

REPRESENTAÇÕES FONÉTICA E FONOLÓGICA: RUMO À PARCIMÔNIA¹

ELEONORA CAVALCANTE ALBANO
(LAFAPE - IEL – UNICAMP)

1. RETROSPECTIVA

Este trabalho versa sobre a questão da representação do conhecimento fônico e da necessidade de seguir ou não a prática, corrente na literatura, de desdobrá-lo em ao menos dois níveis, o fonológico e o fonético. Vou defender a perspectiva de que esse desdobramento é desnecessário e de que tanto a produção como a compreensão da fala supõem apenas um nível *permanente e específico* de representação lingüística, embora possam mobilizar representações temporárias menos específicas, de natureza auditiva, cinestésica, visual, etc.

A visão, hoje quase consensual, de que, para um mesmo enunciado, há uma representação fonológica e uma representação fonética distintas não é antiga. Ela data do esforço, aliás monumental, da Fonologia Gerativa para dar um tratamento explícito à questão da alofonia. Chomsky e Halle (1968) foram os primeiros a atribuir um estatuto teórico à distinção, já praticada desde a década de 20, entre transcrição “fonêmica” ou fonológica e transcrição fonética. O famoso Capítulo 7 desse livro, intitulado “*The Phonetic Framework*”, afirma textualmente:

“Given the surface structure of a sentence the phonological rules of a language interact with certain universal phonetic constraints to derive all grammatically determined facts about the production and perception of the sentence. These facts are embodied in the “phonetic transcription”.”

(Chomsky, N. & M. Halle. *The Sound Pattern of English*. New York: Harper & Row, 1968, p. 293)

A cisão estabelecida por Chomsky e Halle entre a fonética e a fonologia enquanto níveis de representação tinha precedente na prática de Sapir de considerar os alofones parte do padrão fônico – ou sônico, para usar a tradução de Mattoso Câmara (Sapir 1969) – da língua em estudo. Eis, por exemplo, como Sapir resolve em dois níveis –

¹ Texto da conferência proferida, à guisa de prova de aula, no concurso para provimento de cargo de Professor Titular na área de Fonética e Fonologia do Departamento de Lingüística do Instituto de Estudos da Linguagem da UNICAMP, realizado a 11 e 12 de novembro de 1999. O ponto sorteado foi: “Representações Fonética e Fonológica”. Agradeço as falas estimulantes dos membros da banca: Rodolfo Ilari (IEL-Unicamp), Roberto Romano (IFCH-Unicamp), Leonor Scliar-Cabral (UFSC), Cesar Ades (Instituto de Psicologia da USP) e Maria Raquel Delgado Martins (Universidade de Lisboa).

bastante distanciados entre si – a “ortografia” do Paiute do Sul, discutida a propósito da necessidade reconhecer o fonema como unidade abstrata, representativa das intuições do falante:

Ortografia Fonética	Ortografia Fonológica
1. pa.	pa.
2. paβa.	papa.
3. paəA.	papa.

(Sapir, E. “A realidade psicológica dos fonemas”. In: M. Dascal, org. *Fundamentos Metodológicos da Lingüística*, vol. II, *Fonologia e Sintaxe*. Campinas: Edições do Autor, 1981, p. 42)

Uma tal ambigüidade das representações fônicas ganha também respaldo, ainda que menos veemente, na definição de Bloomfield de “forma fonética”:

“Any combination of phonemes that occurs in a language, is pronounceable in this language, is a phonetic form.”

(Bloomfield, L. *Language*. New York: Holt, Rinehart & Winston, 1933, p. 138)

Em resumo, a Fonologia Gerativa formalizou e estendeu a prática do estruturalismo americano de descrever o padrão fônico de uma língua como um conjunto de relações entre fonemas e alofones. Nessa perspectiva, tanto os fonemas como os alofones são entidades discretas, seqüenciais e podem ser representados por uma notação aproximadamente alfabética, que serve também para assinalar as suas correspondências uns com os outros. O que o modelo gerativo traz de novo, além do refinamento da teoria dos traços distintivos, é afirmar que os alofones, como os fonemas, fazem parte das intuições lingüísticas do falante nativo, e devem, portanto, ser tratados não como epifenômenos mas como fenômenos, fatos em si próprios.

Vou, aqui, endossar a visão gerativa de que os alofones são fatuais, isto é, podem ser descritos explicitamente por um modelo científico rigoroso. Vou endossar, ainda, a reivindicação de que esse modelo é de natureza computacional.

