

# CIÊNCIA DA FALA: DESAFIOS TEÓRICOS E METODOLÓGICOS

## *Speech Science: theoretical and methodological challenges*

Thaís Cristófaró Silva\*

Maria Cantoni\*\*

Liliane Barbosa\*\*\*

### RESUMO

Este artigo tem por objetivo principal discutir aspectos teóricos e metodológicos relacionados com a Ciência da Fala na expectativa de apontar desafios atuais e indicar possíveis caminhos para avanço nesta área do conhecimento. Na área da Linguística a fonética e a fonologia são as disciplinas que tradicionalmente se ocupam da investigação da fala. Outros domínios do conhecimento também dedicam atenção à Ciência da Fala, dentre estes a psicologia, a fonoaudiologia, a neurociência e a tecnologia de fala. Assim, este artigo deverá contribuir com o debate teórico e metodológico na área da Linguística e também em outros domínios do conhecimento.

Palavras-chave: *Ciência da Fala; Cognição; Sistemas dinâmicos.*

### ABSTRACT

This paper intends to discuss theoretical and methodological aspects related to Speech Science hoping to indicate current challenges and to suggest possible pathways to advance research in this area of knowledge. In Linguistics the disciplines of phonetics and phonology are traditionally the ones that study

\* UFMG, CNPq, FAPEMIG.

\*\* UFMG, CAPES.

\*\*\* UFMG, UNIMONTES, FAPEMIG.

speech. Other areas of research also address issues related to Speech Sciences such as psychology, speech therapy, neurosciences, and speech technology. Thus, this article should contribute towards the theoretical and methodological debate in Linguistics as well as in other areas of knowledge.

Keywords: *Speech Science; Cognition; Dynamic Systems.*

## 1. INTRODUÇÃO

A Ciência da Fala é o domínio do conhecimento de caráter inter-, trans- e multidisciplinar que tem por objetivo investigar a produção, a transmissão e a percepção da fala enquanto ruído que tenha interpretação linguística e que seja utilizado na comunicação humana. Adicionalmente, a Ciência da Fala, em sentido amplo, investiga qualquer aspecto relacionado com a interpretação linguística dos padrões sonoros utilizados na comunicação humana, como, por exemplo, os movimentos ou gestos relacionados com a expressividade linguística (levantamento das sobrelhas, movimento dos braços, etc.), bem como aspectos relacionados com os correlatos neurofisiológicos da linguagem humana. Embora o estudo do conhecimento gramatical não seja foco central de investigação da Ciência da Fala, observa-se que os resultados obtidos pelas pesquisas nessa área oferecem subsídios importantes para a compreensão da organização gramatical da linguagem humana. Este artigo tem por objetivo discutir aspectos teóricos e metodológicos relacionados com a Ciência da Fala na expectativa de apontar desafios atuais e indicar possíveis caminhos para o avanço nesta área do conhecimento.

Além desta seção inicial, este artigo conta com outras quatro seções. A segunda seção discute a relação das ciências cognitivas com a Ciência da Fala. A terceira seção discute desafios teóricos impostos à investigação do ruído da fala, que é utilizado para propósitos de comunicação. A quarta seção, intitulada Desafios Metodológicos, apresenta vários equipamentos que são atualmente utilizados na investigação da fala e que podem contribuir

com avanços importantes para a Ciência da Fala. Por fim, a quinta seção apresenta as conclusões do artigo e é seguida da bibliografia.

## 2. CIÊNCIAS COGNITIVAS E CIÊNCIA DA FALA

As Ciências Cognitivas englobam várias linhas de investigação de natureza inter-, trans- e multidisciplinar que têm em comum a busca em compreender os vínculos existentes entre a mente e o cérebro. Neste sentido, as Ciências Cognitivas estabelecem diálogo entre diferentes áreas do conhecimento na formulação de elementos teóricos sobre a cognição humana, ou seja, como o ser humano pensa, manifesta-se, percebe, aprende e apreende. A inter-, trans- e multidisciplinaridade das Ciências Cognitivas se realiza pelo estabelecimento de relações entre várias áreas do conhecimento como a Linguística, a Psicologia, a Filosofia, as Neurociências e a Inteligência Artificial. Mais recentemente, o escopo das Ciências Cognitivas foi ampliado para as áreas da Educação (ensino-aprendizagem), da Engenharia (em várias modalidades), das Ciências Sociais (Sociologia, Antropologia e História do Conhecimento), das Ciências da Saúde (Saúde Coletiva, Psiquiatria), das Ciências Sociais Aplicadas (Gestão de Informação, Estética de Massa e Linguagens), dentre outras áreas que estudam o gerenciamento do conhecimento pela espécie humana.

