

# O papel das dicas articulatórias na percepção da fala. Uma concepção remodelada de uma antiga questão.

*The role of articulatory cues in perception of speech.  
A remodeled conception of an old issue.*

Wagner Ferreira Lima<sup>1</sup>  
Universidade Estadual de Londrina

♦ **RESUMO:** Sons da fala são realidades variáveis dependentes do contexto, uma das causas sendo a coarticulação vocal. Mas apesar disso a fala se mostra inteligível ao ouvinte, que supera as inconsistências acústicas por compensar a coarticulação. O objetivo desta exposição é fazer conhecer, tanto quanto possível, a clássica teoria motora da percepção da fala, para a qual a compensação é vista como dependendo do conhecimento vocal dos ouvintes. Uma vez que isso é essencialmente intermodal, essa teoria vem recebendo plausibilidade biológica das neurociências. Esse fato, ao lhe conferir generalidade, pode revigorar o papel dessa teoria na explicação da variação e percepção da fala.

♦ **PALAVRAS-CHAVE:** Coarticulação vocal; compensação de coarticulação; teoria motora da percepção da fala.

♦ **ABSTRACT:** Speech sounds are variable realities depending on the context, one of the causes being vocal coarticulation. But despite this, the speech proves to be intelligible to the listener, who overcomes the acoustic inconsistencies by compensating for coarticulation. The purpose of this exhibition is to make known, as much as possible, the classic motor theory of speech perception, for which compensation is seen as depending on the listeners' vocal knowledge. Since this is essentially intermodal, this theory has been receiving biological plausibility from the neurosciences. This fact, giving it generality, can invigorate the role of this theory in explaining the variation and perception of speech.

♦ **KEYWORDS:** Vocal coarticulation; coarticulation compensation; motor theory of speech perception.

## Introdução

A percepção dos sons da fala requer mais que exclusivamente estímulos acústicos para ser bem-sucedida. Uma pessoa ouvindo o áudio da pronúncia de [b<sup>h</sup>a] reiteradamente pode perceber igualmente /ba/ ([b<sup>h</sup>a]) ou /fa/ ([fa]), conforme seja a visão da articulação vocal que é pareada com o áudio. Se a visão da articulação for bilabial, a pessoa perceberá

---

<sup>1</sup> Doutor em Letras pela Universidade Estadual Paulista Júlio de Mesquita Filho, UNESP, Brasil. Professor Associado do Departamento de Letras Vernáculas e Clássicas, da Universidade Estadual de Londrina, UEL, Brasil.

a pronúncia de /ba/ ([b<sup>h</sup>a]); se for labiodental, ela ouvirá a de /fa/ ([fa]) (mesmo seu ouvindo captando o espectro acústico da pronúncia real)<sup>2</sup>. O que aconteceu aqui? O cérebro “notou” um conflito entre as modalidades auditiva e visual na condição de [b<sup>h</sup>a]. E, para ajustar a percepção da pronúncia, ele dissolveu esse conflito apelando às dicas visuais disponíveis.

Esse é um tipo de ilusão conhecido por efeito McGurk: Uma ilusão perceptual multimodal, gerada pela interferência da visão de gestos articulatórios sobre a audição de um padrão sonoro da fala (TIIPPANA, 2014; ROSENBLUM, 2008). Conquanto limitado em seu escopo, esse efeito nos revela que a percepção (sonora) é regida pelo princípio de integração intermodal. Logo, se podemos decidir entre um ou outro fonema em meio a inconsistências de estímulos, como no exemplo inicial; então fazemos em alguma medida uso desse princípio durante a percepção. Assumimos que isso é o caso quando lidamos com o fenômeno de “compensação de coarticulação” (para outros processos perceptuais, veja DAVIS; JOHNSRUDE, 2007). Como veremos abaixo, esse conceito, tal como empregado por Liberman (1985; 1989) e colaboradores, reforça a importância do papel da intermodalidade na percepção.

Assim sendo, o objetivo desta exposição é esclarecer como, durante a percepção da fala, nossa mente recorre a dicas outras que não só as acústicas para estabelecer perceptos sonoros estáveis (fonemas). Faremos isso, em primeiro lugar, revelando como a teoria motora da percepção da fala lida com o fenômeno da compensação de coarticulação. E, em segundo, conferindo plausibilidade biológica a essa teoria.

Assim, na seção 1, apresentamos o tema da coarticulação dos sons vocais focalizando seus efeitos sobre os estímulos sonoros da fala. Como pretendemos mostrar em 2, esses efeitos de coarticulação são compensados durante a percepção para se manter a consistência perceptual. Entre as teorias clássicas que estudam esse mecanismo está, como veremos em 3, a teoria de Liberman e colaboradores, que enfatizam o papel da intermodalidade na compensação de coarticulação. Os achados neurocientíficos descritos na seção 4 são trazidos à cena para conferir plausibilidade biológica ao pensamento de Liberman, e para justificar a alegação acima de que a percepção (linguística) é intermodal.

## 1. Coarticulação dos sons da fala

Coarticulação é *grosso modo* o processo pelo qual movimentos motores, ou dicas cinemáticas, que caracterizam uma ação se sobrepõem aos gestos da ação subsequente (FOWLER, 2006). Coarticulação ocorre comumente na produção da linguagem oral. Por exemplo, no português falado no Brasil, uma fricativa alveolar como /s/, em posição final de sílaba, seguida de africada alveopalatal [tʃ], como “lástima”, pode se realizar como fricativa alveopalatal [ʃ]; em virtude da influência do traço alveopalatal da consoante seguinte. Em outros ambientes, como em “casta”, essa influência não ocorre porquanto seguida de segmento oclusivo alveolar (/t/). Aconteceu aqui que a articulação de /s/ antecipou a articulação alveopalatal do fone seguinte; coarticulando-se com ele e gerando uma mistura de traços.

