

**COMPARANDO A PERCEPÇÃO DE CRIANÇAS MONOLÍNGUES E BILÍNGUES**

*Sabrina Gewehr-Borella<sup>1</sup>*

*Márcia Cristina Zimmer<sup>2</sup>*

*Ubiratã Kickhöfel Alves<sup>3</sup>*

**RESUMO**

Neste trabalho, analisamos e comparamos os resultados de um teste de percepção aplicado em 30 alunos, organizados em três grupos distintos: monolíngues da cidade de Rio Grande, que não possuem contato com falantes de outra língua (grupo MR); monolíngues da cidade de Picada Café, que possuem contato com falantes bilíngues (grupo MP) e bilíngues da cidade de Picada Café (grupo B). Ao trabalharmos a percepção de oclusivas surdas e sonoras por grupos que apresentam produções distintas nesses segmentos, buscamos encontrar diferenças também na percepção de nossos participantes. Embasados na Fonologia Gestual, partimos de uma perspectiva dinâmica da percepção da fala. O teste, elaborado pelos autores, foi dividido em duas partes: alternativas com VOTs não-manipulados e manipulados (6 questões). A partir da análise dos dados, constatamos mais erros de percepção nas alternativas não-manipuladas no grupo MP, seguido do grupo MR e do grupo B. Diante de um VOT zero - alternativa manipulada - os grupos de Picada Café optaram pela opção 'A'(surda), enquanto que o grupo de Rio Grande preferiu a alternativa 'C'(sonora). Foi encontrada diferença significativa em apenas uma das análises realizadas, mostrando haver uma diferença existente entre o grupo bilíngue e os demais grupos. Acreditamos que as diferenças encontradas nos grupos de Picada Café ocorram em razão do contato de seus participantes com a língua de imigração Hunsrückisch, o que vai ao encontro da concepção dinâmica de aquisição de linguagem aqui explicitada.

**Palavras-chave:** Percepção de oclusivas do PB. *Voice Onset Time*. Transferência interlinguística.

## INTRODUÇÃO

Até o início dos anos 60, poucas investigações sobre a relação entre percepção e produção dos sons da fala haviam sido realizadas. Já na década de 70, Stevens e House (1972) compartilharam a ideia de que os processos de percepção e produção estão estreitamente interligados (GAMA, 1994). No entanto, há relatos de que inter-relações entre percepção e produção são *eventuais*, o que demonstra que caminhos distintos também podem ser traçados (MUNRO, BOHN, 2007).

A partir da observação de resultados díspares apresentados pela literatura a respeito da relação entre percepção e produção, buscamos, no presente artigo, comparar a percepção de falantes monolíngues (português brasileiro) e bilíngues (português brasileiro e Hunsrückisch) que apresentam produções orais diferentes, para verificar se há, de fato, indícios de relação estreita entre a interface percepção/produção.

Em razão da escassez de pesquisas referentes à percepção de alunos monolíngues e bilíngues, este trabalho tem caráter relevante, uma vez que a partir dele poderemos ter um melhor entendimento da relação 'percepção'/ 'produção' e do funcionamento do sistema bilíngue em si.

A seguir, apresentamos um pequeno referencial teórico, seguido do método utilizado na pesquisa. Depois disso, são apresentados os dados analisados, além das conclusões obtidas a partir da aplicação do teste realizado e de nossa visão teórica baseada na Fonologia Gestual.

## 2. REFERENCIAL TEÓRICO

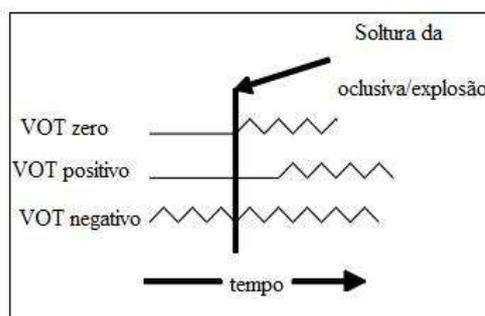
### 2.1 A LÍNGUA DE IMIGRAÇÃO HUNSRÜCKISCH

A língua de imigração Hunsrückisch, “uma variedade supra-regional do alemão falado no sul do Brasil (...)” (ALTENHOFEN, 1996, p.27), foi introduzida em nosso país pelos pioneiros da imigração alemã a partir do início do século XIX, com a chegada dos imigrantes alemães na chamada província de São Pedro do Rio Grande (MÜLLER, 1999).

Uma das características dessa língua reside no fato de que os padrões de vozeamento de suas oclusivas são diferentes daquelas do português brasileiro, o que pode gerar transferências interlinguísticas, tanto na produção como na percepção, causando, assim, potenciais falhas na comunicação entre falantes monolíngues e bilíngues.