Embora seja bastante amplo o escopo de investigação das Ciências Cognitivas, em suas diversas modalidades, há foco preciso de estudo em cada área do conhecimento. A Linguística Cognitiva investiga os mecanismos envolvidos no processamento e categorização da linguagem e, portanto, compartilha pontos de interesse com a Ciência da Fala. Destaca-se aqui que o rótulo de Linguística Cognitiva é também adotado com regularidade em modelos formais de investigação da linguagem como, por exemplo, os modelos gerativos e as demais teorias denominadas *chomskyanas*. A Linguística Cognitiva compartilha com a Linguística Gerativa o fato de ambas assumirem que a linguagem é um fenômeno relacionado com a mente. Contudo, a Linguística Gerativa investiga a linguagem do ponto de vista formal, por meio da representação simbólica categórica, assumindo a

existência de um módulo específico responsável pela faculdade da linguagem, cuja natureza é primordialmente sintática e é regida por princípios formais específicos. Por outro lado, a Linguística Cognitiva investiga a linguagem do ponto de vista conceitual e funcional, por meio de representações dinâmicas e multidimensionais, assumindo que o conhecimento linguístico compartilha de propriedades gerais de organização que operam no gerenciamento do conhecimento pela espécie humana. Portanto, a dinamicidade e a multirrepresentacionalidade das representações, bem como a complementaridade do conhecimento linguístico com o conhecimento geral da espécie humana, são aspectos da Linguística Cognitiva que a separam da tradição gerativista ou chomskyana.

O estudo do ruído linguístico utilizado na comunicação humana é objeto de investigação da Linguística Cognitiva por incorporar o caráter dinâmico, multirrepresentacional e complementar da linguagem, e é parte da Ciência da Fala por investigar a produção, a transmissão e a percepção de padrões sonoros com interpretação linguística utilizados na comunicação humana.

Elman (1998) cita três abordagens cognitivas distintas e recentes voltadas para o processamento cognitivo, sendo elas: Conexionismo, Vida Artificial e Sistemas Dinâmicos. Tais abordagens têm recebido grande atenção nas últimas décadas e muitas de suas ideias refletem as mudanças paradigmáticas que tiveram lugar na investigação do gerenciamento do conhecimento pela espécie humana. Essas três abordagens, apesar de diferirem entre si, possuem diversos pontos em comum, o que possibilita que muitos pesquisadores trabalhem simultaneamente com elas (ELMAN, 1998). Neste artigo, trataremos do Conexionismo e dos Sistemas Dinâmicos com o objetivo de contextualizar os desafios teóricos e metodológicos impostos à Ciência da Fala.

Há diversos modelos incluídos sob o rótulo de Conexionismo, porém quase todos compartilham certas características e aderem ao esquema geral: nesses modelos, o processamento é conduzido por elementos, nós ou unidades, que possuem uma dinâmica análoga à dos neurônios. Cada nó recebe *input* de outros nós, responde a esse *input* e excita ou inibe outros nós aos quais está conectado e que estão ao seu redor. O processamento da

linguagem se dá através do ajuste entre as conexões feitas entre os nós na recepção de *inputs* e produção de *outputs* que determinam uma rede. Com esta abordagem, pretende-se construir uma modelagem de base biológica que reproduza a forma como a mente humana categoriza, armazena e gerencia informações linguísticas.

As abordagens conexionistas foram marcantes para a Psicologia Cognitiva e para a consolidação de modelos de organização do conhecimento em rede. A proposta de Rumelhart e McClelland (1986) e os trabalhos de Elman (1991, 1998, 1998) são marcos importantes do Conexionismo. Avanços no modelo original sugerem algoritmos de aprendizagem que possibilitam que uma determinada rede se programe, ou seja, as redes seriam expostas a exemplos de um comportamento-alvo e, através da aprendizagem, a rede seria capaz de ajustar seus pesos ao longo do tempo; assim, as respostas da rede seriam aperfeiçoadas. O ponto central a ser frisado neste artigo com relação ao Conexionismo é a importância de as relações entre vários níveis de uma rede serem concebidas como paralelas e distribuídas.

Com relação aos Sistemas Dinâmicos, por outro lado, o ponto central que nos interessa neste artigo é o desdobramento temporal dos eventos. O conceito de Sistemas Dinâmicos surgiu da necessidade de se conceber um arcabouço teórico para analisar os sistemas que se desenvolvem de acordo com uma estrutura que relaciona um estado presente a estados passados. Ou seja, um modelo que considere o fator temporal em suas análises, já que no decorrer do tempo um sistema pode passar por estados diferentes. O modelo de sistemas dinâmicos teve origem na matemática e na física matemática e parte da premissa de que processos acontecem em tempo real e com dinamicidade.

Port e van Gelder (1995) apontam razões para se pensar a cognição como um sistema dinâmico: os comportamentos cognitivos não são atemporais (na verdade, existem e desdobram-se no tempo) e os sistemas cognitivos alteram-se frequente e continuamente (os sistemas dinâmicos admitem modelar e formalizar as diversas e concomitantes interações em um sistema). Assim, um sistema dinâmico teria a habilidade de se auto-organizar e se desenvolver por meio de sua evolução própria.

Thelen e Bates (2003) sugerem que o conexionismo e os sistemas dinâmicos seriam antes abordagens diferentes de uma mesma teoria nova e sintética do desenvolvimento humano, por reunirem evidências da neurobiologia, física, matemática e ciências computacionais. Os dois modelos difeririam muito pouco e, principalmente, difeririam em decorrência de especificidades de sua trajetória histórica: enquanto os Sistemas Dinâmicos teriam emergido primeiramente do campo relacionado à ação e percepção, tendo obtido uma maior repercussão em estudos sobre o desenvolvimento sensório-motor, o Conexionismo surgiu a partir da modelagem computacional, enfocando, principalmente, mapeamentos internos cerebrais e seus conteúdos representacionais.

