

# Capítulo 17

## Produção e percepção de contrastes fônicos em crianças com distúrbio fonológico

Larissa C. Berti

Universidade Estadual Paulista – UNESP/Marília

Este capítulo mostra como a investigação da produção e percepção da fala em crianças com distúrbio fonológico pode fornecer evidências para elucidar questões sobre a unidade envolvida na representação/produção e percepção da fala. Assume-se a existência de uma unidade comum de representação, produção e percepção como sendo o gesto articulatório, conforme preconiza a Fonologia Gestual. Persegue-se a hipótese de que haveria uma correspondência gestual nos erros de produção e de percepção em crianças com o diagnóstico de distúrbio fonológico, na medida em que essas unidades compartilhariam a mesma unidade. Apresenta-se uma nova classificação dos erros de produção e percepção de fala com base no gesto articulatório. Discute-se, a partir do estudo realizado pela autora, quais indícios, advindos da realização de experimentos de produção e percepção de fala, nos permitiriam supor a unidade de representação fonológica. Destaca-se como implicação clínica a necessidade de se considerar as habilidades de produção e percepção (do outro e de si), tanto na avaliação quanto na intervenção de crianças com distúrbio fonológico.

**Palavras-chave:** produção de fala; percepção de fala; aquisição fonológica; fonética articulatória; fonologia gestual.

### 1 Introdução

O uso da linguagem por indivíduos de uma mesma comunidade linguística requer a existência de uma unidade comum para cumprir a sua função comunicativa.<sup>1</sup>

---

<sup>1</sup>A norma adotada na escrita deste capítulo foi a do português brasileiro.



Ou seja, uma mesma unidade deve ser compartilhada pelos indivíduos no plano da representação, da produção e da percepção da fala, de modo a cumprir a chamada “condição de paridade” (Liberman & Whalen 2000).

Um tema intrigante aos estudiosos da linguagem diz respeito a qual unidade linguística cumpriria a condição de paridade. A discussão sobre essa unidade comum tem emergido de pesquisas e de teorias sobre domínios que naturalmente são interligados, mas que são tradicionalmente estudados separadamente, a saber: produção, percepção e representação (Goldstein & Fowler 2003).

Uma abordagem teórica, que parte da premissa da existência de uma unidade comum que integra os domínios da representação, produção e percepção, refere-se à Fonologia Gestual (doravante FonGest) (Albano 2012, 2020), originalmente conhecida como Fonologia Articulatória (Browman & Goldstein 1992, Goldstein & Fowler 2003, Fowler & Galantucci 2005). Assume-se que a unidade comum compartilhada nesses diferentes domínios seja o chamado gesto articulatório.

Na próxima seção, apresentaremos a definição de gesto articulatório, bem como exporemos estudos que trazem evidências de que o gesto seria a unidade compartilhada nos domínios de representação, produção e percepção de fala.

## 2 Gesto articulatório: Unidade de representação, produção e percepção da fala

O gesto articulatório refere-se tanto a uma ação de constrição dos diferentes articuladores quanto à representação dessa manobra articulatória necessária para que essa ação se concretize. Nesse sentido, o gesto articulatório não está ligado apenas aos movimentos dos articuladores envolvidos, como o próprio nome sugere, mas também a uma oscilação abstrata da dinâmica de tarefa nas atividades de fala (Saltzman & Kelso 1987).<sup>2</sup>

Assume-se, nesta abordagem, que os gestos articulatórios se combinam sistematicamente para formar estruturas maiores, tais como os segmentos e as sílabas. No entanto, essa combinação gestual ocorre de modo mais elaborado do que simples sequências lineares, permitindo, conseqüentemente, que haja tanto uma sobreposição entre eles no tempo quanto uma variação de sua magnitude (em termos de amplitude de movimento), decorrentes de ativações parciais durante a produção da fala (Goldstein & Fowler 2003, Goldstein et al. 2006).

---

<sup>2</sup>A dinâmica de tarefa aplicada na fala envolve as ações coordenadas de vários articuladores. A preocupação está no movimento das variáveis do trato e não nos movimentos individuais realizados por determinados articuladores. Para maior aprofundamento, consultar as obras de Browman & Goldstein (1992) e Albano (2001, 2012).

Pressupõe-se, portanto, que os gestos articulatórios sejam as unidades selecionadas no curso da produção da fala, que, ao mesmo tempo, carregam unidades de ação e de informação, ou seja, trata-se de “caracterizações abstratas de eventos articulatórios com tempo intrínseco” (Browman & Goldstein 1992: 155-156).

No curso da percepção da fala, analogamente, presume-se que seja o gesto articulatório o objeto distal resgatado pelos ouvintes, uma vez que a análise do sinal físico seria insuficiente para marcar o processo de decisão perceptual, necessitando recorrer ao conhecimento sobre a produção da fala. Isso significa assumir, portanto, uma ligação direta entre os processos de produção de fala e a percepção (Fowler 1996, Goldstein & Fowler 2003).

Pelo fato de essas unidades informacionais e de ação incorporarem o simbólico e o concreto, não há necessidade de tradução entre os domínios da representação, produção e percepção nas tarefas de fala. Isto é, as unidades de representação/produção e percepção da fala não são destruídas ou distorcidas durante a sua produção e, posteriormente, reconstruídas na mente do ouvinte; ao contrário, o ouvinte depreenderia diretamente da fala os padrões gestuais ou os modos de ordenações gestuais, revelando as informações fonológicas (Goldstein & Fowler 2003).

Evidências empíricas de diferentes naturezas têm sido utilizadas para sustentar a premissa de que o gesto articulatório seja a unidade comum para a representação, produção e percepção (Rizzolatti & Arbib 1998, Studdert-Kennedy & Goldstein 2003, Rizzolatti & Craighero 2004, Pouplier & Goldstein 2005, Galantucci et al. 2006, Gentilucci & Corballis 2006, Goldstein et al. 2007, Pouplier 2008, McMillan & Corley 2010, Falahati Ardestani 2013).

Para suportar a alegação de que o gesto articulatório é a unidade selecionada no curso do processo de produção de fala (representação/produção), a FonGest ancora-se, fundamentalmente, nos achados empíricos reportados na literatura sobre os erros de fala (Pouplier & Goldstein 2005, Goldstein et al. 2007, Pouplier 2008, McMillan & Corley 2010, Falahati Ardestani 2013).

Essa série de estudos experimentais tem demonstrado, como uma tendência comum, que os erros de fala induzidos em adultos, julgados pelos ouvintes como sendo substituições segmentais, podem não ser decorrentes de um processo em que um segmento substitui o outro. Ao contrário, nas condições experimentais empregadas nessas pesquisas, um erro de produção, julgado auditivamente como uma substituição segmental, é interpretado em função dos diferentes modos pelos quais os gestos se sobrepõem no tempo e/ou em virtude de suas diferentes magnitudes de ativação, caracterizando os chamados erros gradientes.