Outro exemplo, também do português brasileiro, é a articulação da palavra “cama”. O fonema /a/ na primeira sílaba pode realizar-se como uma nasal [ã] em [ˈkã.mə]. Aí a úvula antecipa a articulação de /m/ (ressoando as cavidades nasais), antes mesmo que a produção de /a/ seja concluída.

Por que a sobreposição de gestos articulatórios de fones adjacentes (vogais e consoantes) frequentemente mistura as fronteiras das categorias fonéticas, o ouvinte está

---

<sup>2</sup> Exemplo adaptado a partir da BBC Horizon Clip (2010).

sempre diante do problema de identificar fonemas em meio a contínuos acústicos (TOSCANO, 2011; KUHL, 2010). O fenômeno curioso quanto a isso é o fato de a mente humana ser capaz de realizar essa tarefa naturalmente. Em parte, essa capacidade se explica em termos de uma operação intuitiva básica denominada compensação de coarticulação.

## 2. Compensação de coarticulação

A busca pelo entendimento de como os ouvintes percebem sinais acústicos sensíveis ao contexto em termos de categorias levou à suposição de um mecanismo de compensação subjacente aos efeitos de categoria (DIEHL; LOTTO; HOLT, 2004). Em um experimento, Mann e Reep (1981) geraram efeitos de compensação. Eles produziram um estímulo ambíguo ao longo de um contínuo /ta-/ka/, com um estímulo precursor representado por /f/ ou /s/. Devido à coarticulação, /f/ movimenta para trás o lugar de articulação das oclusivas, fazendo o estímulo ambíguo soar mais como /ka/; enquanto /s/ movimenta para frente os articuladores, fazendo-o soar mais como /ta/. Apesar disso, eles notaram que ouvintes tendiam a perceber mais /ta/ quando o estímulo seguia /f/ do que quando seguia /s/. Os participantes estavam compensando coarticulação durante percepção.

Trata-se de uma tendência dos ouvintes a atribuir pontos ambíguos de constrição (aqueles nas zonas intermediárias do contínuo) a efeitos de coarticulação (VISWANATHAN; MAGNUSON; FOWLER, 2010). Se, de um lado, efeitos de contextos misturam os segmentos vocais; de outro, eles contribuem para a percepção dos fonemas. Conforme Nguyen (2005),

[...] segundo numerosos autores, esses fenômenos de coarticulação não são obstáculos para a identificação dos fonemas pelo ouvinte. Pelo contrário, supõe-se que o ouvinte explora esses efeitos de contexto para identificar os fonemas mais facilmente. Isso lhe é possível na medida em que os efeitos de contexto assumem um caráter sistemático e regular. Considera-se que o ouvinte utiliza essas regularidades a seu favor, reportando-se, de uma forma ou de outra, ao contexto para identificar cada fonema<sup>3</sup>. (p. 8-9)

Isso significa, por outras palavras, que o aparato cognitivo está preparado para perceber sons em contextos dinâmicos de comunicação; como os resultantes de coarticulação. Nesse sentido, ele é capaz de “normalizar” as flutuações acústicas; contando, para tanto, com o conhecimento intuitivo dos processos intrínsecos de variação. As abordagens sobre percepção da fala são unânimes em admitir que a compensação de coarticulação é inerente à identificação e discriminação dos sons. Entretanto, elas divergem quanto às dependências necessárias para o desempenho dessas tarefas. Vejamos então como as teorias motoras da percepção da fala lidam com a compensação.

## 3. Teorias motoras da percepção da fala

Teorias motoras da percepção da fala são aquelas que consideram a modalidade motora como um mediador da percepção dos sons vocais. Elas se dividem em duas: a “teoria do realismo direto da percepção da fala” e a “teoria motora de percepção da fala”

---

<sup>3</sup> Selon de nombreux auteurs, ces phénomènes de coarticulation ne font pas obstacle à l'identification des phonèmes par l'auditeur. Il est au contraire supposé que l'auditeur exploite ces effets de contexte pour identifier les phonèmes plus facilement [PER 86]. Cela lui est possible dans la mesure où les effets de contexte revêtent un caractère systématique et régulier. On considère que l'auditeur utilise ces régularités à son profit en se rapportant d'une manière ou d'une autre au contexte pour identifier chaque phonème.

(DIEHL; LOTTO; HOLT, 2004). Em geral, ambas partilham a suposição de que o sistema auditivo, sozinho, não é capaz de reconhecer segmentos sonoros da fala. Assim sendo, eles alegam que para isso os ouvintes contam com intuições dos gestos articulatórios da fala (movimento da língua, da mandíbula, arredondamento dos lábios etc.). O corolário disso é tomar a percepção da fala como uma capacidade exclusiva do homem.

Mas as semelhanças param por aí. Com efeito, elas diferem quanto a como os ouvintes se valem do conhecimento motor na audição. A primeira, a teoria do realismo, considera que os ouvintes usam o conhecimento imediato e direto dos articuladores vocais para compensar coarticulação e, então, reconhecer as unidades discretas da fala. Já a segunda, a propriamente motora, defende que a categorização dos espectros sonoros é determinada, não pela observação direta dos articuladores, mas sim pelo conhecimento fonológico – mais especificamente, pelo conhecimento de gestos vocais intentados.

A teoria do realismo direto sustenta que os ouvintes fazem uso da estrutura acústica nas emissões vocais como informação para a fonte causal dessa estrutura (o efeito McGurk é um exemplo disso). Eles buscam os gestos articulatórios que produzem o sinal acústico da fala e, assim, reconhecem a categoria fonética em causa. Por exemplo, qualquer consequência acústica de efeito de “posteriorização” de /ʃ/ no domínio do lugar de articulação de /t/ é atribuído a /ʃ/, e não a /t/ (MANN; REEP, 1981).