Entendemos que a Ciência da Fala se beneficia de resultados de pesquisas tanto do Conexionismo, quanto dos Sistemas Dinâmicos. Sendo a Ciência da Fala responsável por investigar a produção, a transmissão e a percepção dos padrões sonoros utilizados na comunicação humana, ela se adequa a uma abordagem dinâmica, que avalie a coordenação de movimentos e seus correlatos cognitivos. A próxima seção avalia o problema da indefinição da unidade de análise da Ciência da Fala e indica modelos fonológicos que se beneficiaram de desenvolvimentos do Conexionismo e dos Sistemas Dinâmicos: Fonologia Gestual, Fonologia de Uso e Teoria de Exemplos.

### 3. DESAFIOS TEÓRICOS

Um dos pontos centrais ao se definir uma área do conhecimento é determinar seu objeto de estudo e sua unidade primária de análise. A Ciência da Fala, por ter natureza inter-, trans- e multidisciplinar, encontra desafios na tarefa de determinar qual seria a unidade básica de análise. Moore (1997), indica a dificuldade em se definir a unidade de análise na percepção. Contudo, as ponderações elencadas por ele são as mesmas observadas em aspectos de produção da fala.

It is still not clear whether the basic unit of perception is the syllable, the phoneme, or some other unit such as the phonetic feature. For rapid connected speech, the

psychoacoustic evidence indicates that the acoustic pattern signaling individual phonetic features or phonemes would occur too rapidly to be separately perceived in correct temporal order. Thus, the recognition of the overall sound pattern corresponding to a longer segment, such as a syllable, is more likely (MOORE, 1997, p. 303).

Um dos temas do debate atual na Ciência da Fala é, justamente, qual seria a unidade básica de investigação e análise. De fato, neste momento não sabemos precisar qual é unidade básica de investigação, como indicado pela citação de Moore (1997). Da busca pela unidade básica de investigação desdobra-se o debate adicional sobre a organização das unidades de investigação e análise. Assim, abordagens conexionistas e de sistemas dinâmicos são aparatos teóricos valiosos na busca de uma maior compreensão sobre a organização temporal e representacional do ruído linguístico que é utilizado na comunicação humana.

As disciplinas da Linguística que tratam dos padrões sonoros utilizados na comunicação humana são a Fonética e a Fonologia. Embora em décadas passadas estas disciplinas tenham sido tratadas como domínios separados do conhecimento, hoje, elas são vistas como disciplinas complementares. Essas duas disciplinas compartilham, com outras áreas do conhecimento, a investigação da Ciência da Fala. No restante desta seção serão apresentados, brevemente, três modelos teóricos que tratam da Fonética e da Fonologia como disciplinas complementares, voltadas para a organização do conhecimento sonoro da linguagem. Serão apresentados os seguintes modelos: Fonologia Gestual (BROWMAN; GOLDSTEIN, 1992), Fonologia de Uso (BYBEE, 2001) e Teoria de Exemplos (PIERRREHUMBERT, 2001). Estas abordagens se aproximam por defenderem o caráter emergente e dinâmico da fala, bem como por proporem algum tipo de modelagem em rede para a organização do conhecimento linguístico. Adicionalmente, estes modelos sugerem uma relação estreita entre a produção e a percepção da fala na construção do conhecimento linguístico e sugerem a investigação de correlatos físicos da linguagem, além de postularem indicadores probabilísticos no gerenciamento linguístico.

### 3.1 FONOLOGIA ARTICULATÓRIA OU FONOLOGIA GESTUAL

A Fono­logia Articula­tória ou Fono­logia Ges­tual pos­tula que a fala é um sistema de produção de gestos articu­la­tórios que se sobrepõem no tempo. O modelo sugere a independência das ações dos órgãos de articulação (lábios; ponta, corpo frontal, corpo dorsal e raiz da língua; subsistemas autônomos do véu palatino e mandíbula) na organização da linguagem. O modelo gestual assume que a fala esteja “sujeita aos mesmos princípios que a motricidade humana/animal e, portanto, passível de descrição análoga à da gestualidade manual, da marcha, da dança, do esporte, etc.” (ALBANO, 2009, p. 3).

O gesto é a unidade básica de análise, caracteriza os eventos articu­la­tórios na dimensão do tempo e é compreendido como uma unidade discreta de análise fônica, cuja natureza é simbólica e, ao mesmo tempo, dinâmica. O gesto pode ser compreendido como uma unidade sensório-motora, que implementa os movimentos dos articuladores e também como uma unidade simbólica, que é fruto da internalização da ação ou gesto. Os níveis simbólico e dinâmico são, portanto, complementares.

Diferentes gestos podem ser sobrepostos, i.e., dois ou mais gestos podem co-ocorrer temporalmente. Enunciados são constelações de gestos que se conectam na organização do componente sonoro. A proposição da teoria foi apresentada inicialmente em Browman e Goldstein (1986, 1990, 1992). Aplicações deste modelo teórico ao português brasileiro, incluindo contribuições teóricas adicionais, são Albano (2001, 2004).

