/s/ -> [z]). A estratégia de análise do linguísta é observar a quantidade de contextos em que os segmentos aparecem. Temos surdas frente a pausas e outros segmentos surdos, e sonoros apenas frente a segmentos sonoros. Logo, assumimos que a regra se aplica em um fonema /s/ que é ‘transformado’ em [z]. Mas como saber se nossa análise reflete o que realmente ocorre em nossa mente, ou se nossa regra é resultado de uma estratégia/análise do pesquisador? Em outras palavras, será que o nosso argumento de quantidade de contextos realmente importa em nossa mente? Pode ser muito bem a idéia de

que a quantidade de contextos não importa na definição da regra e, nesse caso, podemos muito bem hipotetizar que a regra é de dessonorização de /z/ frente a pausa ou segmento surdo.

No entanto, os dados de aquisição de Silva (2008) trazem evidências independentes de que a proposta atual da lingüística (de que a regra é de sonorização) está correta. A autora mostra que, muito embora as crianças adquiram os segmentos fricativos sonoros antes dos surdos, elas têm dificuldades é com os sonoros quando no contexto da regra, o que indica que os surdos é que estão sujeitos a processo, que eles é que sofrem mudanças.

Um outro caso em que os dados de aquisição fonológica trazem evidências para uma ou outra análise das línguas naturais diz respeito à estrutura dos ditongos crescentes (que são sempre resultado de um processo de ditongação). Há duas possíveis posições para o glide dos ditongos crescentes (CglV) nas diferentes línguas: uma é que o glide preenche o ataque (ANDRADE; MATEUS, 1998 para o português europeu; BOOJ, 1989 para o frisão; KAYE; LOWENSTAMM, 1984 para o francês; HARRIS, 1983 para o espanhol – nestas duas últimas, apenas em sílabas sem consoante inicial), e outra que preenche o núcleo (HARRIS, 1983 para o espanhol; KAYE; LOWENSTAMM, 1984 – para sílabas com consoante inicial).

A questão é saber qual destas estruturas (ou se as duas) é a que ocorre em português brasileiro. Quando ditongos decrescentes (CVgl) sofrem nasalização, tanto a vogal quando o glide são nasalizados (por exemplo em ['pãjna] 'paina'), o que indica que a nasalização é um processo que ocorre na rima. No entanto, os glides de ditongos crescentes não são nasalizados (por exemplo em [pãnu] 'piano'), portanto só pode ser o caso do glide estar fora da rima, no ataque, como propõem López (1979) e Bisol (1999). Os dados de aquisição trazem mais evidências para esta análise: os ditongos decrescentes CVgl surgem ao mesmo tempo que as sílabas de rima ramificada CVC, enquanto que os ditongos crescentes CglV surgem ao mesmo tempo que as sílabas de ataque ramificado CCV (SANTOS, 2001, 2007).

Considerações Finais

Como vimos, embora pareça simples, a aquisição da fonologia de uma língua é bem complexa, como têm apontado os estudos. Antes mesmo de nascer, as crianças já têm acesso à estrutura de sua língua, e com horas de vida já são capazes de perceber regularidades nos segmentos, nas estruturas e nas regras. Mesmo assim, a aquisição não é instantânea: há um longo caminho a percorrer.

Os resultados encontrados levam-nos de volta ao problema de Platão (como sabemos tanto, tendo tão poucas evidências?) e à adequação explanatória nos

termos de Chomsky (1981). Isto é, qualquer teoria sobre a linguagem que se queira adequada deve dar conta de explicar como a criança a adquire. As tecnologias atuais têm-nos permitido descobrir mais sobre o conhecimento lingüístico das crianças e aberto novas fronteiras de investigação, auxiliando-nos nesta tarefa.

Agradecimentos: Agradeço a Jairo Nunes a leitura prévia e discussão deste texto.

SANTOS, R. S. Acquiring Language Phonology: Production, Perception and Phonological Representation. *Alfa*, São Paulo, v.52, n.2, p.465-481, 2008.

- *ABSTRACT: This paper presents an overview of what we know about first language phonology acquisition. What does it mean to acquire phonology? What kind of knowledge should children acquire? What do studies on production and perception tell us about children's initial abilities? How can studies on phonology acquisition help us develop phonological theories?*
- *KEYWORDS: Phonology. Acquisition. Production. Perception.*

Referências

ADAM, G.; BAT-EL, O. The trochaic bias is universal: new evidence from hebrew. In: GAVARRÓ, A.; FREITAS, M. J. *Language Acquisition and Development: proceedings of GALA 2007*. Newcastle: Cambridge Scholars Publishing, 2008. p.12-24.