## 2.2 O VOICE ONSET TIME

De acordo com Lisker e Abramson (1964, *apud* REIS, NOBRE-OLIVEIRA, 2007), o *Voice Onset Time* (VOT) é o período de surdez entre a soltura/explosão da consoante e o início da periodicidade de vozeamento do segmento seguinte. A partir da Figura 1, que segue, podemos compreender melhor tal explicação:



**FIGURA 1:** Três padrões de VOT<sup>4</sup>.

Conforme podemos observar na Figura 1, os padrões de vozeamento tendem a ser caracterizados a partir de três categorias de VOT:

- 1<sup>a</sup>) negativa, que nos mostra um pré-vozeamento, ou seja, uma vibração das cordas vocais antes da soltura da oclusiva, em média de -100ms;
- 2<sup>a</sup>) zero, que apresenta um período de surdez pequeno, em que o início de vozeamento e a soltura ocorrem em um período bastante próximo, em torno de + 10ms;

3ª) positiva, apresentando um período de surdez mais longo, isto é, um retardo no início da vibração das pregas vocais, em torno de + 75ms (LISKER, ABRAMSON, 1964, *apud* REIS, NOBRE-OLIVEIRA, 2007). Ocorre nesta categoria a aspiração.

Como os valores de VOT sofrem variações influenciadas por características como idade, velocidade da fala, dentre outras, não há consenso entre os pesquisadores a respeito dos seus valores médios para cada consoante. Contudo, alguns estudiosos apresentam algumas classificações. De acordo com Reis e Nobre-Oliveira (2007), os fonemas surdos do PB apresentam as seguintes médias: /p/- 12ms, /t/- 18ms e /k/- 38ms – todos classificados, portanto, dentro da categoria ‘zero’, anteriormente descrita. Já as oclusivas sonoras possuem pré-vozeamento, o que as enquadra dentro da categoria ‘negativa’.

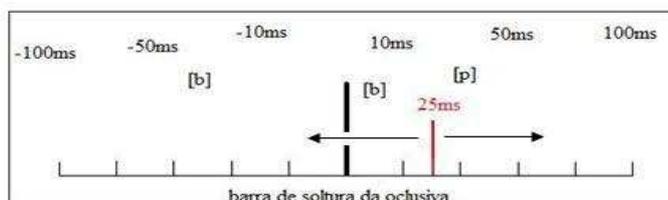
Quanto aos padrões de vozeamento das línguas de imigração alemãs, nas quais se inclui o Hunsrückisch, Braun (1996) relata que a grande maioria “apresenta um período de surdez longo nas oclusivas /p/, /t/ e /k/, interpretadas como aspiradas, e um curto período de surdez nas oclusivas /b/, /d/ e /g/, concebidas como surdas não-aspiradas” (BRAUN, 1996, p. 30). Jessen e Ringen (2002) acrescentam que o VOT negativo, ou seja, o pré-vozeamento, é bastante raro nas línguas de imigração alemãs.

A partir de uma análise realizada com falantes de Hunsrückisch da cidade de Picada Café/RS, chegamos às seguintes médias de valores de VOT: ([p<sup>h</sup>]=87ms, [t<sup>h</sup>]=81ms, [k<sup>h</sup>]=91ms, [p]=27ms, [t]=23ms, [k]=40ms) (GEWEHR-BORELLA, 2010), sendo os três primeiros fones classificados dentro da categoria ‘positiva’ e os três últimos da categoria ‘zero’. Vejamos, agora, o que diz a literatura a respeito da percepção dos padrões de vozeamento.

### 2.3 A PERCEPÇÃO DOS PADRÕES DE VOT

Quanto à percepção dos padrões de VOT formados pelas diferentes línguas, Lieberman e Blumstein (1988) afirmam que ela é categórica. No par [p/b], por

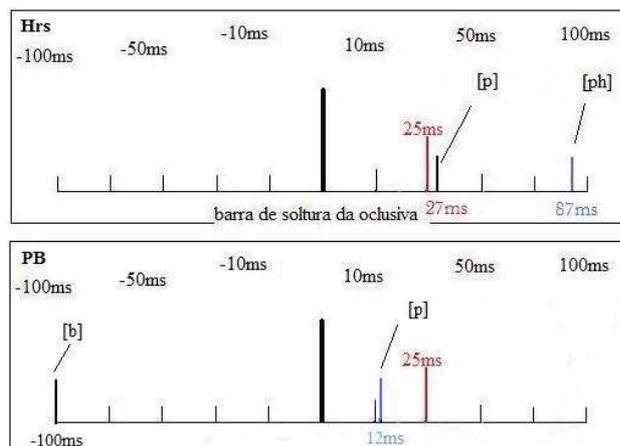
exemplo, as oclusivas que têm atraso de vozeamento, isto é, menores de 25ms, são percebidas como um [b], enquanto que as maiores de 25ms são percebidas como um [p] (LIEBERMAN, BLUMSTEIN, 1988, p. 197), conforme podemos visualizar na Figura 2.