### 3.2 FONOLOGIA DE USO

A Fono­logia de Uso (BYBEE, 2001) sugere que a Gramática de uma língua emerge do uso. A palavra é a unidade representacional, embora a categorização de unidades linguísticas possa operar em vários níveis (por exemplo, silábico, morfológico, semântico, etc.). A Fono­logia de Uso propõe que todas as variantes de uma palavra a que um dado falante foi exposto (percepção e produção da fala) estão armazenadas e categorizadas em sua



memória. Efeitos probabilísticos atuam na organização lexical, sendo que o protótipo de uma variante é definido por mecanismos de frequência. Assim, as representações são múltiplas e encontram-se organizadas em redes que operam em vários níveis (silábico, segmental, acentual, etc.). As conexões ou redes contêm informações linguísticas e não linguísticas, ou seja, aspectos sociais e mesmo características da voz específica de uma determinada pessoa são armazenadas e utilizadas no processamento linguístico. Como foi mencionado acima, efeitos de frequência caracterizam o protótipo de uma variante levando-se em consideração vários parâmetros linguísticos e não linguísticos.

A Fonologia de Uso se adequa aos modelos gramaticais baseados no uso por assumir que a experiência é crucial para a organização do conhecimento linguístico e fonológico em particular (BARLOW; KEMMER, 2000; BYBEE, 1999, 2001; KRISTIANSEN, 2006; LANGACKER, 1988). Adicionalmente, a Fonologia de Uso sugere que o conhecimento linguístico seja organizado probabilisticamente e adota a Teoria de Exemplos, que será apresentada na seção seguinte, como modelo representacional. Aplicações desse modelo ao português brasileiro são Cristóvão Silva (2003, 2005, 2008) e Cristóvão Silva e Gomes (2005, 2008).

### 3.3 TEORIA DE EXEMPLARES

De maneira análoga à Fonologia de Uso, a Teoria de Exemplos assume que a palavra é a unidade representacional. O detalhe fonético das palavras é armazenado pelo falante e é utilizado no processamento do conhecimento linguístico. As representações são múltiplas e gerenciadas probabilisticamente. A frequência lexical é de fundamental importância para o armazenamento, organização e processamento do conhecimento linguístico. A proposição da Teoria de Exemplos é apresentada em Pierrehumbert (2001), seguindo proposta de Johnson (1997) e Goldinger (1997).

A grande crítica à Teoria de Exemplos e também à Fonologia de Uso diz respeito ao gerenciamento de toda e qualquer informação linguística

a que o falante é exposto. Embora seja controverso, este é um tema que demanda maiores estudos, uma vez que a memória humana parece ser capaz de armazenar uma enorme quantidade de informações (JOHNSON, 1997; BYBEE, 2001).

### 3.4 SOBRE A CIÊNCIA DA FALA

Os modelos teóricos que foram apresentados nas seções precedentes são compatíveis com princípios de investigação da Ciência da Fala. Para finalizar esta seção, gostaríamos de apontar a contribuição que a Linguística, e suas várias disciplinas, pode dar para a Ciência da Fala.

Um dos pontos principais da investigação é a busca e identificação de padrões recorrentes, que expressem generalizações do comportamento linguístico. De maneira similar, é também de interesse da linguística a investigação de padrões anômalos ou inexistentes. A Linguística pode, portanto, contribuir com a identificação de aspectos distribucionais relacionados com a linguagem e em particular com a sonoridade. Mesmo que ainda tenhamos dúvidas quanto às unidades representacionais da fonologia, faz-se necessária a busca de generalizações que expressem o comportamento do componente fonológico. A identificação de padrões fonológicos específicos permite a investigação das mudanças sonoras, as quais expressam o caráter dinâmico da sonoridade e oferecem maior compreensão sobre a evolução das línguas.

A investigação do detalhe fonético na produção e na percepção são domínios da Ciência da Fala que podem contribuir para melhorar o entendimento de parâmetros essenciais e determinantes de padrões sonoros específicos (sejam estes morfemas, sílabas, etc.). O conhecimento da motricidade envolvida na articulação da fala deve contribuir com um melhor conhecimento das rotinas motoras e dos possíveis ajustes finos, que permitem a criação e o gerenciamento de categorias sonoras.

Possivelmente, rotinas motoras devem definir mecanismos que se relacionam com categorias específicas que vinculam a produção com a

percepção. Tais rotinas motoras lembrarão, talvez, a ideia de uma categoria abstrata semelhante ao que hoje se postula como fonema. Contudo, tais categorias não serão estáticas, como a noção atual de fonema, mas terão dinamicidade e inter-relação entre si, permitindo explicações sobre a variação e a mudança linguística.

Neste estágio de investigação, mesmo com toda a precariedade inerente a uma ciência que se desenvolve faz apenas poucas décadas, podemos vislumbrar um futuro desafiador e gratificante para a Ciência da Fala. Desafiador porque novos conhecimentos gerarão novas perguntas que deverão ser respondidas. Gratificante porque a colaboração entre as várias áreas do conhecimento deverá oferecer respostas sobre a linguagem, com e sem patologia, e sobre a organização das línguas naturais.