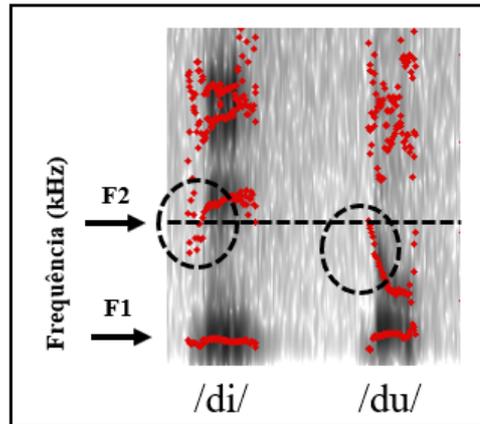
Já a teoria motora da percepção da fala enfatiza o papel *top-down* da mediação fonológica durante o reconhecimento das categorias funcionais da fala (LIBERMAN; MATTINGLY, 1985). Proposta por Liberman e colaboradores, essa teoria sofreu alguns incrementos ao longo de sua existência, desde seu aparecimento no final da década de 50 (GALANTUCCI; FOWLER; TURVEY, 2006). Contudo, o objetivo central se manteve relativamente o mesmo desde sua origem: fundamentar a percepção categorial na base dos aspectos invariantes da fala, vale dizer, dos aspectos subfonêmicos-articulatórios (traços distintivos) dos sons vocais.

A seguir vamos enfatizar esta última, entre outras coisas, por oferecer apoio empírico à alegação inicial de que a percepção da fala se assenta em um princípio de integração intermodal.

### 3.1 *A teoria motora da percepção da fala (TMPF)*

Essa teoria está intimamente associada ao nome de A. Liberman, um dos pioneiros do estudo experimental da percepção dos sons da fala. O fenômeno de coarticulação se evidenciou a Liberman et al. (1957a; 1967) em um estudo sobre a percepção de /d/ em contexto acústico variável. Os autores exibiram espectrogramas de duas sílabas sintéticas, /di/ (“dee”) e /du/ (“doo”). Cada espectrograma apresentava dois formantes (F1 e F2). Nesses padrões, os formantes de estado-estacionário correspondiam aos valores das vogais meta /i/ e /u/ (Figura 1).

**Figura 1 – Espectrogramas das sílabas.**



Fonte: Adaptado de Liberman *et al.* (1957a; 1967)

Diferenças na transição de formantes (traços fortes, indicados entre círculos) da consoante /d/ em combinação com vogais distintas. /i/ apresenta um segundo formante (F2) mais alto, de modo a elevar o F2 da consoante (1814Hz para 2217Hz). /u/ contém um F2 mais baixo, abaixando o F2 da consoante (1814Hz para 1025Hz). Conforme notou Liberman, apesar dessas diferenças acústicas, ouvintes notam as consoantes como instâncias do fonema /d/.

Já a transição de formantes no início de cada sílaba transmitia importante informação acerca da consoante inicial. De um lado, a transição ascendente de F1 de ambas as sílabas indicava que se tratava de uma oclusiva vozeada, como /b/, /d/ ou /g/. De outro, a transição ascendente de F2 de /di/ e a descendente de F2 de /du/ passavam informação crítica sobre o lugar de articulação (tratava-se de /d/ mais do que de /b/ ou /g/). Apesar dessas diferenças espectrais na transição de F2 em cada sílaba, os ouvintes reconheceram as consoantes como instâncias de /d/ (LIBERMAN *et al.*, 1957a; LIBERMAN, 1957b; LIBERMAN *et al.*, 1967). Para Liberman e colaboradores, esse fato se tornou uma prova praticamente irrefutável de que os ouvintes contam com o conhecimento fonológico para perceber os sinais acústicos como categorias invariáveis (DIEHL; LOTTO; HOLT, 2004).

Além disso, os autores postularam a existência de um módulo fonético no cérebro que incorpora competência motora a respeito da fala. Tal módulo inclui, basicamente, conhecimento de coarticulação e suas consequências acústicas (FOWLER, 2006), ou seja, o conhecimento de gestos articulatorios intentados (SITEK, JOHNSON, 2012). Segundo Liberman e Mattingly (1985),

os objetos da percepção da fala são os gestos fonéticos intentados do falante, representados no cérebro como comandos motores invariantes que exigem movimentos dos articuladores através de certas configurações linguisticamente significativas. Esses comandos gestuais são a realidade física subjacente às noções fonéticas tradicionais – por exemplo, “posterioridade da língua”, “arredondamento dos lábios” e “elevação da mandíbula” – que fornecem as bases para categorias fonéticas [os fonemas]. Eles são os eventos elementares da produção e percepção da fala<sup>4</sup>. (p. 2)

<sup>4</sup> [...] the objects of speech perception are the intended phonetic gestures of the speaker, represented in the brain as invariant motor commands that call for movements of the articulators through certain linguistically significant configurations. These gestural commands are the physical reality underlying the traditional phonetic notions-for example, ‘tongue backing,’ ‘lip rounding,’ and ‘jaw raising’- that provide the basis for phonetic categories. They are the elementary events of speech production and perception.

De acordo com isso, compensação ocorre à medida que os ouvintes fazem uso dessa competência para perceber os sons da fala. Eles “precisam rastrear gestos separadamente durante intervalos nos quais coarticulação produz consequências acústicas que misturam informação para gestos sobrepostos”<sup>5</sup> (FOWLER, 2006, p. 162).