Não obstante, vou opor-me à visão, pressuposta pelo gerativismo, de que todas as computações envolvidas na alofonia têm um caráter simbólico. Para isso vou mostrar que muitas, embora nem todas, dentre essas computações têm um caráter numérico.

Essa demonstração abre caminho à tese de que apenas a informação de caráter simbólico qualifica-se como representação lingüística, no sentido forte de **esquema cognitivo permanente**, armazenado na memória a longo prazo.

Trata-se de praticar a parcimônia excluindo a possibilidade de computações numéricas *off-line* e, portanto, proibindo a existência de representações fonéticas detalhadas permanentes, salvo em casos excepcionais (V. Albano no prelo, cap. 7). Isso requer, por outro lado, que as computações numéricas *on line* sejam facilitadas.

Eis, em suma, o que vou defender: (1) a memória lingüística só retém algo que, embora mais complexo, é, ao menos quanto ao grau de abstração, aparentado ao velho fonema; (2) a variabilidade dos alofones, quando contínua, é calculada *on line* sobre uma representação fonológica permanente, simbólica, que, não obstante, é

comensurável com uma representação motora temporária gerada por um modelo fonético dinâmico; e (3) essa **representação fônica única** vale tanto para a produção como para a percepção da fala.

2. UM CASO EXEMPLAR: A ALOFONIA DA SIBILANTE DE CODA NO DIALETO CARIOCA

Para mostrar como a minha posição resgata idéias do estruturalismo quanto à concretude da representação lexical sem abdicar da tese gerativa da computabilidade dos alternantes fônicos, vou usar um exemplo célebre de Joaquim Mattoso Câmara Jr. (1969, p. 29; 1976, p. 70), que envolve as expressões “pá sólida”, “paz sólida”, “paz armada” e similares. O foco recairá sobre o dialeto carioca, não por ser o meu², mas por ser o que mais problemas coloca a uma descrição minuciosa da alofonia da sibilante de coda.

Eis algumas transcrições fonéticas, razoavelmente consensuais, das expressões em questão na fala carioca:

“pá sólida”	pa ¹ sólidɐ
“paz sólida”	pa ¹ sólidɐ, pa ¹ s:ólidɐ
“paz armada”	pa ¹ zafi ¹ madɐ
“paz pura”	pa ¹ f ¹ pure
“paz bela”	pa ¹ ʒ ¹ belɐ

Todos conhecem a solução mattosiana (Câmara 1977, p. 80-81) para o problema, evidente na segunda coluna, da variabilidade da sibilante de coda, a saber: a postulação do arquifonema /S/, inespecífico não só quanto ao ponto de articulação, que pode alternar entre palatal e dental, mas também quanto ao vozeamento, que pode alternar entre surda e sonora, dando o leque de alofones [s, z, ʃ, ʒ].

Todos sabem também que a solução gerativa consiste em resolver essa variabilidade em favor de uma forma básica, que, em algumas análises (p. ex., Leite 1974), é a sibilante surda /s/ e, em outras, mais preocupadas com a economia do processo de vozeamento, é a sua contraparte sonora /z/ (p. ex., Lopez 1979).

Ambos os tipos de solução têm méritos e exprimem o essencial do pensamento fonológico, que é captar a invariância na variabilidade. Ambos padecem, entretanto, da mesma superficialidade quanto à descrição dos fatos fonéticos. Além de relegar a inserção do [j] na palavra “paz” a um mero detalhe fonético, negligenciam outros detalhes, mais resistentes à observação de outiva, em que a sibilante de coda tem um aspecto ainda mais transicional, se examinada instrumentalmente.

Vou mostrar abaixo, através de espectrogramas, que essa sibilante de coda do carioca pode apresentar mais de um grau de vozeamento, mais de um grau de