**FIGURA 2-** Percepção do VOT.

Os autores apontam que não há diferença perceptível, por exemplo, nos valores de 40 e 60ms; ambos são percebidos como [p]. No entanto, se escutarmos a mesma diferença, de 20ms, nos valores de 10 e 30ms, a distinção será notada, o primeiro sendo percebido como um [b] e o segundo como um [p] (LIEBERMAN, BLUMSTEIN, 1988). Parece haver, aí, uma *fronteira fonética* que faz com os ouvintes distingam um ou outro fone. De acordo com a concepção dinâmica de fonologia que rege este artigo, sabemos que tais fronteiras apresentam “lugares” variados não somente de língua para língua, mas, também, que dentro de um próprio sistema os valores não são, tampouco, estanques. O fonema /p/ do PB, por exemplo, é produzido com menos de 25ms, e nem por isso é percebido como um [b]. Provavelmente, a grande quantidade de pré-vozeamento, aliada à ausência de aspiração apresentada no português brasileiro, venha a “transferir” a chamada *fronteira fonética* mais para perto da barra de soltura da oclusiva, o que faz com que, ao percebermos um fone de 10ms, ou até menor, já o classifiquemos como um /p/ no PB. Com base no apresentado, temos que deixar claro que não concordamos com a visão “categórica” apresentada por Lieberman e Blumstein. A partir da concepção dinâmica defendida neste artigo, vemos a percepção, portanto, como um *continuum*.

Apesar de partimos deste *continuum*, não podemos negar que existam diferenças entre a percepção de falantes monolíngues e bilíngues, uma vez que as produções (médias de padrões de vozeamento) de suas línguas-mães são bastante diversas, como podemos observar na Figura 3, apresentada a seguir:



**FIGURA 3:** Média do VOT de [p,p<sup>h</sup>] do Hunsrückisch (parte superior) e de [p,b] do português brasileiro (parte inferior).

Por apresentarem valores de VOT positivo, alguns falantes de Hunsrückisch tendem a transferir um pouco desta elevação de VOT para suas produções também em PB, realizando, por exemplo, um [p] com média de 28ms, bem maior do que o valor de 12ms apresentado pela literatura (REIS, NOBRE-OLIVEIRA, 2007). Em razão da diversidade de produções, acreditamos que possa haver, também, diferença na percepção de alunos monolíngues e bilíngues.

Pesquisas têm demonstrado que a percepção influencia a produção dos VOTs (WANG, BEHNE, 2007; BIRDSONG, 2007; FLEGE, 1987 *apud* SANCIER, FOWLER, 1997). É importante apontarmos também que, ao escutarmos os sons da fala, não percebemos apenas os padrões de VOT, mas todos os segmentos vizinhos presentes na informação acústica (BORDEN *et al.*, 1994). Portanto, diferentemente das letras no mundo escrito, discretas e separáveis, tanto a produção oral como a percepção devem ser entendidas e analisadas levando em conta o seu conjunto. Tal visão de conjunto nos remete à perspectiva dinâmica apresentada pela Teoria dos Sistemas Dinâmicos, discutida a seguir.

## 2.4 A DINAMICIDADE DA FALA E DA PERCEPÇÃO

Diferentemente das teorias de aquisição de linguagem que consideram a aquisição como um processo que segue um desenvolvimento linear, a Teoria dos

Sistemas Dinâmicos (TSD) vai buscar no comportamento não-linear<sup>5</sup>, muitas vezes imprevisível, o desenvolvimento da linguagem, através da observação de aspectos cognitivos e sociais.

Uma das principais características dos sistemas dinâmicos é a questão da interconexão das variáveis. Todo o sistema formado está ligado, o que significa que, quando uma variável é modificada, todo o sistema é impactado. Podemos observar com isto que os estados de um sistema se modificam no transcorrer do tempo, através da orquestração dos gestos (GOLDSTEIN *et al.*, 2007).

Os modelos de Fonologia Articulatória (FAR) e Acústico-Articulatória (FAAR), modelos de Fonologia Gestual, estão enquadrados dentro de uma perspectiva dinâmica, onde os padrões criados dentro de um sistema estão em constante modificação. Em oposição à visão binária da Fonologia Clássica, os novos modelos de Fonologia (FAR- FAAR) primam pelo destaque conferido à gradiência na produção dos segmentos.

A FAR busca analisar os processos contínuos que podem ocorrer durante a produção da fala através da observação dos gestos, tomados como primitivos de análise dotados de tempo intrínseco. De acordo com Browman & Goldstein (1992), os gestos já são produzidos na fase do balbucio, período inicial da aquisição da fala. A FAAR (ALBANO, 2001), por sua vez, além de analisar a parte articulatória, já presente na FAR, acrescenta ao modelo as informações de ordem acústica, obtidas através da realização do gesto.

Temos que ter claro que, ao percebermos um som, não estamos ouvindo uma série de segmentos isolados consecutivos, mas um fluxo de elementos gradientes sendo articulados, muitas vezes, sobrepostamente. Dependendo das pistas do *input*, os seres humanos podem modificar suas falas. Esta modificação é bem característica da formação dos sotaques, mesmo na fase adulta. Um exemplo de modificação ocorrida com frequência, em razão do contato com uma segunda língua, diz respeito aos padrões de VOT (*Voice Onset Time*), anteriormente citados.