Destaca-se, ainda, um estudo sobre erros da fala infantil em contexto de patologia que tem corroborado alguns achados descritos na análise dos erros de adultos

típicos de que muitas das chamadas substituições segmentais constituem, na verdade, intrusões gestuais e/ou ativações parciais dos gestos articulatórios (Gick et al. 2007). Os autores deste estudo investigaram, com o uso da ultrassonografia, os erros de produção (caracterizados tradicionalmente como substituição, omissão e distorção) envolvendo as líquidas do Inglês em três diferentes populações: duas crianças pequenas em processo de aquisição das líquidas; três adolescentes com problemas de produção nesses sons; e quatro adolescentes com perda auditiva também com problemas de produção nas líquidas. Primeiramente, os autores descreveram o padrão gestual identificado nas produções típicas das líquidas do Inglês, a saber: a presença simultânea de dois gestos articulatórios, um relativo à ponta da língua e o outro relativo ao dorso ou raiz de língua; posteriormente, analisaram os erros envolvendo esses sons. De forma geral, os autores relataram que os erros interpretados como substituições, omissões e distorções segmentais envolvem dois mecanismos de simplificações gestuais: a) omissão ou redução de magnitude de um dos gestos das líquidas, resultando auditivamente em uma substituição ou omissão; b) fusão gestual, resultando auditivamente em uma substituição ou distorção.

Por outro lado, achados, sobretudo de estudos neurofisiológicos, sobre a existência de um conjunto comum de mecanismos cerebrais que subjazem tanto às funções da percepção quanto da produção da fala têm fortemente sugerido a existência de uma unidade comum entre a produção e a percepção (Rizzolatti & Arbib 1998, Rizzolatti & Craighero 2004, Galantucci et al. 2006, Gentilucci & Corballis 2006, Pulvermüller et al. 2006). A maior parte desses achados diz respeito às evidências cerebrais de que a percepção da fala ativa representações de ações compatíveis com o input recebido; isso significa dizer que o sistema motor seria acessado *online* durante a percepção de fala (Galantucci et al. 2006).

Outro conjunto de evidências recai sobre a função dos neurônios espelhos, indicando a existência de um sistema de correspondência entre percepção e ação. Os neurônios espelhos referem-se a um pequeno subconjunto de neurônios (encontrados no córtex pré-motor e lóbulo parietal anterior de macacos), que são ativados tanto durante a execução de uma ação quanto na observação desta ação (Rizzolatti & Craighero 2004). Um sistema de neurônios espelhos em humanos tem sido proposto para destacar o papel fundamental da produção da fala na sua percepção, fornecendo um mecanismo neurofisiológico que cria a paridade entre falante e ouvinte (Rizzolatti & Arbib 1998, Gentilucci & Corballis 2006).

Em concordância com essa visão, Pulvermüller et al. (2006) identificaram a ativação de áreas cerebrais envolvidas no planejamento/ execução da produção da fala (como, por exemplo, giro frontal inferior esquerdo, córtex pré-motor ventral

e córtex motor primário) e na propriocepção dos movimentos da boca (córtex somato-sensorial) durante a percepção visual, audiovisual e auditiva da fala.

Porém, destaca-se que, embora as evidências empíricas utilizadas para sustentar os pilares da FonGest a respeito de uma unidade comum de representação, produção e percepção sejam bastante contundentes, elas advêm, fundamentalmente, de pesquisas com sujeitos adultos típicos e, ainda, em condições experimentais bastante específicas. Consequentemente, acredita-se não ser possível a completa generalização desses achados para outros contextos, como, por exemplo, o de aquisição e o de patologias da linguagem, uma vez que pesquisas dessa natureza são bastante escassas e com número reduzido de sujeitos. Essa generalização, pois, seria relevante considerando não somente o alcance explicativo do modelo teórico, mas também o fortalecimento do chamado “núcleo duro” deste paradigma científico (Lakatos 1979), cujo fio condutor mais visível é entender a linguagem como um sistema dinâmico (Albano 2020).

Dentre as patologias da linguagem, acredita-se que uma investigação conduzida, particularmente, com crianças que apresentam o diagnóstico de distúrbio fonológico possa ser um lugar estratégico para elucidar questões concernentes à unidade comum de representação, produção e percepção da fala. Explica-se o porquê.

Embora historicamente o distúrbio fonológico seja caracterizado, fundamentalmente, por um problema de representação e uso dos sons da fala, isto é, uma dificuldade primordialmente de ordem simbólica (Ingram 1997), estudos fonéticos recentes têm descrito a presença de um componente motor associado, ou seja, uma dificuldade de ordem motora. Especificamente, a literatura fonética mais recente tem descrito não somente a presença de padrão anormal de movimento em crianças com distúrbio fonológico (tais como gestos articulatórios indiferenciados (Gibbon 1999),<sup>3</sup> realocação do ponto articulatório (Gibbon & Wood 2002), valores cinemáticos alterados (Goozée et al. 2007), gestos articulatórios indissociados entre ponta e corpo de língua (de Oliveira & Berti 2018)), como também a presença de erros subfonêmicos ou contrastes encobertos (Scobbie et al. 2000, Munson et al. 2010, Berti et al. 2016).

Embora esses estudos citados não investiguem de modo concomitante aspectos da representação, produção e percepção da fala em crianças com distúrbios fonológicos, pode-se depreender desses achados empíricos que tais crianças apresentam problemas nesses três domínios.

---

<sup>3</sup>Faz-se necessário esclarecer ao leitor que a autora utiliza o termo “gesto articulatório indiferenciado” como sinônimo de movimento articulatório, concepção bastante distante da adotada no presente capítulo.

Assim sendo, ao assumir, tal como preconiza a FonGest, uma mesma unidade de representação, produção e percepção, seria esperado que haveria uma correspondência gestual nos erros de produção e de percepção em crianças com o diagnóstico de distúrbio fonológico.

Nesse sentido, a proposta contida no presente capítulo é investigar os erros de produção presentes na fala de crianças com diagnóstico de distúrbio fonológico, bem como investigar a percepção auditiva destes sujeitos em relação às próprias produções.

### **3 Metodologia**

#### **3.1 Participantes**

Participaram deste estudo dez crianças com diagnóstico de Distúrbio Fonológico, recrutadas do Centro de Estudos em Educação e Saúde (CEES) – UNESP/Marília, especificamente do Estágio Supervisionado em Terapia Fonoaudiológica: Foniologia Clínica. Os seguintes critérios foram satisfeitos para a seleção das crianças: ausência de alterações intelectuais e neurológicas; ausência de alterações anátomo-morfológicas que comprometam o processo de produção de fala (como, por exemplo, fissura lábio-palatina); ausência de alterações otológicas/auditivas; nenhuma história de terapia fonoaudiológica prévia. Foram utilizados protocolos de rastreio para a identificação das alterações de linguagem e alterações anátomo-morfológicas, enquanto para a triagem auditiva foi utilizado o audiômetro Interacoustic AD-28, com fones de ouvido TDH-39, no interior de uma cabine acústica. Investigamos as frequências de 1.000, 2.000 e 4.000 Hz, em uma intensidade de 20 dB NA (nível de audição).