Consideremos, por exemplo, o estudo clássico de Mann (1980) sobre os dissílabos /arda/ e /arga/. Ele notou a compensação (e efeitos categoriais) durante a percepção do contínuo sonoro entre /da/ e /ga/, usando como estímulos precursores os segmentos /ar/ e /al/. A coarticulação produz os segmentos homorgânicos /ld/ e /rg/. Em termos acústicos, o espectro de /r/ abaixa a frequência de F3 da consoante seguinte; enquanto /l/ eleva a frequência de F3 desse segmento. Apesar disso, o ouvinte tende a notar mais /da/ seguindo /ar/ (/arda/) do que seguindo /al/, que favorece a percepção de /ga/ (/alga/). Do ponto de vista da teoria motora de Liberman, o ouvinte atribui as consequências do abaixamento de F3 dos gestos de /r/, durante a articulação de /d/ e /g/, aos próprios gestos tentados de produção de /r/.

O efeito McGurk também serve *mutatis mutandis* para apoiar empiricamente a alegação da teoria motora. Como já vimos no início, a produção desse efeito requer que a percepção da fala seja mediada por integração intermodal. (Se o ouvinte é “enganado” pelas dicas articulatórias, ouvindo /ada/ no lugar de /aba/, é porque o conhecimento de gestos tentados para a produção do percepto /ada/ foi eliciado pela imagem visual da vocalização desse segmento.) Para a teoria de Liberman, as dicas motoras dos gestos vocais não são externas, como no efeito McGurk, mas sim internas; fazendo parte de uma competência fonética do ouvinte.

As assunções das teorias motoras da percepção da fala são, em geral, questionadas no âmbito da abordagem auditiva da percepção (KUHL, 2000). Especialmente pesquisas sobre aquisição da linguagem não consideram o sistema motor (seja como conhecimento ou como movimentos efetivos) necessário para a percepção dos sinais acústicos (KUHL, 2007; MAYE, WERKER; GERKEN, 2002). De modo geral, defensores da abordagem auditiva sustentam que percepção auditiva é uma capacidade cognitiva autossuficiente, e está sujeita aos mesmos processos envolvidos na percepção em geral (para uma revisão, veja GOLDSTONE; HENDRICKSON, 2009; HOLT; KLUNDER, 2000).

Outra crítica forte proveniente da abordagem auditiva é quanto ao papel efetivo das representações motoras para o reconhecimento da fala. Para alguns estudiosos tais representações, sejam imagens diretas ou tentadas, funcionam mais como uma variável moderadora do que mediadora (HICKOK; HOUDE; ROMG, 2011). Assim, elas são requisitadas apenas quando a percepção auditiva vai mal e demanda o auxílio de dicas do sistema motor.

Outra questão de desavença é quanto à natureza do conhecimento motor recrutado durante percepção. Para Liberman e Mattingly (1989), percepção da fala supõe um módulo fonético especializado para mapear sinais acústicos em gestos articulatórios intencionados. Esse módulo difere, assim, de outros módulos dedicados ao processamento acústico, como a identificação da fonte sonora (LIBERMAN; MATTINGLY, 1989). A teoria do realismo direto não assume esse postulado, pois, para ela, a observação dos gestos vocais é o conhecimento que os ouvintes empregam durante identificação e discriminação da fala (FOWLER, 2006; DIEHL; LOTTO; HOLT, 2004.)

Seja lá como for, atualmente alguns achados neurocientíficos podem ser empregados para justificar a TMPF em termos biológicos; o que pode conferir mais

---

<sup>5</sup> [...] must track gestures separately during intervals in which coarticulation causes acoustic consequences that blend information for the overlapping gestures.

generalidade a ela. Como veremos abaixo, um desses achados são os chamados “neurônios-espelho”. Essa descoberta reforça a tese de que a percepção em geral, e não apenas a da fala, é intermodal por natureza.

#### 4. Bases neuronais da teoria motora

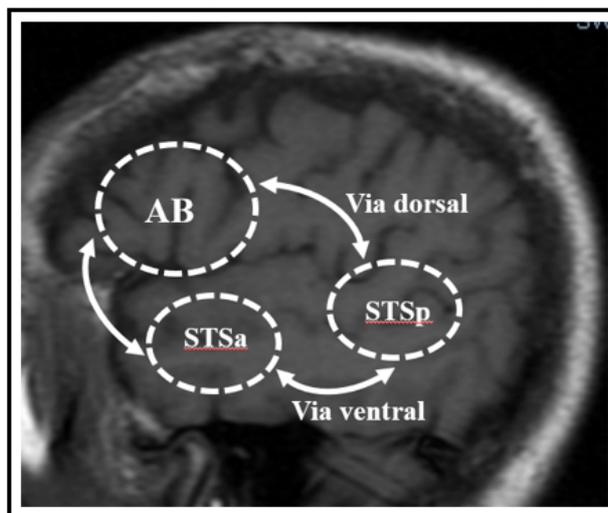
O que fizemos até agora foi descrever como a TMPF explica a percepção da fala em termos de compensação de coarticulação. Mas, diferente de suas rivais, essa teoria alega que a compensação ocorre com base no conhecimento da fala motora. Contudo, os dados apresentados por ela, apesar de robustos, não são suficientes para tornar essa alegação livre de críticas. E é improvável que isso venha a acontecer. Nesse sentido, acreditamos que a TMPF pode reduzir seu grau de falibilidade recorrendo a outros argumentos, além dos dados experimentais. Isso pode ser o caso ao se atribuir plausibilidade biológica a sua alegação. As neurociências oferecem uma avenida de possibilidades. Examinemos então duas dessas, as interconexões neuronais e o comportamento dos neurônios-espelho.