Após esta breve revisão teórica, passamos, agora, para a visualização do método utilizado neste trabalho.

### 3. MÉTODO

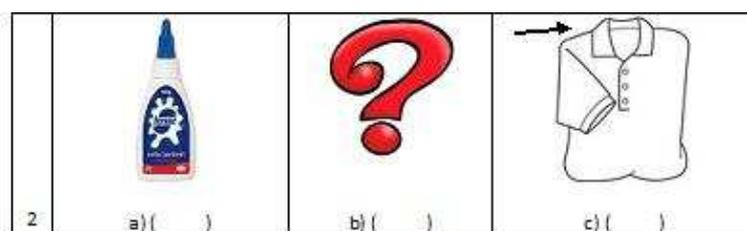
A partir do objetivo geral da pesquisa, apresentado na introdução, buscamos:

a) Visualizar qual grupo de participantes comete o maior número de erros na percepção de questões sem manipulação de VOT (que apresentam os padrões de VOT encontrados no PB); b) Verificar onde ocorrem mais erros na percepção de questões sem manipulação de VOT, isto é, nos segmentos sonoros ou nos segmentos surdos do PB, por parte dos três grupos participantes e, finalmente, c) Investigar, nas questões com manipulação de VOT, a direção escolhida pelos participantes (mais para o surdo ou mais para o sonoro) frente a um segmento iniciado por VOT 0, artificialmente produzido.

Para a presente pesquisa, elaboramos um teste de percepção que foi aplicado em 30 participantes (7-11 anos), divididos em três grupos distintos:

- 1º) Grupo MR: 10 crianças monolíngues, falantes do português brasileiro (PB), que não possuem contato com crianças bilíngues (da cidade de Rio Grande/Sul do RS);
- 2º) Grupo MP: 10 crianças monolíngues, falantes do português brasileiro (PB), que possuem contato com crianças bilíngues (da cidade de Picada Café/Nordeste do RS), cidade essa na qual a língua de imigração Hunsrückisch faz-se presente a todo instante no falar de seus habitantes.
- 3º) Grupo B: 10 crianças bilíngues, falantes do português brasileiro (PB) e da língua de imigração alemã Hunsrückisch (também da cidade de Picada Café).

O teste, por nós elaborado, consistia na percepção auditiva de 18 palavras, apresentadas dentro de uma frase veículo (A palavra é...). Ao iniciar o teste, cada aluno recebia três folhas contendo um total de 18 tiras, como a apresentada a seguir.



**FIGURA 4:** Exemplo de questão do teste de percepção.

O aluno era instado a ouvir a frase-veículo e, logo depois, assinalar a opção que havia escutado. No exemplo apresentado (Figura 4), os alunos, sentados em conjunto em uma sala de aula, foram expostos à seguinte frase: ‘a *palavra* é...*cola*’. Neste caso, marcariam: (a) se escutassem [kɔla]; b) se ficassem na dúvida; e c) se escutassem [gɔla]. O teste foi montado em cima dos seguintes pares mínimos:

<b>Contexto</b>	<b>Pares mínimos</b>
<b>[p,b] seguido de [a]</b>	[pala] - [bala]
<b>[p,b] seguido de [ɔ]</b>	[pɔtʃɪ] - [bɔtʃɪ]
<b>[t,d] seguido de [a]</b>	[tata <sup>6</sup> ] - [data]
<b>[t,d] seguido de [ɔ]</b>	[tɔka] - [dɔka]
<b>[k,g] seguido de [a]</b>	[kalo] - [galo]
<b>[k,g] seguido de [ɔ]</b>	[kɔla] - [gɔla]

**QUADRO 1:** Pares mínimos utilizados no teste de percepção.

Cada um dos seis pares mínimos foi utilizado em três frases, totalizando, assim, as 18 sentenças do teste. Além de trabalharmos com pares mínimos (que nos possibilitariam duas frases), manipulamos os VOTs da palavra-alvo para gerar uma terceira palavra (e, por conseguinte, uma terceira frase) para cada par mínimo. Assim, a partir da palavra-alvo do par que continha a oclusiva surda, realizamos a manipulação de seu VOT, encurtando-o para obtermos um VOT próximo de zero. Obtivemos, assim, mais um valor de VOT (primeiro valor); acrescentando este com o VOT positivo (segundo valor) e o VOT negativo (terceiro valor), tínhamos os três valores que queríamos para o teste. Optamos por manipular o segmento surdo, e não o sonoro, pelo fato de que permaneceriam resquícios de sonoridade se apenas cortássemos o pré-vozeamento do segmento sonoro. Como exemplo, demonstraremos, no que segue, o que fizemos com o par mínimo (‘pala’/‘bala’), para facilitar a compreensão do procedimento.