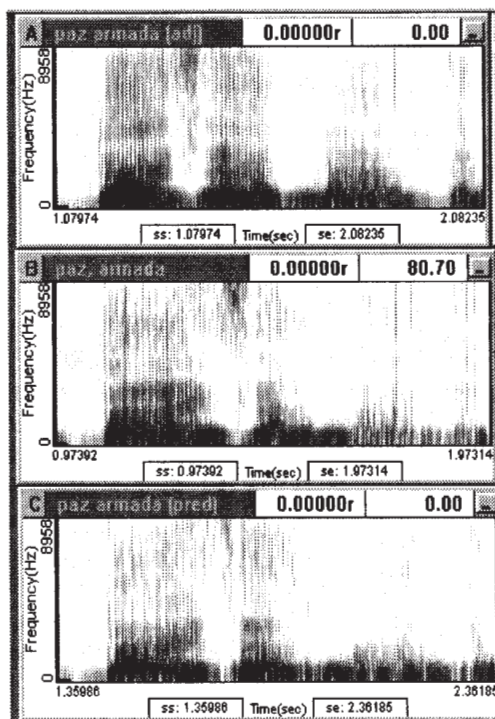
² O prazo de 24h para o preparo desta conferência me fez tomar a mim mesma como informante. Mantenho, aqui, as ilustrações originais, já corroboradas pela inspeção de dados de outros cariocas.

palatalidade, mais de um grau de iotização da vogal adjacente – tudo isso em múltiplas combinações e em resposta sistemática à estrutura prosódica do enunciado, que está correlacionada à sua estrutura gramatical. Daí decorre que essa variabilidade **contínua** não é, como queriam os estruturalistas, aleatória, assistemática, alheia ao padrão fônico. Decorre também que ela também não é, como querem até hoje certos gerativistas, expressável em termos simbólicos por uma simples conversão dos traços distintivos binários em gradações discretas ou escalas.

3. ALGUNS EXEMPLOS DE GRAUS DE PALATALIDADE, VOZEAMENTO E IOTIZAÇÃO NA SIBILANTE CARIOCA

Os espectrogramas abaixo fazem revelações surpreendentes sobre a inacuidade da observação de outiva dos fatos fonéticos concernentes à alofonia da coda sibilante no dialeto carioca. Neles, a expressão “paz armada” foi analisada em três diferentes contextos prosódico/gramaticais, representados pelas seguintes frases:

- 1) Que tipo de paz você cultiva? _Cultivo a paz armada.
- 2) Eu disse aos soldados: _ Cultivo a paz, armada.
- 3) Como você cultiva a paz, amiga? _ Cultivo a paz armada.



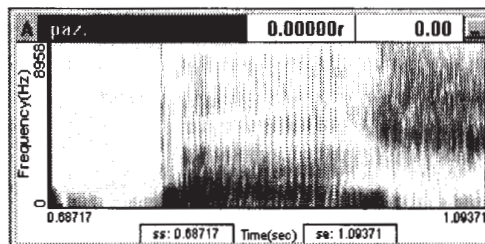
Observemos a janela (a) da figura acima. O [a] é o menos iotizado e o [z] o mais vozeado dos três. Mas, surpreendentemente, tem energia acústica abaixo de 4kHz, o que é um sinal de palatalidade. O que, então, o faz soar como [z] e não como [ʒ]? Provavelmente, a baixa intensidade do ruído fricativo nessa faixa, que contrasta com a alta intensidade de F₂ e F₃ da vogal seguinte, localizados na mesma faixa de frequência. Há aqui, segundo Stevens (1985), uma fronteira auditiva quântica, isto é, uma resposta auditiva que faz a percepção do ponto de articulação da fricativa saltar para um outro patamar, devido ao mascaramento do ruído pelos formantes, mais intensos, da vogal.

Como transcrever: [z] ou [ʒ]? Esse não é o ponto. O ponto é que essa fricativa é ambígua e distinta de outros [z, ʒ]. Uma tal ambigüidade não é certamente perceptível conscientemente, para fins da atividade metalingüística de transcrição. Mas é perceptível inconscientemente, na medida em que um falante desse dialeto *aprende-a e pratica-a* sistematicamente.

Queria isso dizer que a perda da palatalização da fricativa de coda é um fenômeno perceptivo e que as suas representações fonológica e fonética devem resumir-se, de fato, a uma palatal?

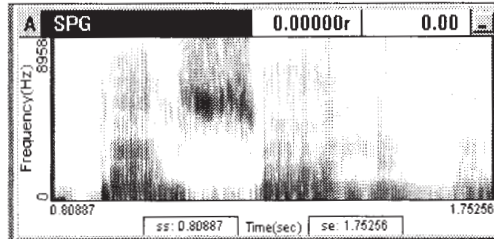
Ledo engano. Observemos agora os espectrogramas das janelas (b) e (c). A energia acústica do ruído fricativo está, em ambos os casos, inequivocamente acima de 4KHz. Também em ambos os casos, o vozeamento, mesmo não sendo contínuo, tem duração suficiente para ser claramente percebido – *a fortiori* porque as vogais precedentes são bastante alongadas em comparação com a da janela (a). Tudo indica tratar-se de verdadeiras ocorrências de [z], ambas ligeiramente desvozeadas, devido à presença de fronteiras prosódicas razoavelmente fortes (correspondentes ao vocativo e ao predicativo do sujeito).