#### **3.2 Experimento de produção**

Utilizámos o instrumento PERCEFAL (Berti 2017) para o levantamento dos dados de produção. Embora o instrumento referido não tenha sido construído com o propósito de se avaliar a produção de fala, optámos por utilizá-lo de modo a viabilizar a comparação e a correlação dos erros de produção e de percepção de fala (a partir da fala no adulto típico e das próprias produções das crianças), uma vez que teríamos as mesmas possibilidades de ocorrência de todos os fonemas analisados.

O PERCEFAL é composto por um subconjunto de quatro experimentos, porém utilizámos neste estudo somente os experimentos que envolviam os contrastes consonantais: o PerceOcl, que avalia a identificação entre os fonemas oclusivos,

a partir de 30 pares mínimos (6 oclusivas x 5); o PerceFric, que avalia a identificação entre as fricativas, considerando 30 pares contrastivos (6 fricativas x 5); e o PerceSon, que avalia a identificação entre as sonorantes, a partir de 42 pares contrastivos (7 sonorantes (3 nasais e 4 líquidas) x 6).

Os pares mínimos são apresentados por figuras, que foram utilizadas para eliciar a amostra linguística das crianças a partir de uma tarefa de nomeação, computando o total de 102 produções. Caso a criança não soubesse nomear a gravura disposta na tela do computador, lhe apresentávamos o modelo para que pudesse repetir.

As gravações foram realizadas com cada participante separadamente, em uma cabine acústica no Laboratório de Análise Acústica (LAAc) – CEES – UNESP/Marília. O tempo de duração de cada gravação foi de, aproximadamente, 30 a 40 minutos com cada criança. Caso ela se mostrasse cansada, a coleta seria interrompida e refeita em duas sessões distintas.

Para a realização dessas gravações, foram utilizados os seguintes equipamentos de alta fidelidade: microfone unidirecional, ultrassom portátil modelo DP 6600 com transdutor linear acoplado a um computador, além do estabilizador de cabeça Ultrasound Stabilisation Headset (ARTICULATE INSTRUMENTS LTD, 2008). Ambos os sinais – de áudio e de imagem – foram capturados com o uso do software Articulate Assistant Advanced – AAA (ARTICULATE INSTRUMENTS LTD, 2012) e sincronizados pela unidade Sync Bright Up (ARTICULATE INSTRUMENTS LTD, 2010), o que reduziu a dessincronização do sinal de áudio com a imagem. As imagens de ultrassom (US) foram adquiridas com uma frequência de imagem de 6.5 MHz, 120° de campo de imagem e 29.97 Hz de taxa de amostragem. Os dados gravados foram salvos individualmente e exportados em arquivos de áudio (.wav) e vídeo (.avi) separadamente.

### 3.3 Experimento de percepção

O experimento de percepção consistiu em dois testes distintos: o experimento de identificação de contrastes fonológicos a partir da produção de fala de um adulto típico, denominado “percepção no outro”; e o experimento de identificação de contrastes fonológicos a partir das próprias produções das crianças, denominado “percepção em si”.

Em ambos os testes, utilizamos o instrumento PERCEFAL (Berti 2017). O procedimento experimental consistiu, em linhas gerais, de uma tarefa de identificação que compreendeu três etapas distintas: reconhecimento das palavras do experimento (sondagem), fase de treino e fase de teste, com o uso do software PERCEVAL (Perception Evaluation Auditive & Visuelle) (André et al. 2009).

Na fase de reconhecimento, apresentámos as figuras do teste às crianças, a fim de familiarizá-las. Elas deveriam reconhecer pelo menos 80% das figuras. Caso as crianças não reconhecessem pelo menos 80% das figuras elas eram excluídas do estudo.

A fase de treino visou garantir a compreensão da tarefa. Seleccionámos, de modo aleatório, dez estímulos para que a criança identificasse a gravura correspondente ao estímulo auditivo apresentado, dentre duas possibilidades dispostas na tela do computador. Porém, as respostas não foram contabilizadas pelo *software*. Repetimos o mesmo procedimento na fase de teste com todos os estímulos dos experimentos, e todas as respostas foram consideradas pelo *software*.

Para a realização dos testes de percepção, as crianças foram dispostas em frente à tela do computador, com fones KOSS acoplados aos seus ouvidos, no interior de uma cabine acústica.

Cabe ressaltar que, no teste de identificação, a partir da fala do outro, foram utilizados os estímulos auditivos correspondentes às gravações em áudio de todas as palavras do PERCEFAL, por um falante adulto típico; enquanto no teste de identificação, em si, foram utilizados estímulos auditivos correspondentes às edições das produções de cada criança das palavras do PERCEFAL (utilizadas no experimento de produção). Ou seja, todas as gravações foram editadas em arquivos individuais e substituídas no *script* do experimento de percepção, de modo que, no teste de percepção, em si, cada criança precisava identificar os contrastes a partir, apenas, das próprias produções.

Em ambos os testes, os padrões de respostas aceitos foram: acerto (quando a criança identificava o estímulo corretamente), erro (quando não o identificava corretamente) e não resposta (quando não apresentava nenhuma resposta durante o período determinado). No caso do teste de percepção em si, considerámos acerto em caso de a criança ser capaz de reconhecer a própria produção, independentemente de essa produção ter sido julgada, na tarefa de produção, como alvo ou substituição. Por exemplo, se a criança apresentasse um erro de produção substituindo /b/ → [p], a produção da palavra /'bɒti/, por exemplo, seria [ˈpɒtʃi], considerada incorreta. Porém, na tarefa de identificação em si, se a criança, após se escutar, indicasse a figura correspondente ao “pote”, sua resposta seria computada como acerto.

Tanto o tempo de apresentação dos estímulos auditivo e visual quanto o tempo de resposta das crianças foram controlados e mensurados automaticamente pelo *software* PERCEVAL.

Portanto, cada criança era submetida a dois testes de percepção: um de identificação na fala do outro, envolvendo o subconjunto de três experimentos (PerceOcl, o PerceFric e PerceSon); e outro de identificação em sua própria fala, consistindo,



analogamente, de três experimentos (PerceOcl, o PerceFric e PerceSon). Os dois testes foram aplicados em cada criança em dois a três dias separadamente, para se evitar que a exaustão prejudicasse os resultados.

### 3.4 Tratamento dos dados

Uma vez coletados os dados de produção e de percepção, houve a necessidade de realizar o processamento dos mesmos.