Primeiramente, solicitamos que um locutor aposentado, com mais de 20 anos de trabalho em rádios da cidade de Rio Grande, gravasse as seguintes frases, em

um ambiente silencioso<sup>7</sup>: ‘A palavra é... pala’ e ‘A palavra é... bala’. Feita essa gravação, já tínhamos duas palavras-alvo<sup>8</sup> para nosso teste: uma com o VOT de 12ms (‘pala’) e uma com o VOT com vibração antes da barra de plosão (‘bala’). A partir deste ponto, utilizamos o VOT de ‘pala’ (12ms) para a manipulação<sup>9</sup> do outro valor desejado. Optamos por utilizar um VOT zero<sup>10</sup>, através do apagamento do total do VOT formado, para ver qual das figuras os participantes iriam marcar, a partir deste estímulo auditivo. O vozeamento da vogal seguinte, portanto, nos VOTs por nós manipulados, começava junto à barra de soltura da oclusiva. Devido a esta inclusão, optamos por colocar a opção da letra ‘b’ em nosso teste (opção de dúvida).

As frases, depois de prontas, foram distribuídas de forma aleatória no teste, cuidando apenas para que os pares mínimos não aparecessem em sequência, de modo a evitarmos, por exemplo, a presença de um /p/ antes ou após um /b/. Procedemos, de forma semelhante, também com os outros pares: /t,d/ e /k,g/.

## 4. ANÁLISE DOS RESULTADOS

### 4.1 ANÁLISE DAS 12 FRASES SEM A MANIPULAÇÃO DE VOT

Na Tabela 1, que segue, apresentamos o número de erros de cada grupo, além dos padrões de VOT em que ocorreram:

**TABELA 1:** Comparação estatística dos erros cometidos pelos participantes dos três grupos analisados.

Grupos	MR	MP	B	Valor de p
Erros	23	28	19	
Média/dp de erros de A	1,80 (1,14)	1,80 (0,79)	1,20 (1,03)	0,3639 <sup>ns*</sup>
Média/dp de erros de C	0,50 (0,71)	1,00 (1,05)	0,70 (1,06)	0,5220 <sup>ns*</sup>

\*Estatística do teste Kruskal-Wallis não apontou diferença significativa entre os grupos no nível de 5%.

Como podemos perceber na Tabela 1, o grupo que cometeu o maior número de erros foi o MP (28), seguido do MR (23) e, por fim, do grupo B (19).

Na percepção do segmento surdo (alternativa 'A'), os dois grupos monolíngues empataram com médias iguais de erros, seguidos do grupo bilíngue. Já na percepção do segmento sonoro, o grupo MR foi o que cometeu o menor número de erros, seguido do grupo B e, por fim, do grupo MP. Como podemos perceber nas comparações feitas, não houve diferença significativa entre os grupos, nas médias de erros referentes à alternativa 'A' e à alternativa 'C'.

Como podemos perceber, todos os grupos cometeram mais erros na percepção do segmento surdo, marcando, por exemplo, a figura do 'galo' para a escuta da palavra [kalo]. A Tabela 2, que segue, possibilita observar em que palavras cada grupo cometeu seus erros de percepção.

**TABELA 2:** Palavras em que ocorreram os erros de percepção.

Palavras	pala	bala	pote	bote	tata	data	toca	doca	calo	galo	cola	gola	Total
MR	7	-	4	1	4	3	2	1	1	-	-	-	23
MP	7	1	4	1	4	2	-	5	2	-	1	1	28
B	6	-	3	2	-	-	1	3	2	1	-	1	19

Como podemos visualizar na Tabela 2, todos os três grupos tiveram mais dificuldade na percepção da palavra 'pala', provavelmente por esta ser uma palavra nova para todas as crianças. Estudos têm demonstrado que a *frequência* de uso de determinadas palavras na língua determina a velocidade do processamento de leitura. As palavras de baixa frequência são processadas mais lentamente e ensejam um número maior de erros do que as de alta frequência (SEIDENBERG, 2005). Podemos verificar, através do exposto, que é mais fácil escutar com acuidade uma palavra mais conhecida do que uma palavra recentemente aprendida, com a qual se tem pouca familiaridade. Por tal motivo, assinalamos de vermelho também as palavras 'tata' e 'doca'. Se fôssemos descartar as três palavras recentemente aprendidas pelos participantes (pala, tata e doca), teríamos este novo resultado: 11 (MR), 12 (MP) e 10 (B). As diferenças diminuiriam bastante, quase que tornando equivalentes os resultados nos três grupos.

A seguir, apresentamos os resultados das alternativas manipuladas. Queremos, primeiramente, enfatizar que tais achados não podem ser vistos com o mesmo olhar dos recém apresentados, pelo fato de terem sido manipulados, tratando-se, portanto, de dados artificiais.

#### 4.2 ANÁLISE DAS 6 FRASES COM VOT MANIPULADO

Almejavamos, com a manipulação, investigar a direção escolhida pelos alunos (mais para o surdo ou mais para o sonoro) frente a um segmento iniciado por VOT 0, artificialmente produzido. Vejamos os resultados:

**TABELA 3:** Comparação estatística das alternativas 'A' e 'C' perante os VOTs manipulados nos três grupos participantes.