O surpreendente nesses dados não é a existência de variantes fonéticas distintas, que até a simples transcrição de outiva é capaz de captar. É a distribuição inesperada das variação de ponto de articulação face à força das fronteiras prosódicas. Um ruído ambíguo entre palatal e dental ocorre em fronteira fraca, enquanto a dental, que se ouve, mas, de fato, não se pronuncia nesse ambiente, onde seria esperada, acaba ocorrendo em fronteira forte não absoluta, onde não o seria. A expectativa de encontrar algo mais próximo da palatal nas fronteiras fortes baseia-se no fato, confirmado pelo espectrograma abaixo, de que ela é a variante de posição final absoluta. Eis o ['paʒ] de "Cultivo a paz".



Nota-se aqui que a iotização do [a] é menor que nos outros casos, mas que a fricativa é decididamente surda e palatal.

Um ruído fricativo também surdo, porém mais anterior – provavelmente dental – aparece, além disso, na geminada mattosiana de “paz sólida”³. Nela o ruído fricativo se inicia na faixa de 4 KHz e não há indício de iotização de [a], conforme indica a trajetória, quase plana e bastante suave, de F2 e F3:



Esses fatos, no seu conjunto, defrontam-nos com um aparente paradoxo: por um lado, temos uma alternância francamente categórica entre um ruído de aparência dental, com energia acima de 4 KHz, e um ruído de aparência palatal, com energia abaixo de 4 KHz; por outro, temos gradientes de palatalidade, vozeamento e iotização dentro dessas categorias.

Como compatibilizar a alofonia, aparentemente discreta, do primeiro caso com a alofonia, aparentemente contínua, do segundo?

Vou defender abaixo que a solução consiste em conjugar, no léxico, uma representação simbólica comensurável com um modelo fonético dinâmico a uma forma não-derivacional de computar a alofonia discreta, isto é, uma computação que relacione as variantes em questão sem apelar para a noção de forma básica.

4. REPRESENTAÇÕES FONOLÓGICAS COMPATÍVEIS COM UM MODELO FONÉTICO DINÂMICO

Lembremos que o caráter sucessivo, pontual do fonema sobreviveu, na teoria fonológica contemporânea, não só na chamada “camada temporal” mas também nos vários traços autosegmentalizados. Uma exceção nesse quadro conceitual é a *Fonologia Articulatória* (Browman & Goldstein 1985; 1986; 1989; 1990a,b; 1992a,b), que postula uma unidade fônica lexical dotada de duração intrínseca, isto é: de um tempo, um intervalo abstrato compatível com o tempo concreto dos movimentos articulatorios, tal como implementado num modelo fonético dinâmico. A solução para a variabilidade descrita acima consiste em diferenciar os alofones da sibilante de coda do dialeto carioca diretamente na representação lexical através desse tipo de unidade, denominado gesto articulatório.

Eis como se representa a sibilante de coda final do carioca na minha própria versão do modelo, que denomino *Fonologia Acústico-Articulatória*, conforme exposição num ensaio, ainda no prelo, que submeti à banca deste concurso, sob o título “O gesto e suas

³ De fato uma geminada, como demonstra a sua maior duração com relação às das fricativas dos espectrogramas do primeiro bloco acima, que estão na mesma escala temporal.

Bordas: Esboço de Fonologia Acústico-Articulatória do Português Brasileiro” (Albano no prelo):

		Sibilante de Coda Final			
		Absoluto	_#C sonora	_#s	_#V
Região Coronal					
Grau de constrição		cr.	cr.	crítico	cr.
Local de constrição					
Região glotal					
		aberta	fechada	aberta	fechada
		[ʃ]	[ʒ]	[s]	[ʒ]