ANDRADE, E. D'.; MATEUS, M. H. M. The syllable structure in portuguese. *D.E.L.T.A*, São Paulo, v. 14, n.1 p.13-32, 1998.

BISOL, L. A sílaba e seus constituintes. In: NEVES, M. H. de M. (Org.). *Gramática do Português Falado: novos estudos*. Campinas: Ed. da UNICAMP, 1999. v.7. p.701-742.

BOOJ, G. On the representation of diphthongs in Frisian. *Journal of Linguistics*, Cambridge, v. 25, p.319-332, 1989.

BOYSSON-BARDIES, B.; SAGART, L.; DURAND, C. Discernible differences in the babbling of infants according to target language. *Journal of child language*, Cambridge v.11, n.1, p.1-15. 1984.

CALABRESE, A. A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry*, Cambridge, v.26, n.3, p. 373-463. 1995.

CHOMSKY, N. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CLEMENTS, G. N, HUME, E. V. The internal organization of speech sounds. In: GOLDSMITH, J. A. (Org.). *The handbook of phonological theory*. Cambridge: Blackwell. 1995. p.245-301.

DEMUTH, K. The prosodic structure of early words. In: MORGAN, J. L., DEMUTH, K. (Ed.). *Signal to syntax: bootstrapping from speech to Grammar in early acquisition*. Hillsdale: Lawrence Erlbaum, 1996. p. 171-184.

EIMAS, P. D. et al. Speech Perception in infants. *Science*, Washington, v.171, n.3968, p.303-306, 1971.

FEEST, S. van der. *Building a phonological lexicon: the acquisition of the dutch voicing contrast in perception and production*. Utrecht: Prince Productions B.V., 2007. Disponível em: <http://webdoc.uhn.ru.nl/mono/f/feest_s_van_der/buila_phl.pdf>. Acesso em: 19 fev. 2007.

FIKKERT, P. Developing representations and the emergency of phonology. Evidence from perception and production. *LabPhon*, Paris, v.10, 2008. No prelo.

_____. *On the acquisition of prosodic structure*. Amsterdam: Holland Institute of Generative Linguistics, 1994.

GRIJZENHOUT, J.; KRÄMER, M. Final devoicing and voicing assimilation in dutch derivation and cliticization. In: STIEBELS, B.; WUNDENRLICH, D. (Ed). *Lexicon in Focus: studia grammatical*. Berlin: Akademie Verlag, 2000. v.45. p.55-82.

HARRIS, J. W. *Spanish phonology*. Cambridge: MIT Press, 1983.

HUANG, X. et al. Temporary threshold shifts induced by low-pass and high-pass filtered noises in fetal sheep in útero. *Hearing Research*, Amsterdam, v.113, n.1-2, p.173-181, 1997.

JAKOBSON, R. *Child Language Aphasia and Phonological Universals*. Paris: Mouton. 1941.

JUSCZYK, P. W.; HOUSTON, D. M.; NEWSOME, M. The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, New York v.39, n.3-4, p.159-207, 1999.

KAYE, J. D.; LOWENSTAMM, J. De la syllabicité. In: DELL, F.; HIRST, D.; VERGNAUD, J-R. (Ed). *Forme sonore du langage: structure des représentations en phonologie*. Paris: Hermann. 1984. p.123-159.

KUIJPERS, C. T. L. Perception of the voicing contrast by dutch children and adults. *Journal of Phonetics*, London, n.24, n.3, p.367-382, 1996.

LAMPRECHT, R. et al. (Org). *Aquisição fonológica do português: perfil de desenvolvimento e subsídios para terapia*. Porto Alegre: Artmed. 2004.

LEE, S-H. *Morfologia e Fonologia Lexical do Português Brasileiro*. 1995. 190f. Tese (Doutorado em Lingüística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.

LENNEBERG, E. H. *Biological foundations of language*. New York: John Wiley. 1967.

LEVELT, C.; VIJVER, R. van de. Syllable types in cross-linguistic and developmental grammars. In: KAGER, R.; PATER, J.; ZONNEVELD, W. (Ed.) *Fixing Priorities: constraints in Phonological Acquisition*. Cambridge University. Press, 2004. p.204-218.

LÓPEZ, B. S. *The sound pattern of brazilian portuguese: cariocan dialect*. 1979. 275f. Dissertation (Doctor of Philosophy) – University of California, Los Angeles, 1979.