#### 4. DESAFIOS METODOLÓGICOS

Um dos problemas enfrentados pelas teorias apresentadas na seção anterior diz respeito aos procedimentos metodológicos da investigação. Em decorrência de avanços tecnológicos verificados nas últimas décadas, uma ampla gama de equipamentos passou a ser utilizada na investigação dos padrões sonoros da comunicação humana. Os equipamentos a serem apresentados ao longo desta seção podem contribuir de maneiras diferentes e, muitas vezes, complementares, para uma maior compreensão dos estudos relacionados com a Ciência da Fala.

Nos modelos teóricos apresentados anteriormente, destaca-se a ideia de que os processos cognitivos envolvidos na linguagem humana constituem e são constituídos pelas rotinas motoras que compõem a articulação da fala, ideia amplamente explorada pelas teorias de base gestual e que é também compartilhada pela Fonologia de Uso e Teoria de Exemplos. Assim, fica cada vez mais patente aos modelos fonológicos a necessidade de integração de dados empíricos acerca da natureza acústico-articulatória da fala. A seguir, serão apresentados equipamentos e recursos tecnológicos capazes de proporcionar medições objetivas e precisas da fala humana, considerada sob

diversos aspectos. Informações mais detalhadas podem ser obtidas, por exemplo, em Baken (1996), Harcastle, Laver e Gibbon (2010), Ladefoged (2003) e Solé, Beddor e Ohala (2007).

#### 4.1 MEDIDAS ACÚSTICAS

##### MICROFONE

Capta o sinal da fala. É comumente classificado de acordo com a direcionalidade de sua captação: se são mais sensíveis ao sinal proveniente de uma única direção (unidirecionais); se mais sensíveis ao sinal proveniente de duas direções opostas (bidirecionais) ou se igualmente sensíveis em todas as direções (onidirecionais). O tipo mais adequado à gravação de um único informante - situação mais comum nos estudos da fala - é o microfone unidirecional, pois ele capta o sinal de fala do informante com precisão e ignora possíveis sons provenientes de outras fontes ao redor (por exemplo, ruídos). O correto posicionamento do microfone é um detalhe importante no protocolo experimental, que, na Ciência da Fala, geralmente prevê a gravação de informantes em posição fixa e assentados. Nessa situação, é recomendável que os microfones mais comuns, que consistem em um globo afixado a uma haste curta, sejam afixados a suportes próprios, a uma altura e distância confortáveis ao informante. Uma excelente alternativa é utilizar microfones de lapela ou microfones de cabeça, que tornam mais livre a movimentação natural do informante ao falar e mantêm constante a distância entre o mecanismo de captação e a fonte de emissão do sinal de fala.

##### AMBIENTE DE GRAVAÇÃO

Na gravação da fala, a qualidade do sinal acústico está diretamente relacionada a uma baixa incidência de ruído. Para tanto, o mais recomendado é, quando possível, realizar o procedimento dentro de uma cabine com isolamento acústico. No caso de pesquisa de campo, deve-se procurar sempre evitar a proximidade de fontes de sons indesejados, como aparelhos eletroeletrônicos, animais, cursos de água, outras pessoas, etc.

Contudo, mesmo as gravações realizadas em cabine acústica apresentam, em geral, algum nível de ruído. Uma solução possível é a posterior filtragem digital do sinal da fala, capaz de remover ruídos bem localizados em certa faixa de frequência. A maioria dos programas de análise acústica oferece filtros com essa finalidade.

#### AQUISIÇÃO E DIGITALIZAÇÃO DO SINAL DE FALA

Geralmente realizadas diretamente em computadores, por meio de programas de análise acústica, ou em gravadores digitais. Deve-se regular a intensidade da aquisição, de forma a conseguir o maior nível de sinal sem que haja, contudo, saturação (ou clipagem, perda de informação). De grande importância é a atenção à taxa de amostragem e de quantização, na digitalização dos dados: são amplamente usados os valores de 22050 Hz e 16 bits, respectivamente, pois, proporcionam um intervalo e níveis suficientes para garantir a fidelidade do sinal digital e são compatíveis com a maioria dos programas.

#### PROGRAMAS DE ANÁLISE E EDIÇÃO DE ÁUDIO

Dentre os programas atualmente mais utilizados por linguistas para análise, edição e manipulação do sinal da fala, podem-se citar, como referência, o Praat (<http://www.fon.hum.uva.nl/praat/>), o Winpitch (<http://www.winpitch.com/>), o Audacity (<http://audacity.sourceforge.net/>) e o Adobe Audition (antigo CoolEdit, <http://www.adobe.com/products/audition/index.html>).

O sinal acústico é apenas uma das diversas dimensões da fala que podem ser medidas. A seguir, serão brevemente apresentados vários equipamentos que, seja há várias décadas ou somente nos últimos anos, têm sido amplamente empregados no estudo experimental da fala humana. Esses equipamentos são geralmente utilizados concomitantemente com um microfone, interno ou externo ao sistema, que irá captar o som da fala para sincronização ou referência.