Essa representação admite componentes simultâneos, à semelhança dos traços distintivos, com uma diferença, decorrente da sua natureza motora: não são meramente pontuais. Cada componente possui um intervalo, um “tempo” característico, que pode coincidir ou não com os dos demais. Assim, o que a figura acima diz é que há movimentos simultâneos na região coronal e na região glotal e que, dependentemente da maior ou menor sincronização daqueles que deslocam a constrição ao longo da ordenada com os que a deslocam ao longo da abscissa do plano sagital do trato vocal, a coda resultante é uma palatal ou uma dental, ou ainda, uma dental com uma transição lenta, que dá uma aparência iotizada à vogal precedente. Uma constrição em que a língua move-se lentamente na região coronal até encontrar o grau de constrição crítico é palatal. Uma constrição coronal que se torna crítica imediatamente é dental. Uma constrição na região coronal que se torna crítica apenas no final iotiza a vogal precedente, devido à lentidão adicional do movimento.

Não tenho espaço aqui para explicar como essa representação difere da de Browman e Goldstein e por que pressupõe que o gesto tenha alvos não só articulatórios mas também acústicos (V. Albano no prelo, capítulos 3 e 4). O importante, porém, é que ela permite, quando interpretada por um modelo dinâmico de implementação fonética, versões gradientes dos quatro alofones acima. Dada a sua natureza temporal, cada bloco corresponde a um movimento que pode ser comprimido, expandido ou deslizar com relação aos demais, a ele alinhados. Assim, o gesto de fechamento glotal pode encolher, dando lugar a uma sonora parcialmente desvozeada. Da mesma forma, o movimento que especifica o grau de constrição pode antecipar-se ou adiar-se, diminuindo ou aumentando a iotização da vogal precedente.

Como na Fonologia Articulatória, a implementação fonética dessa representação consiste em fazer os pontos que especificam o grau e o local de constrição deslocarem-se no trato vocal de acordo com uma equação diferencial do tipo que descreve o deslocamento de uma massa atrelada a uma mola. Os detalhes de como isso se faz estão

explicados em Kelso, Saltzman & Tuller (1986) e são muito complexos para serem reproduzidos aqui.

Um modelo assim prescinde de representação fixa do detalhe fonético, discreta ou escalar, exatamente porque pode gerar representações temporárias, variáveis, contínuas do movimento dos articuladores. Toda computação fonética, numérica é realizada *on line*, mediante especificação de outras variáveis tais como estilo e taxa de elocução.

Quanto à computação fonológica, não faz sentido estabelecer correspondências entre gestos derivando-os uns dos outros, uma vez que quaisquer representações simbólicas usadas no léxico destinam-se apenas a reduzir a dimensionalidade dos sistemas dinâmicos envolvidos e devem ser passíveis de interpretação e implementação numéricas. Assim, não se pode fazer mais do que tabelar, registrar, no léxico, toda alofonia discreta, congelada, tal como a refletida, nos dados acima, na escolha do [ʒ] final para assinalar que, apesar dos índices de pausa, o enunciado continua numa vogal.

Alofones no léxico? Mesmo quando realizam o *sandhi* externo? Para quem foi educado na tradição gerativa, isso parece um retrocesso, uma volta ao estruturalismo.

A ameaça à parcimônia não é, contudo, tão grande quanto parece. Uma representação lexical compatível com um modelo fonético dinâmico economiza a representação fonética e, sobretudo, a expressão simbólica de detalhes fonéticos com função lingüística, tais como os graus de iotização vistos acima. Um modelo fonológico de traços jamais seria capaz de dar conta deles, como mostraram Albano *et al.* (1998).

Como dar conta das disjunções lexicais? Isso é matéria para perspectivas que trabalham com restrições de boa formação e não com regras, como a teoria da otimalidade (Prince & Smolensky 1993). O tema é longo demais para irmos aqui além de mencionar a urgência e a oportunidade do seu desenvolvimento.

Para terminar, vou tecer algumas considerações sobre outros aspectos do processamento lingüístico em que a representação gestual traz uma considerável economia.

5. INDÍCIOS FAVORÁVEIS À REPRESENTAÇÃO GESTUAL COMO REPRESENTAÇÃO FÔNICA ÚNICA

Há, na literatura, uma controvérsia entre autores como John Kingston e Randy Diehl (1994, 1995), que julgam que as representações envolvidas na percepção de fala são acústico-auditivas, e autores como Carol Fowler (1996) e Douglas Whalen (1999), do famoso Haskins Laboratories, berço da Teoria Motora da percepção da fala (Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967), que julgam que elas são motoras.