Os dados resultantes da tarefa de produção de fala foram exportados em dois arquivos individuais: um arquivo de áudio no formato .wav e um arquivo de vídeo no formato .avi. A análise dos dados de produção de fala foram feitos a partir do julgamento-perceptivo auditivo com o apoio da análise acústica (inspeção do oscilograma e espectrograma) e do julgamento perceptivo-visual das imagens ultrassonográficas, com apoio da análise articulatória (inspeção do frame ultrassonográfico correspondente ao ponto máximo de constrição da língua na produção do fonema alvo).

Os resultados dos testes de percepção foram salvos automaticamente em arquivos .txt e, posteriormente, organizados em planilha do excel.

### 3.5 Forma de análise

Com base em estudos anteriores (Goldstein 2003, van Lieshout & Goldstein 2008), que utilizaram o gesto articulatório como sendo a unidade de análise, fizemos uma interpretação dos erros a partir do uso dos descritores gestuais propostos pela Fonologia Gestual, considerando as seguintes categorias:

- a) erro de local e/ou de constrição gestual: caracteriza-se pela inadequação quanto ao local e/ou a forma em que a constrição gestual é realizada, tal como ocorre, por exemplo, nas substituições de /s/ → [ʃ], /t/ → [k]; ou quando ocorre nas substituições entre as líquidas coronais (/ɾ/ → [l]).<sup>4</sup>
- b) erro de ativação do gesto laríngeo: refere-se aos diferentes graus de ativação laríngea na produção dos segmentos vozeados julgados, em geral, como desvozeados ou surdos. Nesses casos, o tempo de ativação do gesto laríngeo está reduzido ou ausente.

---

<sup>4</sup>Uma vez que as líquidas são produzidas por duplos gestos linguais (um de ponta e outro de língua) (da Silva Barberena et al. 2014, de Castro Nunes de Lima et al. 2018), as substituições das líquidas coronais podem ser caracterizadas pelo deslocamento do local de constrição de um dos gestos (em geral do gesto de corpo de língua) ou pela mudança da forma de constrição da ponta da língua (Esperandino 2020).

- c) erro de grau de constrição gestual: ocorre quando há adequação do gesto que envolve a língua (seja o gesto de ponta, corpo ou raiz de língua), porém há uma inadequação de ativação quanto à magnitude deste gesto, como é o caso das substituições de fricativas por oclusivas (tal como a substituição de /s/ → [t]).
- d) erro de grau e local de constrição gestual: caracteriza-se por mudanças gestuais que ocorrem tanto em relação ao grau de ativação da magnitude gestual quanto ao local em que se deveria ocorrer a constrição gestual, o que levaria, conseqüentemente, a uma mudança quanto à forma de constrição do gesto. São erros que envolvem, tradicionalmente, o ponto e o modo de articulação. Um exemplo para ilustrar esse tipo de erro seria a substituição de /ʌ/ → [y].

A Tabela 1 exemplifica a interpretação dos erros a partir de uma análise gestual.

Tabela 1: Interpretação dos erros a partir de uma análise gestual (Berti 2019: 90)

Exemplos de erros	Interpretação gestual
/t/→[k]; /s/→[ʃ]; /r/→[l]	Erro de local e/ou forma de constrição gestual
/b/→[p]; /v/→[f]	Erro de ativação do gesto laríngeo
/s/→[t]	Erro de grau de constrição gestual
/ʃ/→[t]; /ʌ/→[y]	Erro de grau e de local/forma de constrição gestual

## 4 Resultados

A Tabela 2, a seguir, apresenta a distribuição do padrão de erro gestual em cada uma das habilidades, conforme as seguintes categorias: erros de local e/ou forma de constrição; erros de ativação do gesto laríngeo; erros de grau de constrição; e erros de grau e de local/forma de constrição gestual.

Em relação à análise do padrão de erros em função das habilidades, a ANOVA de Medidas Repetidas mostrou um efeito significativo somente para o padrão de erros ( $F(3,45)=9,65$ ,  $p>0,00$ ), porém não revelou um efeito significativo para as habilidades ( $F(2,15)=0,58$ ,  $p=0,57$ ), nem para a interação entre padrão de erros e as habilidades ( $F(6,45)=1,94$ ,  $p=0,09$ ).

A análise *post hoc* com o uso do teste Bonferroni mostrou que os erros envolvendo o local de constrição e os erros envolvendo local + grau/forma de constrição foram os mais frequentes ( $p<0,05$ ), comparativamente aos erros de ativação

Tabela 2: Distribuição da média de erros gestuais de produção e de percepção (no outro e em si) (Berti 2019: 96)

Habilidades	Padrão de Erro Gestual (Média e desvio-padrão)							
	Local e/ou forma de constrição		Ativação gesto laríngeo		Grau de constrição		Grau/local e forma de constrição	
Produção	4,8	±3,45	4,6	±3,91	6	±6,54	5,2	±7,82
Percepção no outro	5,88	±4,25	3,42	±1,13	2,57	±2,22	5,1	±4,62
Percepção em si	10,8	±3,85	5,7	±2,11	2,44	±1,33	11,2	±4,96

do gesto laríngeo e de grau de constrição, os quais não se diferenciaram entre si ( $p>0,05$ ). As Figuras 1 e 2 ilustram os resultados obtidos.

Adicionalmente, fizemos uma análise de correlação entre as habilidades, com o uso do Coeficiente de Concordância de Kendall, considerando cada tipo de erro gestual, separadamente. O teste de Kendall é similar ao teste de correlação de Spearman, exceto pelo fato de expressar a correlação entre múltiplas variáveis.

Observamos a presença de uma correlação positiva significativa ( $p<0,05$ ) entre os erros de local de constrição gestual e os erros de ativação do gesto laríngeo, com um coeficiente de Concordância de 0,53 para os erros de local de constrição e 0,31 para os erros de ativação do gesto laríngeo. Nos erros envolvendo o grau de constrição, a correlação mostrou-se significativa ( $p<0,05$ ) apenas entre as habilidades de percepção no outro e de produção, com um coeficiente de Concordância de 0,60. Para os erros envolvendo o local/grau e forma de constrição, não houve correlação entre nenhuma das habilidades.

Em síntese, na reinterpretação dos erros que ocorreram nas habilidades de produção, percepção no outro e percepção em si, a partir das categorias gestuais, não houve diferença entre tais habilidades, e, ainda, observamos, no geral, a presença de correlação das três habilidades, exceto para os erros que envolveram o local/grau e forma de constrição.

Na próxima seção, apresentaremos a discussão dos resultados obtidos.