Grupos	Resposta A	Resposta C	Valor de p
MR	2,10 (1,60)	3,30 (1,64)	0,0769
MP	3,00 (1,63)	2,00 (0,82)	0,1164
B	3,30 (1,06)	2,20 (1,03)	0,0301*

\* Estatística do teste Wilcoxon resultou significativa no nível de 5%

Os VOTs manipulados foram mais percebidos como surdos no seguinte ordenamento: B (média 3,30), MP (média 3,00) e MR (média 2,10). No que diz respeito ao julgamento como sonoro, o ordenamento foi o seguinte: MR (média 3,30), B (média 2,20) e MP (média 2,00). Como demonstrado na última coluna da Tabela 3, o único grupo que apresentou diferença significativa foi o grupo 'B'.

O interessante do resultado apresentado na Tabela 3 é que ambos os grupos de Picada Café marcaram em maior número a alternativa 'A' (surda) do que a alternativa 'C' (sonora), de modo oposto ao ocorrido com os alunos do grupo MR. Debateremos tais achados em nossa conclusão, apresentada a seguir.

#### CONCLUSÃO

Neste artigo, buscamos: a) Visualizar qual grupo de participantes cometeria o maior número de erros na percepção de questões sem manipulação de VOT; b) Verificar onde ocorreriam mais erros na percepção de questões sem manipulação de VOT, isto é, nos segmentos sonoros ou nos segmentos surdos, dos três grupos participantes; e c) Investigar, nas questões com manipulação de VOT, a direção escolhida pelas crianças (mais para o surdo ou mais para o sonoro) frente a um segmento iniciado por VOT 0, artificialmente produzido.

Quanto ao grupo que apresentou o maior número de erros (Questão 'A'), tivemos primeiramente o grupo MP (28), seguido do grupo MR (23) e, por fim, o grupo B (19). No entanto, ao excluirmos as palavras de baixa frequência de nosso teste, ficamos com resultados muito parecidos, fato que não pode ser desconsiderado, uma vez que a frequência das palavras atuou fortemente nos resultados.

Dos erros cometidos nas alternativas 'A' e 'C' (Questão 'B'), como pudemos verificar, o maior número ocorreu, nos três grupos participantes, nas respostas com alternativa 'A', como, por exemplo, na escolha da figura de um 'bote' para a escuta da palavra ['pɔtʃi]. A partir desta constatação, poderíamos levantar três possíveis explicações. A primeira delas seria a de que o valor produzido pelo locutor, '12ms', o mesmo descrito como média do segmento [p] do PB na literatura (REIS, NOBRE-OLIVEIRA, 2007), poderia ser muito pequeno quando comparado à média de produções de [p] (24ms) de nossos participantes, média essa verificada em outro trabalho (GEWEHR-BORELLA, 2010). Ao compararmos tais médias, percebemos que o valor realizado pelo grupo é o dobro do formado pelo locutor. Por tal razão, o pequeno valor de VOT poderia estar sendo percebido não como um segmento surdo, mas como um segmento sonoro (LIEBERMAN, BLUMSTEIN, 1988). Uma segunda provável explicação, somente para os alunos com algum contato com a língua de imigração Hunsrückisch, seria de que, em razão da ausência de aspiração na palavra, os participantes estariam percebendo-a também como sonora. Uma última alternativa poderia estar no maior número de palavras desconhecidas surdas (2) do que sonoras (1). É muito mais fácil perceber uma palavra conhecida do que uma desconhecida. Tal justificava foi comprovada na palavra ['pala], uma vez que 20

dos 80 erros cometidos (25%), no total dos três grupos, foram desta palavra. Vemos com este resultado a importância da *frequência* das palavras (SEIDENBERG, 2005).

Quanto aos resultados relativos às alternativas de VOT 0 (Questão 3), levantamos duas possíveis justificativas para um maior número de alternativas sonoras escolhidas pelo grupo MR e um maior número de alternativas surdas escolhidas pelos grupos MP e B. Diante do resultado encontrado, poderíamos dizer que as crianças de Rio Grande (monolíngues) optaram mais vezes pela alternativa sonora, por perceberem sonoridade apenas em um VOT curto, enquanto que as crianças de Picada Café (monolíngues e bilíngues) optaram mais pela alternativa surda, pelo fato de estarem mais habituadas à percepção dos segmentos sem pré-vozeamento. Ao falarem e ao perceberem o português brasileiro, poderíamos dizer que as crianças de Picada Café, principalmente as bilíngues, acabaram por transferir os padrões de sua L1 (Hunsrückisch) para sua L2 (português brasileiro). Baseados na abordagem da Fonologia Gestual, vemos, a partir do exposto, a percepção emergindo através de uma interação dinâmica dentro de um contexto social (BEST, TYLER, 2007).

Lembramos que a única diferença significativa encontrada nos testes de percepção ocorreu, com o grupo dos alunos bilíngues, frente aos valores de VOT 0 ( $p=0,0301$ ). Esse resultado pode indicar que realmente há alguma diferença na percepção de tais alunos, em comparação com os alunos dos outros dois grupos participantes.