Começemos pelas interconexões. O sistema auditivo integra uma vasta rede de conexões mútuas, córtico-corticais, córtico-talâmicas, córtico-estriatais etc. (“córtex auditivo distribuído” (VINNER; LEE, 2007; PATTERSON *et al.*, 2002)). O centro desse sistema é o córtex auditivo primário (A1), no giro transversal de Heschl, ponto de chegada das aferências oriundas da cóclea. A1 apresenta uma topografia que reflete o mapa tonotópico da membrana basilar no Órgão de Corti (HUMPHRIES; LIEBENTHAL; BINDER, 2010). O auditivo primário é o responsável pelo processamento dos traços do estímulo acústico.

O córtex auditivo primário faz conexões com as áreas auditivas secundárias, “*belt*” e “*parabelt*” (PICKLES, 2015; HACKETT, 2008). Estas estabelecem conexões mútuas com regiões anteriores e posteriores do temporal superior; responsáveis pela geração de “objetos sonoros”, como os fonemas. Na direção rostral, tais áreas estabelecem ligações com sulco temporal superior anterior (STSa) e giro temporal superior (GTS); e, na direção caudal, com o plano temporal (PT), sulco temporal superior posterior (STSp) e a área de Wernicke (MESULAM *et al.*, 2015). Portanto, como as informações visuais, distribuindo-se ventral e dorsalmente pelo córtex (UNGERLIEDER; HAXBY, 1994); as informações auditivas formam, anteriormente, a via auditiva ventral; e, posteriormente, a via auditiva dorsal (Figura 2).

Esses caminhos, ventral e dorsal, constituem o que hoje é conhecido como “o quê” e o “onde” dos sons, respectivamente (AHVENINEN *et al.*, 2006; BARRETT; HALL, 2006; ALAIN *et al.*, 2001). A **via ventral** representa o substrato neuronal dos processos de identificação das categorias fonéticas (SCOTT; JOHNSRUDE, 2003). E, porque ela está associada com os significados das palavras, é possível que sofra a modulação de processos lexicais. Já a **via dorsal**, o caminho do “onde”, implementa a relação intrínseca entre percepção e articulação dos sons vocais (DAVIS; JOHNSRUDE, 2007).

**Figura 2 – Integração sensório-motora para a audição.**



Fonte: Autor

Vias ventral e dorsal do sistema auditivo. As regiões estão representadas entre círculos e as conexões entre setas bilaterais. Área de Broca (AB); Sulco temporal superior anterior (STSa); Sulco temporal superior posterior (STSp).

O sistema auditivo se conecta à Área de Broca igualmente pela via ventral e pela via dorsal. As conexões entre eles são mútuas e formam um sistema multimodal integrado (HICKOK; HOUDE; RONG, 2012). Esse fato relativiza a suposta tese de que efeitos de percepção categorial são decorrentes exclusivamente da via do “o quê”. Fatores articulatórios, ou fonéticos, podem estar envolvidos também na percepção, e ser mesmo necessários para ela. Isso está em linha com a tese de Liberman.

A via dorsal do sistema auditivo parece ser primordial ao cumprimento dessa função. Com efeito, ela estabelece conexões também como o parietal inferior (giro supramarginal e opérculo parietal), regiões cruciais para o reconhecimento da fonte sonora – a localização do som no espaço –; tendo o *self* corporal como ponto de referência (OVERWALLE; BAETENS, 2009). Uma vez que os articuladores constituem a fonte dos sons vocais, é de se esperar que o sistema auditivo recrute a via dorsal mais fortemente durante a percepção.

Além disso, pela via dorsal corre o fascículo arqueado, um feixe de fibras ligando o temporal superior posterior e o premotor ventral. Avaria do fascículo arqueado dá lugar à síndrome conhecida como afasia de condução (ARDILA, 2010): um mal que leva os pacientes a cometer erros sistemáticos de articulação dos sons ouvidos; mesmo estando seus sistemas auditivo e motor preservados (BUCHSBAUM *et al.*, 2011). As conexões mútuas aí existentes criam uma via de interação entre os sistemas auditivo e motor, tal que seu comprometimento afeta a memória fonológica de curto prazo e, por extensão, a capacidade de produção correta dos sons linguísticos.

Essa ideia da percepção sensório-motora ganha ainda mais plausibilidade quando consideramos o funcionamento dos neurônios-espelho. Descobertos na década de 90 (DEVLIN; AYDELOT, 2009), neurônios-espelho são uma classe específica de neurônios bimodais. Eles estão envolvidos na execução de uma ação autogerada, mas também na observação dessa ação realizada externamente (RIZZOLATTI; CRAIGHERO, 2004). Originalmente, eles foram encontrados no premotor ventral de macacos (área F5) e designados neurônios visuomotores. Em humanos, o correspondente anátomo-funcional

é o centro da linguagem falada, conhecido por Área de Broca (**AB**) (KILNER; LEMON, 2013). Anatomicamente, a Área de Broca está situada no premotor ventral esquerdo, e fisiologicamente está ligada ao sistema auditivo por conexões mútuas.

Atualmente, sabemos que esses neurônios se distribuem para além das áreas motoras, e estão envolvidos no reconhecimento do objetivo das ações a partir de diferentes meios (KILNER; LEMON, 2013), como os “neurônios audiovisuais”, ativados pela audição de sons de ações (p.e., “som de amendoim quebrando-se”) (GAZZOLA; AZIZ-ZADEH; KEYSERS, 2006; KEYSERS *et al.*, 2003).

Em paralelo com os córtices auditivos, neurônios-espelho do premotor podem responder aos sons da fala ao gerar informações sobre as articulações desses sons. Tratar-se-ia de um processo de compreensão multimodal. Como no caso dos neurônios audiovisuais, em que um som informa a natureza da sua fonte (o “quebrar do amendoim”), o som da fala poderia informar a natureza dos articuladores vocais que lhes dão origem. Assim, a afirmação de Liberman *et al.* (1985) segundo a qual perceber a fala é mapear dicas acústicas em gestos articulatorios pode ser interpretada à luz dessa constatação neurocientífica.