MEHLER, J. et. al. A precursor of language acquisition in young infants. *Cognition*, Amsterdam, v.29, n.2, p.143-178, 1988.

MEZOMMO, C.; MENEZES, G. Comparação entre a aquisição da estrutura silábica no português brasileiro e no português europeu. *Letras de Hoje*, Porto Alegre, v.36, n.125, p. 691-698, 2001.

MOON, C.; COOPER, R. P.; FFER, W. P. Canonical and non-cnonical syllable discrimination by two-day-old infants. *Journal of Child Language*, Cambridge, v.19, n.1, p.1-17, 1992.

MOTA, H. B. Aquisição segmental do português: um modelo implicacional de complexidade de traços. *Letras de hoje*, Porto Alegre, v.32, n.4, p. 23-47, 1997.

NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986.

NEWSOME, M.; JUSCZYK, P.W. Do infants use stress as a cue for segmenting fluent speech? In: MacLAUGHLIN, D.; MacEWEN, S. *Proceedings of the 19th Annual Boston University Conference on Language Development*. Somerville: Cascadilla Press, 1995. v.2. p.415-426.

NEWTON, C, WELLS, B. Between-word junctures in early multi-word speech. *Journal of Child Language*, Cambridge, v.29, n.2, p. 275-299, 2002.

NEWTON, C, WELLS, B. The development of between-word processes in the connected speech of children aged between three and seven years. In: MAASSEN, B., GROENEN, P. *Pathologies of speech and language: advances in clinical phonetics and linguistics*. Londres: Whurr, 1999. p. 67-75.

OLLER, D. K. et al. Pre-speech vocalizations of a deaf infant: a comparison with normal metaphnological development. *Journal of speech and hearing research*, Rockville, v.28, n.4, p.47-63. 1985.

PETITTO, L. A., MARENTETTE, P. F. Babbling in the manual mode: evidence for the ontogeny of Language. *Science*, Washington, v.251, n.5000, p. 1493-1496, 1991.

PUJOL, R.; LAVIGNE-REBILLARD, M.; UZIEL, A. Development of the human cochlea. *Acta Oto-laryngologica*, Olso, v.111, n.482, p. 7-123, 1991.

QUERLEU, D. et al. Fetal hearing. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, Amsterdam, v.29. p.191-212, 1988.

ROCA, I.; JOHNSON, W. *A course in phonology*. Oxford: Blackwell, 1999.

SANTOS, R. S. *A aquisição prosódica do português brasileiro de 1 a 3 anos: padrões de palavra e processos de sândi externo*. 2007. 214 f. Tese (Livre-docência) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007.

_____. *Aquisição do acento primário em português brasileiro*. 2001. 316f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2001.

_____. A aquisição da estrutura silábica. *Letras de Hoje*, Porto Alegre. n.112, p. 91-98, 1998.

SANTOS, R. S.; FIKKERT, J. P. M. The prosodic shape of early words: Comparing evidence from Romance and Germanic languages. In: THE ROMANCE TURN, 2., 2006, Utrecht. *Program and abstracts...* Utrecht, 2006. p.23-24.

SHAHIDULLAH, S.; HEPPEL, P. G. Hearing in the fetus: Prenatal detection of deafness. *International journal of prenatal and perinatal studies*, London, v.4, p.235-240, 1992.

SILVA, C. C. *Aquisição da regra de assimilação de vozeamento em português brasileiro*. 2008. 160f. Dissertação (Mestrado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008.

SMITH, N. V. *The acquisition of phonology: a case study*. Cambridge: Cambridge University Press, 1973.

SMITH, B. L.; BROWN-SWEENEY, S.; STOEL-GAMMON, C. A quantitative analysis of reduplicated and variegated babbling. *First Language*, London, v.9, n.6, p.175-189, 1989.

STEMBERGER, J.P. Speech errors in early child language production. *Journal of Memory and Language*, Cambridge, v.28, n.2, p.164-188, 1989.

WERKER, J. F. et al. Developmental aspects of cross-language speech perception. *Child Development*, Chicago, v.52, n.1, p. 349-355. 1981.

WERKER, J. F.; TESS, R. C. Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psycholinguistics*, Toronto, v.37, n.2, p. 278-286, 1983.

ZAMUNER, T. S.; KERKHOFF, A.; FIKKERT, P. *Children's knowledge of how phonotactics and morphology interact*. 2006. Manuscrito não publicado.

Recebido em fevereiro de 2008

Aprovado em junho de 2008