Antes de concluirmos, precisamos apontar duas limitações referentes ao teste de percepção aplicado. A primeira diz respeito à realização de apenas um teste. Acreditamos ser necessária a aplicação de mais testes numa pesquisa futura. A segunda limitação diz respeito ao ambiente em que foi realizado o teste de percepção. As escolas de Picada Café e Rio Grande eram distintas com relação ao ruído externo, havendo um silêncio maior na primeira cidade, enquanto que, na segunda, o mesmo grau de silêncio não foi conseguido. Tal fato pode ter sido a causa de um número elevado de trocas nos alunos de Rio Grande.

Para finalizar, temos que lembrar que a interpretação dos resultados discutidos não deve seguir uma perspectiva de caráter estanque, uma vez que nossa percepção é refinada ao longo de nossas vidas (BEST, TYLER, 2007). A partir da visão dinâmica defendida neste artigo, temos que ter em mente que os

padrões criados dentro de um sistema estão em constante modificação. As manobras articulatórias, os chamados *gestos*, produzidos no trato vocal, e as manobras acústicas, que geram “atributos perceptuais relevantes para a diferenciação dos contrastes dos sons ocorridos na linguagem” (LIEBERMAN, BLUMSTEIN, 1988, p. 196), são observados através de um contínuo físico durante a realização de um determinado som da fala. Em oposição à visão binária da Fonologia Clássica, a partir da qual poderia ser concebida uma percepção de caráter categórico, modelos baseados em uma visão dinâmica da fala/percepção primam pelo destaque conferido à gradiência na produção dos segmentos. A partir dos dados observados, podemos concluir que a existência dessa gradiência pode ser considerada um dos fatores que influenciam a percepção tanto de alunos monolíngues quanto de alunos bilíngues, que apresentam resultados diferenciados por apresentarem, também, sistemas de origem distintos. Tais diferenças, apontadas ao longo do presente trabalho, reforçam a perspectiva dinâmica de aquisição dos sons, aqui defendida.

## **COMPARING THE PERCEPTION OF MONOLINGUAL AND BILINGUAL CHILDREN**

### **ABSTRACT**

In this study, we analyzed and compared the results of a perception test administered to 30 students, divided in three distinct groups: monolinguals from a city called Rio Grande, who are not in contact with speakers from another language (group MR); monolinguals from a town called Picada Café, who are in contact with bilingual speakers (group MP) and bilinguals from the same town (group B). By testing the perception of voiceless and voiced plosives in groups that have distinct production in these segments, we hope to find differences in the perception of the students investigated. Based on Gestual Fonology, we adopt a dynamic speech perception perspective. The test, developed by the authors, was divided into two parts: non-manipulated VOTs and manipulated ones. As for the data analysis, in the non-manipulated stimuli, we found more perception mistakes in the MP group,

followed by the MR and the B groups. From the perspective of a zero VOT – manipulated options- both groups from Picada Café chose option ‘A’ (voiceless), while the group from Rio Grande preferred option ‘C’ (voiced). A significant difference was only found in one of the analyses carried out. This leads us to the conclusion that there is a difference between the bilingual group and the other ones. We believe that the differences found in the groups from Picada Café exist in view of the contact of its participants with the Hunsrückisch immigration language, which is in accordance with the dynamic conception of language acquisition guiding this study.

**Keywords:** Perception of PB plosives. Voice Onset Time. Interlinguistic transfer.

## NOTAS

- <sup>1</sup> Mestre em Letras/Linguística Aplicada pela Universidade Católica de Pelotas (2010). Aluna especial do doutorado de Letras/Linguística Aplicada da Universidade Federal do Rio Grande do Sul.
- <sup>2</sup> Professora Adjunta do Programa de Pós-Graduação em Letras da Universidade Católica de Pelotas.
- <sup>3</sup> Professor Adjunto do Departamento de Línguas Modernas da Universidade Federal do Rio Grande do Sul.
- <sup>4</sup> Figura adaptada de Cohen (2004, p.13).
- <sup>5</sup> O conceito de não-linearidade refere-se ao comportamento não previsível, presente na teoria dos sistemas dinâmicos. Não há, neste tipo de aprendizado, uma relação direta de causa e efeito. Após uma informação ser recebida, ela é processada de diferentes formas, dependendo de diversas variáveis que atuam no processo de aprendizagem. A partir de uma modificação inicial, outras vão ocorrendo ao longo do processo, com distintos graus de efeito. Poderíamos citar como não-linear a aquisição de vocabulário, que é contínua ao longo da vida, sendo fortemente afetada na infância e tendo uma redução de impacto na fase adulta (ELMAN *et al.*, 1996; ADOLPH *et. al.*, 2008).
- <sup>6</sup> Tata: forma carinhosa de se chamar ‘pai’.
- <sup>7</sup> Especificações da gravação: 44.100 Hz ; 24 Bits; Stereo; realizada através do programa Sony Sound Forge 8.0
- <sup>8</sup> Apesar de termos gravado todas as 12 palavras dentro da frase veículo, utilizamos apenas uma gravação da frase veículo. A partir desta, fomos modificando as palavras-alvo.
- <sup>9</sup> Para tal manipulação, utilizamos o programa computacional Praat- versão 5.0.3.2 (BOERSMA, WEENINK, 2008).