## 4.2 MEDIDAS ARTICULATÓRIAS E AERODINÂMICAS

### CÂMERA DE VÍDEO

A gravação em vídeo tem estado cada vez mais presente nos protocolos experimentais dos estudos de fala, não só como forma de documentação da coleta de dados, mas como recurso ativo para medição dos movimentos da face do informante na produção de sons. Por meio de um bom programa de análise, é possível, por exemplo, isolar o movimento de articuladores individuais e comparar objetivamente suas trajetórias no tempo, de forma sincronizada com o sinal acústico.

### PROGRAMA DE ANÁLISE E EDIÇÃO DE VÍDEO

Um excelente programa que tem sido amplamente utilizado para análise de vídeos é o Anvil (<http://www.anvil-software.de/>), mas há também outras opções disponíveis, como o VCode & Vdata (<http://social.cs.uiuc.edu/projects/vcode.html>) e o Elan (<http://www.lat-mpi.eu/tools/elan>).

### CAPTURA DE MOVIMENTO

Além da gravação em vídeo, o movimento dos articuladores na fala pode ser rastreado diretamente por equipamentos de captura tridimensional de movimento, como os que serão brevemente apresentados a seguir.

O *Optotrak Certus* (<http://www.ndigital.com/lifesciences/certus-motioncapturesystem.php>) é um mecanismo de captura de movimento que consiste em um sistema de três câmeras acopladas linearmente que captam o deslocamento tridimensional de marcadores ativos posicionados previamente em um corpo em movimento, com resolução de até 0,1mm. O sinal do *Optotrak*, portanto, pode oferecer um registro acurado da movimentação de articuladores faciais envolvidos na realização da fala, através de marcadores posicionados em regiões estratégicas da face dos informantes.

O *EMA* (Articulografia Médio-Sagital Eletromagnética, <http://www.articulograph.de/>), em sua versão 3D, consiste em um sistema de

seis bobinas que geram um campo eletromagnético e captam a posição de sensores em uma esfera de medição de 300mm. Os sensores são captados digitalmente pelas bobinas em três dimensões e dois ângulos. Duas grandes vantagens do EMA para pesquisa em fonética são: os sensores podem ser colocados dentro da boca e não têm que estar visíveis para que sua movimentação seja captada.

#### ULTRASSONOGRAFIA

Além de suas aplicações clínicas, o ultrassom pode ser também aplicado à análise de fala. Tem como vantagens ser um método não invasivo, que não apresenta riscos à saúde, e ser de baixo custo. Seu funcionamento se baseia na transmissão e recepção de sinal de alta frequência (acima de 1 MHz). O sinal enviado pelo transmissor vai sendo refletido progressivamente à medida que atravessa diferentes camadas de tecido. Ao atingir uma região preenchida com ar, que não é um bom condutor de ultrassom, o restante do sinal é completamente refletido e é finalizada a transmissão. Na análise de fala, esse equipamento é empregado com sucesso para monitorar os movimentos da língua, proporcionando imagens de todo o seu contorno. O aparelho transmissor é posicionado abaixo da mandíbula do informante, apontado para o topo da cabeça. Outra aplicação da ultrassonografia é a análise da movimentação das cordas vocais, com o aparelho transmissor posicionado no pescoço do informante. Contudo, a pequena dimensão das cordas vocais e a rapidez de seus movimentos exigem uma resolução que desafia os limites do equipamento, que tem dificuldades em fornecer detalhes finos da abertura e fechamento e da movimentação vertical das bordas das cordas vocais.

#### PALATOGRAFIA

Utilizada para realizar medidas articulatórias do contato da língua com a região superior do trato vocal na realização de sons com algum nível de articulação palatal.

Na *Palatografia Estática* (Fotopalatografia), cobre-se a língua do informante com alguma substância escura, atóxica (como uma mistura de

azeite de oliva e carvão em pó, que, por exemplo, é usada atualmente na UCLA (2003)). Essa mistura será transferida para as regiões do palato envolvidas na articulação de um som palatal. Com auxílio de um espelho em ângulo de 45° inserido dentro da cavidade oral, tira-se uma fotografia do palato do informante, que estará marcado nos pontos que sofreram contato com a língua. Se for feito o inverso (cobrir o palato, ao invés da língua, com substância escura), o método passa a ser uma Linguografia, em que a língua trará marcas nas regiões que entraram em contato com o palato. Nesse caso, a fotografia é realizada diretamente, sem necessidade de espelho. Tem como vantagens ter baixo custo e poder ser utilizada por diversos falantes. Já entre suas desvantagens estão a lentidão e perda da naturalidade da fala – a cada contato com o palato, é necessário limpar a língua e o palato do informante e aplicar novamente a substância – e assim proceder a uma nova coleta de dados e de aquisição das imagens através de equipamentos específicos.

Na *Palatografia Dinâmica* (Eletropalatografia), o contato da língua no palato é captado por um palato artificial (uma placa fina contendo cerca de 70 eletrodos), confeccionado sob medida. Uma vez posicionado no palato do informante, o palato artificial é conectado a um equipamento que irá captar e processar os dados. Entre suas vantagens, estão a possibilidade de realizar os sons palatais em sequência, sem interrupção, e a aquisição e processamento automático dos dados. Suas desvantagens consistem no alto custo e na necessidade de confeccionar um palato artificial para cada informante.