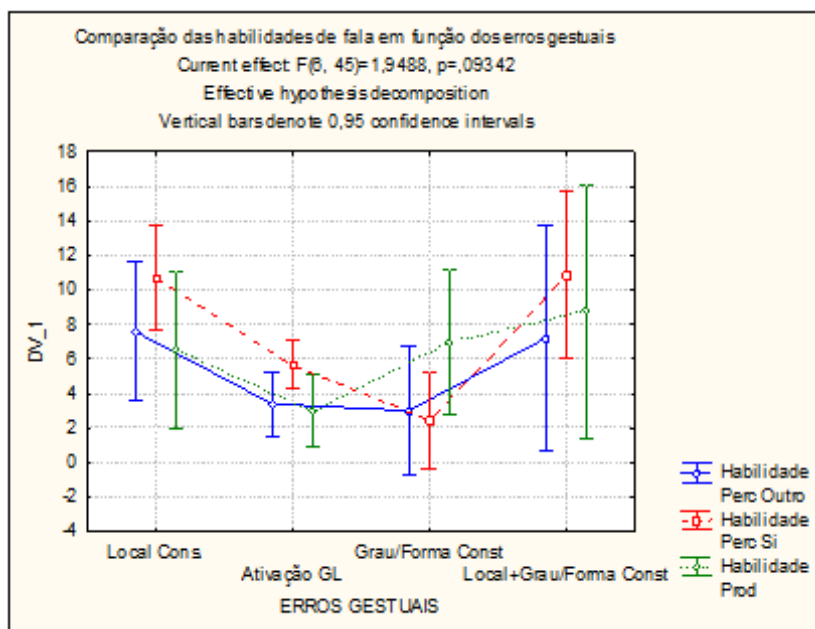


Figura 1: Comparação das habilidades de fala em função dos erros gestuais (Berti 2019: 97).

## 5 Discussão

O presente capítulo apresentou uma investigação dos erros de produção e percepção de fala em crianças com diagnóstico de distúrbio fonológico. Ao assumir, conforme preconiza a FonGest, uma mesma unidade de representação, produção e percepção, tal como o gesto articulatório, seria esperado uma correspondência gestual nos erros de produção e de percepção em crianças com o diagnóstico de distúrbio fonológico.

Os erros foram analisados de acordo com as seguintes categorias gestuais: erros de local e/ou forma de constrição; erros de ativação do gesto laringeo; erros de grau de constrição; e erros de grau e de local/forma de constrição gestual.

Na comparação entre os erros não observamos diferença entre as habilidades de percepção no outro, percepção em si e produção, conforme evidenciado pelo não efeito significativo das habilidades e da interação entre habilidades\*padrão de erro (resultados obtidos pela ANOVA de Medidas Repetidas).

Esses resultados concordaram integralmente com a previsão de que haveria correspondência entre os erros nas habilidades de produção e percepção de fala

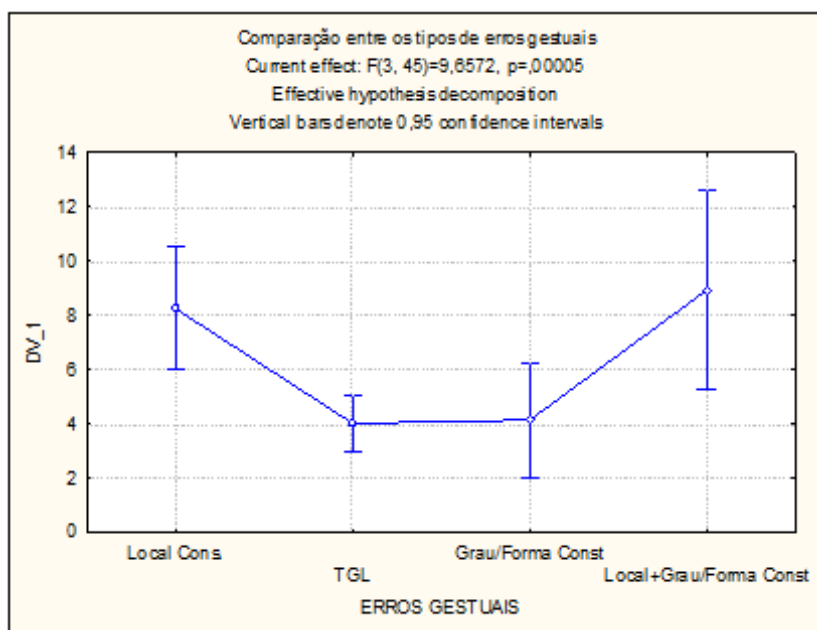


Figura 2: Comparação entre os tipos de erros gestuais (Berti 2019: 97)

(no outro e em si), na medida em que compartilhariam a mesma unidade – o gesto articulatório.

Lieberman & Whalen (2000) utilizam o termo “paridade” para se referirem a três requisitos da linguagem a serviço de sua função comunicativa. Os dois primeiros requisitos relacionam-se à comunicação entre as pessoas: o primeiro prevê que tanto o falante quanto o ouvinte devem contar com a mesma forma de linguagem para viabilizar a comunicação; o segundo requisito diz respeito à equivalência na forma da linguagem enviada e recebida. Ou seja, para a linguagem servir a propósitos comunicativos, ouvintes precisam perceber acuradamente o que os falantes produzem; portanto, deve haver suficiente equivalência entre a mensagem fonológica enviada e recebida. Finalmente, o terceiro requisito refere-se a uma questão filogenética: a especialização para produção e para a percepção da fala codesenvolveram-se, uma vez que nenhuma especialização seria útil sem a outra.

Pressupomos, portanto, que os gestos articulatórios sejam as unidades selecionadas no curso da produção da fala, que, ao mesmo tempo, carregam unidades de ação e de informação, ou seja, são “*caracterizações abstratas de eventos articulatórios com tempo intrínseco*” (Browman & Goldstein 1992: 155-156).

Pelo fato de o gesto articulatório (unidade informacional e de ação) incorporar o simbólico e o motor, parece não haver necessidade de tradução entre os domínios da representação, produção e percepção nas tarefas de fala.

A equivalência de erros entre as habilidades de produção e percepção da fala (no outro e em si), em crianças com distúrbio fonológico, portanto, constitui, a nosso ver, uma forte evidência para a presença de uma única unidade comum compartilhada nos domínios da produção/percepção e representação. Esta unidade, conforme preconizada pela Fonologia Gestual, seria, supostamente, o chamado gesto articulatório.

Outro resultado que fortalece as premissas da Fonologia Gestual refere-se à prevalência de erros envolvendo local+grau/forma de constrição juntamente com erros de local de constrição.