- <sup>10</sup> Diferentemente da literatura, que classifica VOTs de até 15-20ms dentro da categoria chamada “zero”, nesta etapa do trabalho denominamos VOT zero o VOT por nós manipulado, isto é, sem um período de surdez entre a explosão e o início do vozeamento da vogal seguinte.

## REFERÊNCIAS

- ADOLPH, K.; et al. What is the shape of developmental change? *Psychological Review*. USA, v. 115, n. 3, p. 527-543, 2008.
- ALBANO, E. C. *O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro*. Campinas: Mercado de Letras, 2001. 272p.
- ALTENHOFEN, Cléo V. *Hunsrückisch in Rio Grande do Sul: Ein Beitrag zur Beschreibung einer deutschbrasilianischen Dialektvarietät im Kontakt mit dem Portugiesischen*. Stuttgart: Franz Steiner Verlag, 1996. 444p.
- BEST, C. T.; TYLER, M. D. Nonnative and second-language speech perception: commonalities and complementarities. In: BOHN, O. S.; MUNRO, M. J (Orgs.). *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins, 2007, 407p. p. 13-34.
- BOERSMA, P.; WEENINK, D. *PRAAT: doing phonetics by computer*. Versão 5.0.3.2 (programa computacional). Disponível em: <http://www.praat.org>. Acesso em: 23 ago. 2008.
- BORDEN, G. J.; HARRIS, K. S.; RAPHAEL, L. J. *Speech Science Primer: physiology, acoustics, and perception of Speech*. 3.ed. USA: Lippincott Williams e Wilkins, 1994. 319.
- BRAUN, A. Zur regionalen Distribution von VOT in Deutschen. In: \_\_\_\_\_. (Org.). *Untersuchungen zu Stimme und Sprache: Papers on Speech and Voice*, Germany, 1996. p. 19-32.
- BROWMAN, C.; GOLDSTEIN, L. Articulatory Phonology: An Overview. *Phonetica*. USA, n. 49, p. 155-180, 1992.
- COHEN, Gustavo Vargas. *The VOT Dimension: a bi-directional experiment with English and Brazilian-Portuguese stops*. 2004. 70 f. Dissertação (Mestrado em Letras/Inglês e Literatura correspondente) – Faculdade de Letras. Universidade Federal de Santa Catarina: Florianópolis, 2004.
- GOLDSTEIN, L. et al. Dynamic action units slip in speech production erros. *Cognition*. USA, n. 103, p. 386-412, 2007.
- GEWEHR-BORELLA. *A influência da fala bilíngue Hunsrückisch- Português brasileiro na escrita de crianças brasileiras em séries iniciais*. 2010. 205f.

Dissertação (Mestrado em Letras/Linguística Aplicada)- Pós-Graduação em Letras. Universidade Católica de Pelotas: Pelotas, 2010.

GAMA, M. R. *Percepção da fala: uma proposta de avaliação qualitativa*. São Paulo: Pancast, 1994. 99p.

ELMAN, J. et al. *Rethinking innateness: A connectionist perspective on development*. MIT Press, London, 1996.

JESSEN, M.; RINGEN, C. Laryngeal features in German Phonology. *Phonology, USA*, n. 19, p. 189-218, 2002.

LIEBERMAN, P.; BLUMSTEIN, S. E. *Speech physiology, speech perception, and acoustic phonetics*. USA: Cambridge University Press, 1988. 249p.

MÜLLER, T. L. *1824 antes e depois: o Rio Grande do Sul e a imigração alemã*, Porto Alegre: Metrópole, 1999. 24p.

MUNRO, M. J.; BOHN, O. S. The study of second language speech. In: BOHN, O. S.; MUNRO, M. J. (Orgs.) *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2007, 407 p. p. 3-11.

REIS, M; NOBRE-OLIVEIRA, D (2008) Effects of perceptual training on the identification and production of English voiceless plosives aspiration by Brazilian EFL learners. In: RAUBER, A. S.; BAPTISTA, B. O.; WATKINS, M.A. (Eds.), *New Sounds: FIFTH INTERNATIONAL SYMPOSIUM ON THE ACQUISITION OF SECOND LANGUAGE SPEECH*. 2007. *Anais...* Florianópolis, SC: UFSC, 2007. p. 398-407.

SANCIER, M. L.; FOWLER, C. A. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, USA*, n. 25, p. 421-436, 1997.

SEIDENBERG, M. S. Connectionist Models of Word Reading. *American Psychological Society*. Madison, USA, v. 14, n.5, p. 238-242, 2005.

WANG, Y.; BEHNE, D. Temporal remnants from Mandarin in nonnative English speech. In: BOHN, O. S.; MUNRO, M. J.(Orgs.) *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins, 2007, 407 p. p. 167-184.