Alguns achados fornecem apoio empírico para essa suposição. Watkins, Strafella e Paus (2003) encontraram integração sensoriomotora durante percepção. Seus testes evidenciaram que atividade da AB, no córtex motor primário, aumentava durante a audição de sons da fala; comparada com sons significativos de não-fala (p.e., “um copo quebrando-se”). Outro trabalho descobriu atividade aumentada no córtex motor da língua, quando falantes italianos ouviam a palavra “terra”, que envolve movimento da língua; mas não quando eles ouviam a palavra “zaffo”, que não engaja esse articulador (FADIGA *et al.*, 2002).

Em síntese, porque distintos circuitos motores associados a lábios e língua são ativados por fonemas específicos que engajam esses articuladores (DEVLIN; AYDELOTT, 2009), regiões do córtex motor subjacentes à produção podem ser recrutadas também durante percepção. Esse mecanismo é ilustrativo do comportamento bimodal dos neurônios-espelho. Nesse caso, suas respostas são do tipo áudio-motores.

Em suma, a descrição ora feita é um indício a mais de que a percepção da fala engaja conjuntamente sistemas a priori considerados como operando isoladamente. Os achados neurocientíficos reforçam, assim, a base de sustentação de um dos principais pilares epistemológicos dessa teoria: a hipótese de que uma das estratégias usadas pela mente humana para compensar coarticulação é integrando o sistema sensorial e o sistema motor.

### **Considerações finais**

Nessa exposição fizemos uma breve descrição do fenômeno da compensação de coarticulação. Enfatizamos particularmente a explicação oferecida a ele pela TMPF e buscamos plausibilidade biológica para tal explicação nas neurociências cognitivas. Depois de tudo, o que podemos extrair dessa descrição é que, como o efeito McGurk, compensação de coarticulação funciona para dissolver inconsistências acústicas nos estímulos da fala. Todavia, diferente desse efeito e de sua rival mais próxima (a teoria realista direto da percepção da fala), ela conta com processos intermodais (sensorio-motores) internos.

A elucidação desse fenômeno, conjugada com dados neurocientíficos, deixa a importante conclusão: Os processos neuropsicológicos sendo intermodais e integrados, a questão da percepção da fala não pode ser indiferente a eles; e desse modo precisa ser esclarecida à luz desses processos.

Atualmente essas considerações são de suma importância especialmente para as pesquisas em aquisição da linguagem verbal. Desenvolvimento da fala e da escuta não se faz senão pela combinação de modalidades sensoriais e motoras diversas. Por exemplo, particularmente no caso da audição, sabemos que a percepção dos sons da linguagem, como consoantes e vogais, depende de ritmo, acento, taxa de fala, entonação etc. (KUHL, 2007); que por sua vez demanda movimentos corporais e manuais correspondentes (MUNHALL et al., 2004). Até mesmo as dificuldades com leitura podem ser superadas apelando para terapias multimodais; para se mencionarem alguns poucos exemplos.

## REFERÊNCIAS

AHVENINEN, J. *et al.* Task-modulated “what” and “where” pathways in human auditory cortex. **PNAS**, v. 103, n. 39, p. 14608-14613, 2006. Disponível em: <http://www.pnas.org/content/103/39/14608> Acesso em: 16 set. 2018.

ALAIN, C. *et al.* “What” and “Where” in the human auditory system. **PNAS**, v. 98, n. 21, p. 12301-12306, 2001. Disponível em: <http://www.pnas.org/content/98/21/12301> Acesso em: 16 set. 2018.

ARDILA, A. A review of conduction aphasia. **Current Neurology and Neuroscience Reports**, v. 10, n. 6, p. 499-503, 2010. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/20711691> Acesso em: 10 out. 2018.

BARRETT, D. J. K.; HALL, D. A. Response preferences for “what” and “where” in human non-primary auditory cortex. **Neuroimage**, v. 32, n.2, p. 968-977, 2006. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1053811906002199> Acesso em: 16 set. 2018.

BBC Horizon Clip. **McGurk Effect**. Auditory Ilusion. 2010. (2m11s). Disponível em: <https://www.youtube.com/watch?v=2k8fHR9jKVM> Acesso em: 23 jun. 2020.

BUCHSBAUM, B. *et al.* Conduction aphasia, sensory-motor integration, and phonological short-term memory – an aggregate analysis of lesion and fMRI data. **Brain and Language**, v. 119, n.3, p. 119-128, 2011. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3090694/> Acesso em: 10 out. 2018.

DAVIS, M. H.; JOHNSRUDE, I. S. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. **Hearing Research**, v. 299, n. 1-2, p. 132-147, 2007. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378595507000263> Acesso em: 16 set. 2018.

DEVLIN, Joseph T.; AYDELOTT, Jennifer. Speech perception: Motoric contributions versus the motor theory. **Current Biology**, v. 19, n. 5, p. R198-R200, 2009. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0960982209005351> Acesso em: 22 ago. 2020.

DIEHL, R. L.; LOTTO, A. J.; HOLT, L. L. Speech perception. **Annual Reviews Psychology**, v. 55, p. 149-179, 2004. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/14744213> Acesso em: 15 set. 2018.

FADIGA, L. *et al.* Speech listening specifically modulates the excitability of tongue muscles: a TMS study. **The European Journal of Neuroscience**, v. 15, n. 2, p. 399-402, 2002. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/11849307> Acesso em: 16 set. 2018.