#### ELETROGLOTOGRAFIA (EGG) OU ELETROLARINGOGRAFIA

Utilizado para avaliar a dinâmica da articulação ao nível glótico, é um dos poucos equipamentos que não se restringe à análise de vogais sustentadas, podendo ser empregado igualmente para análise de fala contínua. O sinal do EGG representa a área e a extensão longitudinal do contato das cordas vocais. Os modelos mais comuns consistem basicamente em dois eletrodos de contato que, posicionados na altura da cartilagem tireóide do informante, medem a diferença de impedância resultante da abertura e fechamento das cordas vocais. É um método não invasivo, de fácil



operação, mas que é pouco preciso quanto aos eventos que ocorrem durante a fase aberta das cordas vocais e é sujeito à interferência pelo deslocamento vertical da laringe e pela presença de muco nas cordas vocais.

#### MÁSCARAS FACIAIS

Envoltórios plásticos ou de borracha cobrem completamente a boca e/ou o nariz, utilizados para medir o fluxo oral e nasal. Adicionalmente, tubos integrados ou não a máscara e conectados a um aparelho transdutor de pressão podem ser empregados para medir a pressão intraoral, faríngea e esofágica (estimativa da pressão subglótica). Ao utilizar máscaras faciais, deve-se ter em mente a necessidade de calibrar os sistemas que medem fluxo de ar e pressão através de equipamentos específicos, bem como a necessidade de isolar no sinal acústico o formante introduzido pela máscara. Como exemplo de máscaras muito utilizadas para análise aerodinâmica da fala, pode-se citar a conhecida *Máscara de Rothenberg* (1973) e a mais recentemente lançada máscara da estação de trabalho *EVA* (GIOVANNI et al., 1996).

A pressão subglótica pode ser medida diretamente por meio de *Punção da Traquéia* com uma agulha de calibre grande, que é então conectada a um tubo ligado a um transdutor de pressão. Trata-se de uma forma de medição bastante acurada, mas com o inconveniente de ser invasiva e de depender de profissional da saúde para realizar a punção.

#### 4.3 MEDIDAS COGNITIVAS

A avaliação das atividades cerebrais durante o processamento da linguagem é de grande interesse para a Ciência da Fala. Idealmente, as medidas de ordem cognitiva oferecerão informações sobre os processos cognitivos envolvidos com a linguagem. Adicionalmente, em casos de patologia que dizem respeito à perda ou alteração da linguagem, as medidas cognitivas podem oferecer informações sobre a plasticidade do cérebro.

## IMAGEAMENTO CEREBRAL

Para concluir essa seção, serão apresentados brevemente dois métodos não invasivos de imagem dinâmica da atividade cerebral, que podem ser usados para avaliar os processos cognitivos envolvidos na produção de fala. Como exibem características complementares, esforços vêm sendo feitos para compatibilizar o uso simultâneo dos dois equipamentos. Além de seu custo elevado, outra grande desvantagem que apresentam para a pesquisa é o fato de terem de ser operados por médicos em protocolos experimentais envolvendo sujeitos humanos. Isso, contudo, não inviabiliza sua utilização para pesquisa em linguística, uma vez que o estudo avançado da linguagem humana tem demandado cada vez mais a formação de equipes multidisciplinares.

A *Ressonância Magnética Funcional (fMRI)* mede a atividade neural a partir das alterações magnéticas que acompanham as alterações locais do fluxo sanguíneo em áreas ativadas. Seu funcionamento se baseia no princípio de que o envio de oxigênio a um grupo de neurônios em atividade causa pequenas alterações locais no campo magnético. Apresenta boa precisão espacial, mas uma precisão temporal reduzida.

A *Eletroencefalografia (EEG)* mapeia a atividade cerebral a partir da amplificação dos padrões de atividade elétrica dos neurônios em funcionamento. Apresenta boa precisão temporal, mas uma precisão espacial reduzida.

## 5. CONCLUSÃO

Este artigo discutiu aspectos teóricos e metodológicos relacionados com a Ciência da Fala. Inicialmente, buscou-se contextualizar a Ciência da Fala em relação às Ciências Cognitivas indicando o caráter inter-, trans- e multidisciplinar dos estudos relacionados com o ruído linguístico utilizado na comunicação humana. Foram apresentados três modelos teóricos da fonologia que são compatíveis com o caráter dinâmico e de organização do conhecimento em redes: Fonologia Gestual (BROWMAN; GOLDSTEIN, 1991), Fonologia de Uso (BYBEE, 2001) e Teoria de Exemplares (PIERREHUMBERT,

2001). Indica-se que dentre os desafios teóricos centrais temos a identificação da unidade básica de análise e o gerenciamento do sistema de tais unidades. Em busca de apoio metodológico para os desafios teóricos foi apresentada uma seção de apoio metodológico que teve por objetivo indicar equipamentos, técnicas e suporte aos estudos da Ciência da Fala. Sendo a fonética e a fonologia as disciplinas relacionadas com a Ciência da Fala, espera-se que este artigo contribua com avanços na área da Linguística.

## REFERÊNCIAS

ALBANO, E. C. *Fonologia gestual e aquisição do sistema fônico hoje*. SEMINÁRIO DE AQUISIÇÃO FONOLÓGICA. Santa Maria: UFSM, 2009.