Nas primeiras versões da Fonologia Articulatória, Browman & Goldstein (1989) descreveram que, durante o processo de aquisição da linguagem, há pelo menos dois desenvolvimentos necessários para a criança atingir o padrão apresentado na fala adulta: a) diferenciação e ajustamento de gestos articulatórios individuais; b) coordenação deles na palavra. A diferenciação e ajustamento ocorrem a partir de um único gesto individual, mais geral, que servirá para produzir oclusivas, fricativas e aproximantes. Já a coordenação desses gestos na palavra significa que não basta a criança conseguir produzir todos os gestos de uma palavra, é preciso também que ela os coordene. Mais recentemente, Namasivayam et al. (2020: 12) descreveram o padrão de erros de produção da fala em crianças com distúrbio fonológico à luz da Fonologia gestual interpretando-os como sendo decorrentes de dificuldades na coordenação inter gestual; na coordenação inter articulador, envolvendo a seleção correta das variáveis de trato; na especificação correta do grau de constrição dos gestos; no planejamento dos osciladores gestuais; ou ainda, a combinação entre eles.

Posteriormente, em relação ao desenvolvimento de gestos articulatórios, Goldstein (2003) afirmou que as crianças aprendem, num período inicial do controle motor da fala, a diferenciação entre articuladores (nos termos do autor “*between-organ contrast*”), como, por exemplo: lábios vs. língua. Em um segundo momento, há um refinamento do controle neuromotor e, assim, as crianças aprendem a diferenciação intra-articuladores (nos termos desse autor “*within-organ contrast*”), como, por exemplo, ponta da língua vs. corpo da língua. Depois, na medida em que as crianças vão sendo capazes de diferenciar e movimentar os principais articuladores independentemente, elas passam ao ajuste do grau e do local de constrição desses articuladores (Goldstein 2003: 86).

Na reanálise de erros de fala de um banco de dados a partir de descritos gestuais, Studdert-Kennedy (2002) e Goldstein (2003) mostraram que crianças acer-

tavam (em uma frequência acima da ocorrência ao acaso) qual o órgão do trato vocal envolvido na produção do gesto articulatório que iniciava a palavra-alvo. Isto é, os erros compartilhavam o órgão correto com a forma adulta e diferiam em algumas outras propriedades, geralmente, em grau de constrição ou local de constrição. Algumas crianças também mostraram correspondência do gesto glotal e vélico com os alvos adultos. No entanto, nenhuma criança mostrou correspondência de diferenciação intra-articulador e do grau de constrição.

Goldstein et al. (2007: 224) reafirmam a previsão de uma estrutura gestual emergente no processo de aquisição da linguagem:

[...] children should acquire between-organ contrasts earlier than within-organ contrasts because organ differentiation requires that the infant must attune to her language environment.

Os resultados obtidos pelos autores citados acima, tomados juntos, evidenciam uma estrutura hierárquica gestual que implica um domínio, por parte das crianças, de diferentes graus de complexidade de coordenação gestual.

Portanto, os resultados dos testes de comparação (entre a ocorrência dos erros nas três habilidades e entre os tipos de erros gestuais mais frequentes) corroboraram integralmente a previsão de que haveria uma correspondência gestual nos erros de produção e de percepção em crianças com o diagnóstico de distúrbio fonológico.

Entretanto, os resultados obtidos na análise de correlação requerem uma discussão mais complexa. A análise da correlação dos erros gestuais, a partir de cada categoria adotada, mostrou uma correlação positiva significativa entre as habilidades de produção e de percepção (em si e no outro) para os erros de local de constrição gestual e para os erros de ativação do gesto laríngeo. Nos erros envolvendo o grau de constrição, a correlação mostrou-se significativa apenas entre as habilidades de percepção no outro e de produção, enquanto, para os erros envolvendo o local/grau e forma de constrição, não houve correlação entre nenhuma das habilidades.

A natureza da relação entre as habilidades de produção e percepção (em si e no outro) mostrou-se mais complexa do que uma relação de natureza isomórfica em que observamos uma correspondência termo a termo.

Conforme anunciado, uma vez assumido que os gestos são unidades fonológicas, esperávamos que eles desempenhassem um papel tanto na produção de fala quanto na percepção da fala (em si e no outro). De fato, a existência de uma correlação positiva entre as habilidades para a maior parte das categorias dos erros gestuais adotados sugere a existência de uma unidade comum compartilhada nesses domínios.

Entretanto, para os erros mais complexos do ponto de vista gestual, ou seja, aqueles envolvendo concomitantemente os ajustes de grau e de local de constrição, não houve qualquer correlação entre as habilidades de produção e percepção (no outro e em si).

Uma possibilidade explicativa para esses resultados “aparentemente” contraditórios advém de estudos que têm buscado investigar as respostas adaptativas ao *feedback* auditivo perturbado na produção da fala (Shiller et al. 2010, Casserly 2011, Chang et al. 2013, de Boer & Bressmann 2017). Esses estudos também têm procurado investigar a relação entre a produção e percepção de fala de uma forma diferente daquela aqui apresentada. Particularmente, busca-se nesses estudos descrever o efeito na produção da fala decorrente da alteração do *feedback* auditivo.

De forma geral, os resultados encontrados têm mostrado uma mudança na produção da fala em resposta ao *feedback* alterado, porém a resposta compensatória no âmbito da produção da fala é apenas parcial, apresentando grande variação de indivíduo para indivíduo e, ainda, variando nas condições testadas.

A compensação parcial vista na produção em decorrência à modificação do *feedback* não é interpretada nestes estudos como uma falta de evidência para uma estreita relação entre produção e percepção. Ao contrário, uma das possibilidades explicativas encontradas nesses estudos refere-se à natureza do *feedback*.

Chang et al. (2013) argumentam que há dois tipos de *feedback* sensorial: o auditivo e o somato sensorial. Depreendemos do trabalho citado que o *feedback* auditivo refere-se ao monitoramento auditivo de estímulos acústicos advindos do ambiente e/ou da própria fala, enquanto o *feedback* somato sensorial diz respeito à condição que permite ao sujeito experimentar sensações advindas das diferentes partes de seu corpo. Para os autores, a compensação parcial é decorrente do conflito de informações, do ponto de vista neural, advindas do *feedback* auditivo e do somato sensorial. Adicionalmente, os autores destacam a dificuldade de se controlar o *feedback* nos sistemas biológicos, uma vez que este ocorre de forma atrasada em resposta ao estímulo recebido, em função do tempo das sinapses no sistema nervoso.

Portanto, dado ao ajuste de processamento de informações neurais advindas tanto do *feedback* auditivo quanto do *feedback* somato sensorial, em nível cortical, a adaptação ou a compensação motora pode ocorrer de forma parcial.

Embora no presente capítulo não tenhamos analisado as adaptações da produção da fala em função da manipulação do *feedback* auditivo, poderíamos supor que, nas tarefas de percepção (no outro e em si), as informações acústicas e as informações somato sensoriais extraídas dos gestos articulatórios (objeto distal da percepção) estariam sendo processadas de modo concorrente em nível cortical, resultando em uma resposta não correspondente ao input auditivo (seja da fala



do outro, seja da própria produção), explicando, portanto, a não correlação entre as habilidades para todos os erros.