- FOWLER, C. A. Compensation for coarticulation reflects gesture perception, not spectral contrast. **Perception & Psychophysics**, v. 68, n. 2, p. 161-177, 2006. Disponível em: <https://blogs.umass.edu/jkingstn/files/2013/12/fowler-2006-p-and-p.pdf> Acesso em: 15 set. 2018.
- FRANCIS, A. L.; BALDWIN, K.; NUSBAUM, H. C. Effects of training on attention to acoustic cues. **Perception and Psychophysics**, v. 62, n. 8, p. 1668-1680, 2000. Disponível em: <https://link.springer.com/article/10.3758/BF03212164> Acesso em: 21 set. 2018.
- GALANTUCCI, B.; FOWLER, C. A.; TURVEY, M. T. The motor theory of speech perception reviewed. **Psychonomic Bulletin & Review**, v. 13, n. 3, p. 361-377, 2006. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2746041/> Acesso em: 12 out. 2018.
- GAZZOLA, V.; AZIZ-ZADEH, L.; KEYSERS, C. Empathy and somatotopic auditory mirror system in humans. **Current Biology**, v. 16, n. 18, p. 1824-1829, 2006. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/16979560> Acesso em: 10 out. 2018.
- GOLDSTONE, R. L.; HENDRICKSON, A. Categorical perception. **WIREs Cognitive Sciences**, p. 1-10, 2009. Disponível em: <https://pdfs.semanticscholar.org/a703/1ba062221cb706d49f0b100a1b1be5773a64.pdf> Acesso em: 24 set. 2018.
- HACKETT, T. A. Anatomical organization of the auditory cortex. **Journal of the American Academy of Audiology**, v. 19, n. 10, p. 774-779, 2008. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/19358457> Acesso em: 24 set. 2018.
- HICKOK, G.; HOUDE, J.; RONG, F. Sensorimotor integration in speech processing: computational basis and neural organization. **Neuron**, v. 69, n. 3, p. 407-422, 2011. Disponível em: [https://www.cell.com/neuron/abstract/S0896-6273\(11\)00067-5](https://www.cell.com/neuron/abstract/S0896-6273(11)00067-5) Acesso em: 16 set. 2018.
- HOLT, L. L.; KLUENDER, K. R. General auditory processes contribute to perceptual accommodation of coarticulation. **Phonetica**, v. 57, n. 2-4, p. 170-180, 2000. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/10992137> Acesso em: 06 out. 2018.
- HUMPHRIES, C.; LIEBENTHAL, E.; BINDER, J. R. Tonotopic organization of human auditory cortex. **Neuroimage**, v. 50, n. 3, p. 1202-1211, 2010. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2830355/> Acesso em: 24 set. 2018.
- KEYSERS, C. *et al.* Audiovisual mirror neurons and action recognition. **Experimental Brain Research**, v. 153, p. 628-636, 2003. Disponível em: [https://www.researchgate.net/profile/Maria\\_Umlita/publication/10598462\\_Audiovisual\\_motor\\_neurons\\_and\\_action\\_recognition/links/547c491c0cf205d16881f7dd/Audiovisual\\_l-motor-neurons-and-action-recognition.pdf](https://www.researchgate.net/profile/Maria_Umlita/publication/10598462_Audiovisual_motor_neurons_and_action_recognition/links/547c491c0cf205d16881f7dd/Audiovisual_l-motor-neurons-and-action-recognition.pdf) Acesso em: 10 out. 2018.
- KILNER, J. M.; LEMON, R. N. What we know currently about mirror neurons. **Current Biology**, v. 23, n. 23, p. 1057-1062, 2013. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3898692/> Acesso em: 24 set. 2018.
- KUHL, P. K. A new view on language acquisition. **PNAS**, v. 97, n. 22, p. 11850-11857, 2000. Disponível em: <http://www.pnas.org/content/97/22/11850> Acesso em: 12 out. 2018.

KUHL, P. K. Is speech learning “gated” by the social brain? **Developmental Science**, v. 10, n. 1, p. 110-120, 2007. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/17181708> Acesso em: 12 out. 2018.

LIBERMAN, A. M. *et al.* The discrimination of speech sounds within and across phoneme boundaries. **Journal of Experimental Psychology**, v. 54, n. 5, p. 358-368, 1957a. Disponível em: <https://pdfs.semanticscholar.org/101c/7d3c3444f048ba7e6f794cd914099cad74be.pdf> Acesso em: 04 out. 2018.

LIBERMAN, A. M. Some results of research on speech perception. **The Journal of the Acoustical Society of America**, v. 29, n.1, p. 117-123, 1957b. Disponível em: [https://www.researchgate.net/publication/232552593\\_Some\\_Results\\_of\\_Research\\_on\\_Speech\\_Perception](https://www.researchgate.net/publication/232552593_Some_Results_of_Research_on_Speech_Perception) Acesso em: 12 out. 2018.

LIBERMAN, A. M. *et al.* Perception of the speech code. **Psychological Review**, v. 74, n. 6, p. 431-461, 1967. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/4170865> Acesso em: 12 out. 2018.

LIBERMAN, A. M.; MATTINGLY, I. G. The motor theory of speech perception revised. **Cognition**, v. 21, p. 1-36, 1985. Disponível em: [http://psych.colorado.edu/~kimlab/Liberman\\_Mattingly.Cognition1985.pdf](http://psych.colorado.edu/~kimlab/Liberman_Mattingly.Cognition1985.pdf) Acesso em: 15 set. 2018.

LIBERMAN, A. M.; MATTINGLY, I. G. A specialization for speech perception. **Science**, v. 243, n. 4890, p. 489-494, 1989. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/2643163> Acesso em: 12 out. 2018.

LISKER, L. “Voicing” in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. **Language and Speech**, v. 29, n. 1, p. 3-11, 1986. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/3657346> Acesso em: 21 set. 2018.