BAKEN, R. *Clinical measurement of speech and voice*. San Diego/London: Singular Publishing Group, 1996.

BARLOW, M. e KEMMER, S. (Ed.). *Usage-Based models of language*. Cambridge: Cambridge University Press, 2000.

BROWMAN, C.; GOLDSTEIN, L. *Articulatory phonology: an overview*. *Phonetica*, n. 49. p. 155-180, 1992.

BYBEE, J. Usage-based phonology. In: DARNELL, M. et al (Ed.). *Functionalism and formalism in linguistics. I: general papers; II: case studies*. Amsterdam: Benjamins, 1999. p. 211-242.

\_\_\_\_\_. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

ELMAN, J. L. 1991. *Distributed representations, simple recurrent networks, and grammatical structure*. *Machine Learning*, n. 7, p. 195-224.

ELMAN, J. L. Language as a dynamical system. In: R. F. Port; T. van Gelder (Orgs.) *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press, 1995. p. 195-223.

ELMAN, J. L. *Connectionism, artificial life, and dynamical systems: new approaches to old questions*. Department of Cognitive Science: University of California, San Diego, 1998.

GIOVANNI, A.; ROBERT, D.; ESTUBLIER, N.; TESTON, B.; ZANARET, M.; CANNONI, M. *Objective evaluation of dysphonia: preliminary results of a device allowing simultaneous acoustic and aerodynamic measurements*. *Folia Phoniatr Logop.*, v. 48, p. 175-185, 1996.

HARDCASTLE, W.; LAVER, J.; GIBBON, F. (Ed.). *The handbook of phonetic sciences*. 2nd ed. Blackwell, 2010.

JONES, D. *The phoneme: its nature and use*. Cambridge: Heffer, 1967.

KRISTIANSEN, G. Towards a usage-based phonology. *International journal of English studies*, p.107-140, 2006.

LADEFOGED, P. *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Wiley-Blackwell, 2003.

LANGACKER, R. W. *Foundations of cognitive grammar*. Stanford: Stanford University Press, 1987.

\_\_\_\_\_. A usage-based model. In: RUDZKA-OSTYN, B. (Ed.). *Topics in cognitive linguistics*. Amsterdam: Benjamins, 1988. p. 127-161.

LIBERMAN, A.; MATTINGLY, I. The motor theory of speech perception revisited. In: *Cognition*. 1985, 21, p. 1-36.

NATHAN, G. S. What functionalists can learn from formalists in phonology. In: *Proceedings of the symposium on formalism and functionalism*. Amsterdam: Benjamins, 1999. p. 305-327.

PIERREHUMBERT, J. B. Exemplar dynamics: Word frequency, lenition and contrast. In: BYBEE, J.; HOPPER, P. (Ed.). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, 2001, p. 137-157.

\_\_\_\_\_. Word specific phonetics. In: GUSSENHOVEN, C.; WARNER, N. (Ed.). *Laboratory phonology 7*. The Hague: Mouton de Gruyter, 2002. p. 101-139.

PORT, R. F. e VAN GELDER, T. Representing aspects of language. In: *Proceedings of the thirteenth annual conference of the cognitive science society*. Hillsdale, N.J.: Erlbaum, 1991, p. 487-492.

\_\_\_\_\_. (Orgs.). *Mind as motion: explorations in the dynamics of cognition*. Cambridge, Massachusetts: The MIT Press, 1995.

PORT, R. F.; LEARY A. P. Against formal phonology. *Language*, vol. 81. n. 4, p. 927-964, 2005.

ROTHENBERG, M. *A new inverse-filtering for deriving the glottal airflow waveform during voicing*. J. Acoust. Soc. Am., v. 53, p. 1632-1645, 1973.

RUMELHART, D.; MCCLELLAND, J.. On learning the past tenses of English verbs. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 2. Psychological and Biological Models (Rumelhart, D.E. and McClelland, J.L., eds), MIT Press. p. 216-271.

SAPIR, E. The psychological reality of phonemes. In: MAKKAI, V. B. (Ed.). *Phonological theory: evolution and current practice*. New York: Holt Rinehart and Winston, 1972. p. 22-31.

SAUSSURE, Ferdinand. *Curso de linguística geral*. São Paulo: Cultrix, 2001.

SOLÉ, M. J.; BEDDOR, P.; OHALA, M. *Experimental Approaches to Phonology*. Oxford University

Press, 2007.

STAMPE, D. On phonological representation. In: DRESSLER, W. U. et al (Eds.). *Phonologica*, 1984. London: Cambridge University Press, 1987. p. 287–300.

THELEN, E.; BATES, E.. Connectionism and dynamic systems compared. *Developmental Sciences*, v. 6, n. 4, p. 378-391, 2003.

UCLA PHONETICS LABORATORY. *How to do Static Palatography*. 2003. Disponível em: <[http://www.linguistics.ucla.edu/faciliti/facilities/physiology/static\\_pal\\_new/webpal.htm](http://www.linguistics.ucla.edu/faciliti/facilities/physiology/static_pal_new/webpal.htm)>. Acesso em: 04/06/2010.

Submetido em: 30/03/2011

Aceito em: 09/06/2011