Outra possibilidade explicativa para a não correlação dos erros gestuais envolvendo grau e local de constrição poderia estar relacionada ao próprio fenômeno perceptual, de modo mais geral. A percepção da fala tem sido considerada um evento auditivo-visual, que envolve a integração de pistas auditivas e visuais em uma entidade fonológica unitária (Dodd et al. 2008).

Há gestos, como os de corpo e raiz da língua, laringe e de véu palatino, que são (em grande parte) invisíveis, sendo comunicados acusticamente e envolvendo, predominantemente, o sistema auditivo do receptor. Por outro lado, para gestos potencialmente visíveis, tais como os de lábio e ponta de língua, informações ópticas e auditivas são combinadas em um único percepto envolvendo, supostamente, a integração de pistas sensório-motoras.

Uma vez que erros gestuais de local+grau de constrição envolvem maior complexidade gestual, seja pela presença de gestos linguais concomitantes (como é o caso das líquidas), seja pelos diferentes graus de informações óticas, presumimos que ocorram ajustes no processamento cerebral desses diferentes tipos de informações, o que poderia resultar em uma resposta não correspondente ao estímulo apresentado às crianças.

Considerar na análise da percepção da fala (tanto no outro quanto em si) a sua natureza multimodal em estudos futuros parece ser necessário e relevante.

## 6 Conclusão

A investigação conduzida sobre os erros de produção e percepção de fala em crianças com distúrbio fonológico mostrou-se, por um lado, como um lugar estratégico para elucidar questões concernentes à unidade comum de representação, produção e percepção da fala, e, por outro lado, suscitou algumas implicações clínicas.

A partir do gesto articulatório, a análise dos erros ocorridos nas habilidades de produção e percepção (no outro e em si) em crianças com distúrbio fonológico trouxe fortes evidências para sustentar as premissas da Fonologia Gestual de que a unidade compartilhada nesses domínios seria o gesto articulatório.

Em termos de implicações clínicas, o fato de os erros apresentados pelas crianças com distúrbio fonológico abrangerem, igualmente, a produção, a percepção no outro e a percepção em si, de modo equiparado (sem diferença entre a ocorrência de cada uma das habilidades) e correlacionado (pelo menos para a maior parte dos tipos de erros), constitui forte evidência de que essas habilidades mantêm uma estreita relação entre si. Isso significa dizer que ações clínicas, tanto na

avaliação quanto na intervenção, com crianças com distúrbio fonológico, devam ser realizadas considerando as três habilidades: produção, percepção da fala do outro e percepção de sua própria fala.

## Referências

- Albano, Eleonora Cavalcanti. 2001. *Gestos e suas bordas: Esboço de fonologia acústico-articulatória do português brasileiro*. São Paulo: Mercado das Letras.
- Albano, Eleonora Cavalcanti. 2012. Uma introdução à dinâmica em fonologia, com foco nos trabalhos desta coletânea. *Revista da ABRALIN* 11(1). 1–30. DOI: 10.5380/rabl.v11i1.32462.
- Albano, Eleonora Cavalcanti. 2020. *O gesto audível: Fonologia como pragmática*. São Paulo: Cortez.
- André, Carine, Alain Ghio, Christian Cavé & Bernard Teston. 2009. PERCEVAL: Perception Evaluation Auditive & Visuelle v. 5.0.30. (Programa de computador).
- Berti, Larissa Cristina. 2017. PERCEVAL: Instrumento de avaliação da identificação de contrastes fonológicos. *Audiology-Communication Research* 22. e1727. DOI: 10.1590/2317-6431-2016-1727.
- Berti, Larissa Cristina. 2019. *Produção e percepção de fala em crianças com alterações dos sons da fala: Indícios da unidade de representação fonológica*. Marília, SP: Universidade Estadual Paulista (Unesp), Faculdade de Filosofia. (Tese de Livre Docência).
- Berti, Larissa Cristina, Gillian de Boer & Tim Bressmann. 2016. Tongue displacement and durational characteristics of normal and disordered Brazilian Portuguese liquids. *Clinical Linguistics & Phonetics* 30(2). 131–149. DOI: 10.3109/02699206.2015.1116607.
- Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology Yearbook* 6(2). 201–251. DOI: 10.1017/S0952675700001019.
- Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49(3-4). 155–180. DOI: 10.1159/000261913.
- Casserly, Elizabeth D. 2011. Speaker compensation for local perturbation of fricative acoustic feedback. *The Journal of the Acoustical Society of America* 129(4). 2181–2190. DOI: 10.1121/1.3552883.
- Chang, Edward F., Caroline A. Niziolek, Robert T. Knight, Srikantan S. Nagarajan & John F. Houde. 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences* 110(7). 2653–2658. DOI: 10.1073/pnas.1216827110.

- da Silva Barberena, Luciana, Márcia Keske-Soares & Larissa Cristina Berti. 2014. Descrição dos gestos articulatórios envolvidos na produção dos sons /r/ e /l/. *Audiology-Communication Research* 19(4). 338–344. DOI: 10.1590/S2317-6431201400040000135.
- de Boer, Gillian & Tim Bressmann. 2017. Influence of altered auditory feedback on oral–nasal balance in speech. *Journal of Speech, Language, and Hearing Research* 60(11). 3135–3143. DOI: 10.1044/2017\_JSLHR-S-16-0390.
- de Castro Nunes de Lima, Fernanda Leitão, Cássio Eduardo Esperandino da Silva, Lídia Maurício da Silva, Aline Mara de Oliveira Vassoler, Eliana Maria Gradim Fabbron & Larissa Cristina Berti. 2018. Ultrasonographic analysis of lateral liquids and coronal fricatives: Judgment of experienced and non-experienced judges. *Revista CEFAC* 20(4). 422–431. DOI: 10.1590/1982-0216201820412317.
- de Oliveira, Aline Mara & Larissa Cristina Berti. 2018. Aquisição fonológica típica e atípica do padrão silábico CCV: Dados acústicos e articulatórios. *Alfa: Revista de Linguística* 62(3). 591–612. DOI: 10.1590/1981-5794-1811-7.
- Dodd, Barbara, Beth McIntosh, D. Dogu Erdener & Denis Burnham. 2008. Perception of the auditory-visual illusion in speech perception by children with phonological disorders. *Clinical Linguistics & Phonetics* 22(1). 69–82. DOI: 10.1080/02699200701660100.
- Esperandino, Cássio Eduardo. 2020. *Análise ultrassonográfica qualitativa e quantitativa das produções das líquidas alveolares de crianças falantes do português brasileiro*. São Paulo: Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências. (tese de mestrado). <https://repositorio.unesp.br/handle/11449/192682> (30 novembro, 2021).
- Falahati Ardestani, Reza. 2013. *Gradient and categorical consonant cluster simplification in Persian: An ultrasound and acoustic study*. University of Ottawa. (tese de doutoramento).
- Fowler, Carol A. 1996. Listeners do hear sounds, not tongues. *The Journal of the Acoustical Society of America* 99(3). 1730–1741. DOI: 10.1121/1.415237.
- Fowler, Carol A. & Bruno Galantucci. 2005. The relation of speech perception and speech production. Em David B. Pisoni & Robert E. Remez (eds.), *The handbook of speech perception*, 633–652. Malden: Blackwell.
- Galantucci, Bruno, Carol A. Fowler & M. T. Turvey. 2006. The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review* 13(3). 361–377. DOI: 10.3758/BF03193857.
- Gentilucci, Maurizio & Michael C. Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews* 30(7). 949–960. DOI: 10.1016/j.neubiorev.2006.02.004.