MAYE, J.; WERKER, J. F.; GERKEN, L. Infant sensitivity to distributional information can affect phonetic discrimination. **Cognition**, v. 82, n. 3, p. 101-111, 2002. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/11747867> Acesso em: 12 out. 2018.

MANN, V. A. Influence of preceding liquid on stop-consonant perception. **Perception & Psychophysics**, v. 28, n. 5, p. 407-412, 1980. Disponível em: <https://link.springer.com/content/pdf/10.3758/BF03204884.pdf> Acesso em: 15 set. 2018.

MANN, V. A.; REPP, B. H. Influence of preceding fricative on stop-consonant perception. **The Journal Acoustical Society of America**, v. 69, n. 2, p. 548-558, 1981. Disponível em: <https://asa.scitation.org/doi/abs/10.1121/1.385483?journalCode=jas> Acesso em: 15 set. 2018.

MESULAM, M-M. *et al.* The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. **Brain: an Journal of Neurology**, v. 138, n. 8, p. 2423-2437, 2015. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/26112340> Acesso em: 24 set. 2018.

MUNHALL, K. G. *et al.* Visual prosody and speech intelligibility: head movement improves auditory speech perception. **Psychological Science**, v. 15, n. 2, p. 133-137, 2004. Disponível em: <https://www.jstor.org/stable/40063940?seq=1> Acesso em: 17 jun. 2020.

NGUYEN, N. Perception de la parole. In: NGUYEN, N.; VAUQUIER, S.; DURANT, J. **Phonologie et Phonétique**, Hermes, p. 1-22, 2005. Disponível em: <https://hal.archives-ouvertes.fr/hal-00142953/document> Acesso em: 10 out. 2018.

OVERWALLE, F. Van; BAETENS, K. Understanding other's actions and goals by mirror and mentalizing system: a meta-analysis. **NeuroImage**, v. 48, n. 3, p. 564-584, 2009. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/19524046> Acesso em: 24 set. 2018.

PATTERSON, R. D. *et al.* The processing of temporal pitch and melody information in auditory cortex. **Neuron**, v. 36, n. 4, p. 767-776, 2002. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/12441063> Acesso em: 24 set. 2018.

PICKLES, J. O. Auditory pathways: anatomy and physiology. **Handbook of Clinical Neurology**, v. 129, p. 3-25, 2015. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/25726260> Acesso em: 24 set. 2018.

RIZZOLATTI, G.; CRAIGHERO, L. The mirror-neuron system. **Annual Reviews Neurosciences**, v. 27, p. 169-192, 2004. Disponível em: <http://psych.colorado.edu/~kimlab/rizzolatti.annurev.neuro.2004.pdf> Acesso em: 24 set. 2018.

ROSENBLUM, L. D. Speech perception as a multimodal phenomenon. **Current Directions in Psychological Science**, v. 17, n. 6, p. 405-409, 2008. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3732050/> Acesso em: 15 set. 2018.

SCOTT, S. K.; JOHNSRUDE, I. S. The neuroanatomical and functional organization of speech perception. **Trends in Neurosciences**, v. 26, n. 2, p. 100-107, 2003. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/12536133> Acesso em: 16 set. 2018.

SITEK, K.; JOHNSON, K. Ipsilateral and contralateral phonetic context effects. **Proceedings from the Annual Meeting of the Chicago Linguistic Society**, v. 48, n. 1, p. 535-546, 2012. Disponível em: <https://www.ingentaconnect.com/content/cls/pcls/2012/00000048/00000001/art00036> Acesso em: 15 set. 2018.

TOSCANO, J. C. **Perceiving speech in context: compensation for contextual variability during acoustic cue encoding and categorization**. 2011. 203p. Teses (Doctor in Philosophy) – University of Iowa, Iowa City. Disponível em: <https://ir.uiowa.edu/etd/1185/> Acesso em: 12 out. 2018.

TIIPPANA, K. What is the McGurk Effect? **Frontiers in Psychology**, v. 5, n. 725, p. 1-3, 2014. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4091305/> Acesso em: 24 set. 2018.

UNGERLIEDER, L. G.; HAXBY, J. V. “What” and “where” in the brain human. **Current Opinion in Neurobiology**, v.4, p. 157-165, 1994. Disponível em: [http://psych.colorado.edu/~kimlab/ungerleider\\_haxby.94.pdf](http://psych.colorado.edu/~kimlab/ungerleider_haxby.94.pdf) Acesso em: 11 out. 2018.

VINNER, J. A.; LEE, C. C. The distributed auditory cortex. **Hearing Research**, v. 229, n. 1-2, p. 3-13, 2007. Disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/17329049> Acesso em: 24 set. 2018.

VISWANATHAN, Navin; MAGNUSON, James S.; FOWLER, Carol A. Compensation for coarticulation: Disentangling auditory and gestural theories of perception of coarticulatory effects in speech. **Journal of Experimental Psychology: Human**

**Perception and Performance**, v. 36, n. 4, p. 1005, 2010. Disponível em: <https://psycnet.apa.org/record/2010-15881-016> Acesso em: 20 ago. 2020.

WATKINS, K. E.; STRAFELLA, A. P.; PAUS, T. Seeing and hearing speech excites the motor system involved in speech production. **Neuropsychologia**, v. 41, n. 8, p. 989-994, 2003 disponível em: <https://www.ncbi.nlm.nih.gov/pubmed/12667534> Acesso em: 17 set. 2018.

**Recebido em:** junho de 2020.  
**Aprovado em:** agosto de 2020.

**Como citar este trabalho:**

---

LIMA, W. F. O papel das dicas articulatórias na percepção da fala. Uma concepção remodelada de uma antiga questão. **Traços de Linguagem**, v. 4, n. 1, p. 61-74, 2020.

---