- Gibbon, Fiona E. 1999. Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research* 42(2). 382–397. DOI: 10.1044/jslhr.4202.382.
- Gibbon, Fiona E. & Sara E. Wood. 2002. Articulatory drift in the speech of children with articulation and phonological disorders. *Perceptual and Motor Skills* 95(1). 295–307. DOI: 10.2466/pms.2002.95.1.295.
- Gick, Bryan, Penelope Bacsfalvi, Barbara M. Bernhardt, Sunyoung Oh, Slade Stolar & Ian Wilson. 2007. A motor differentiation model for liquid substitutions in children's speech. *Proceedings of meetings on acoustics* 1(1). 060003. DOI: 10.1121/1.2951481.
- Goldstein, Louis. 2003. Emergence of discrete gestures. Em Maria Josep Solé, Daniel Recasens & Joaquin Romero (eds.), *Proceedings of the 15th international congress of phonetic sciences*, 85–88. Barcelona: International Phonetic Association. [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15\\_0085.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_0085.pdf) (30 novembro, 2021).
- Goldstein, Louis, Dani Byrd & Elliot Saltzman. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. Em Michael A. Arbib (ed.), *Action to language via the mirror neuron system*, 215–249. Cambridge: Cambridge University Press.
- Goldstein, Louis & Carol A. Fowler. 2003. Articulatory phonology: A phonology for public language use. Em Niels O. Schiller & Antje S. Meyer (eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*, 159–207. New York: De Gruyter Mouton.
- Goldstein, Louis, Marianne Pouplier, Larissa Chen, Elliot Saltzman & Dani Byrd. 2007. Dynamic action units slip in speech production errors. *Cognition* 103(3). 386–412. DOI: 10.1016/j.cognition.2006.05.010.
- Goozée, Justine, Bruce Murdoch, Anne Ozanne, Yan Cheng, Anne Hill & Fiona Gibbon. 2007. Lingual kinematics and coordination in speech-disordered children exhibiting differentiated versus undifferentiated lingual gestures. *International Journal of Language & Communication Disorders* 42(6). 703–724. DOI: 10.1080/13682820601104960.
- Ingram, David. 1997. The categorization of phonological impairment. Em Barbara W. Hodson & Mary Louise Edwards (eds.), *Perspectives in applied phonology*, 19–41. Gaithersburg: Aspen.
- Lakatos, Imre. 1979. O falseamento e a metodologia dos programas de pesquisa científica. Em Imre Lakatos & Alan Musgrave (eds.), *A crítica e o desenvolvimento do conhecimento*, 109–243. São Paulo: Cultrix.

- Lieberman, Alvin M. & Doug H. Whalen. 2000. On the relation of speech to language. *Trends in Cognitive Sciences* 4(5). 187–196. DOI: 10.1016/S1364-6613(00)01471-6.
- McMillan, Corey T. & Martin Corley. 2010. Cascading influences on the production of speech: Evidence from articulation. *Cognition* 117(3). 243–260. DOI: 10.1016/j.cognition.2010.08.019.
- Munson, Benjamin, Jan Edwards, Sarah K. Schellinger, Mary E. Beckman & Marie K. Meyer. 2010. Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of vox humana. *Clinical Linguistics & Phonetics* 24(4-5). 245–260. DOI: 10.3109/02699200903532524.
- Namasivayam, Aravind K., Deirdre Coleman, Aisling O'Dwyer & Pascal van Lieshout. 2020. Speech sound disorders in children: An articulatory phonology perspective. *Frontiers in psychology* 10(1). 1–22. DOI: 10.3389/fpsyg.2019.02998.
- Poupier, Marianne. 2008. The role of a coda consonant as error trigger in repetition tasks. *Journal of Phonetics* 36(1). 114–140. DOI: 10.1016/j.wocn.2007.01.002.
- Poupier, Marianne & Louis Goldstein. 2005. Asymmetries in the perception of speech production errors. *Journal of Phonetics* 33(1). 47–75. DOI: 10.1016/j.wocn.2004.04.001.
- Pulvermüller, Friedemann, Martina Huss, Ferath Kherif, Fermin Moscoso del Prado Martin, Olaf Hauk & Yury Shtyrov. 2006. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences* 103(20). 7865–7870. DOI: 10.1073/pnas.0509989103.
- Rizzolatti, Giacomo & Michael A. Arbib. 1998. Language within our grasp. *Trends in neurosciences* 21(5). 188–194. DOI: 10.1016/S0166-2236(98)01260-0.
- Rizzolatti, Giacomo & Laila Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27(1). 169–192. DOI: 10.1146/annurev.neuro.27.070203.144230.
- Saltzman, Elliot & J. A. Scott Kelso. 1987. Skilled actions: A task-dynamic approach. *Psychological Review* 94(1). 84–106. DOI: 10.1037/0033-295X.94.1.84.
- Scobbie, James M., Fiona Gibbon, William J. Hardcastle & Paul Fletcher. 2000. Covert contrast as a stage in the acquisition of phonetics and phonology. Em Michael B. Broe & Janet B. Pierrehumbert (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon*, 194–207. Cambridge: Cambridge University Press.
- Shiller, Douglas M., Vincent L. Gracco & Susan Rvachew. 2010. Auditory-motor learning during speech production in 9-11-year-old children. *PloS One* 5(9). e12975. DOI: 10.1371/journal.pone.0012975.

- Studdert-Kennedy, Michael. 2002. Mirror neurons, vocal imitation, and the evolution of particulate speech. Em Maxim I. Stamenov & Vittorio Gallese (eds.), *Mirror neurons and the evolution of brain and language*, 207–227. Amsterdam, Philadelphia: John Benjamins.
- Studdert-Kennedy, Michael & Louis Goldstein. 2003. Launching language: The gestural origin of discrete infinity. Em Morten H. Christiansen & Simon Kirby (eds.), *Studies in the evolution of language*, 235–254. Oxford: Oxford University Press.
- van Lieshout, Pascal H. H. M. & Louis M. Goldstein. 2008. Articulatory phonology and speech impairment. Em Martin J. Ball, Michael R. Perkins, Nicole Müller & Sara Howard (eds.), *The handbook of clinical linguistics*, 467–479. Malden: Blackwell.