O que se discutiu acima sugere fortemente que ambos os lados estão equivocados, na medida em que pressupõem o armazenamento de representações fonéticas detalhadas, motoras ou auditivas, das palavras ou dos componentes sublexicais envolvidos.

Uma representação lexical gestual mínima como a acima proposta pode funcionar na percepção como um mecanismo de análise por síntese porque contém a informação lingüística necessária para a geração *on line* de movimentos articulatórios, abertos ou

encobertos, cujos efeitos acústicos podem ser também avaliados aberta ou encobertamente.

Essa representação permitiria uma interpretação percepto-motora (i. e., pautada por critérios acústico-articulatórios e, portanto, mais coerente com o processo de produção) do chamado modelo LAFS (*Lexical Access From Spectra*), de Dennis Klatt (1979). É compatível com efeitos perceptuais conhecidos, tais como a restauração fonêmica (Warren 1976), assim como com outros, mais controversos e menos consistentes, tais como a compensação para a coarticulação, extensamente estudada pelo grupo de Haskins (V., p. ex., Fowler, Brown & Mann 1999).

Por que os efeitos coarticulatórios de uma consoante sobre uma vogal seriam “descontados” em alguns casos, mas não em outros (Kingston 1999), em testes de identificação de fones?

O que se disse acima permite prever que os resultados desses testes dependem da maior ou menor facilidade e confiabilidade do acesso lexical. Palavras muito freqüentes e que disputam esse acesso com poucos candidatos devem ser reconhecidas sem recurso a uma “fala encoberta” gerada pelo mecanismo de produção. Suas características coarticulatórias devem, portanto, ser pouco perceptíveis, pois o sinal acústico retido temporariamente na memória auditiva é posto em correspondência direta com a representação gestual. Palavras mais raras e pertencentes a coortes maiores (Marslen-Wilson 1989) devem dar lugar a uma computação fonética interna em que a coarticulação torna-se acessível à percepção.

O fato de a maioria dos experimentos de percepção de fala utilizar logatomas e não palavras não altera essa previsão. Existem dados na literatura mostrando que logatomas constituídos de componentes sublexicais bem formados e freqüentes despertam reações de familiaridade, enquanto aqueles compostos de componentes sublexicais mal formados e infreqüentes despertam reações de estranhamento (Coleman & Pierrehumbert 1997).

Infelizmente, ainda não tenho dados que me permitam corroborar a hipótese de que a percepção da coarticulação no português brasileiro é sensível a variáveis lexicais. Ela se coloca aqui, à guisa de conclusão, como um novo passo possível para respaldar a hipótese, atraente porque parcimoniosa, de que uma só representação gestual abstrata serve tanto à percepção como à produção da fala.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBANO, E. C. no prelo. *O gesto e suas Bordas: Esboço de Fonologia Acústico-Articulatória do Português Brasileiro*. Campinas: Editora Mercado de Letras.
- _____; P. Barbosa.; S. Madureira; A. Gama-Rossi; A. Silva. 1998. A interface fonética-fonologia e a interação prosódia-segmentos. *Estudos Lingüísticos XXVII: Anais do XLV Seminário do Grupo de Estudos Lingüísticos do Estado de São Paulo (GEL)*. São José do Rio Preto: UNESP-IBILCE, 135-143.
- BROWMAN, C. & L. Goldstein. 1985. Dynamic modeling of phonetic structure. In V. Fromkin (org.). *Phonetic Linguistics*. Nova Iorque: Academic Press, 35-53.
- _____. 1986. Towards an articulatory phonology. *Phonology Yearbook 3*: 219-252.
- _____. 1989. Articulatory gestures as phonological units. *Phonology Yearbook 6*: 201-251.

- _____. 1990a. Tiers in articulatory phonology. In: J. Kingston & M. Beckman (orgs.) *Papers in laboratory phonology: between the grammar and the physics of speech*. Cambridge: Cambridge University Press, pp. 341-376.
- _____. 1990b. Gestural specification of dynamically determined articulatory structures. *Journal of Phonetics*, 18: 299-320.
- _____. 1992a. Articulatory phonology: an overview. *Phonetica*, 49: 155-180.
- _____. 1992b. "Targetless" schwa: an articulatory analysis. In: G. Docherty e R. Ladd (orgs.) *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press, pp. 26-56.
- CÂMARA Jr., J. M. 1969. *Problemas de Lingüística descritiva*. Petrópolis: Vozes.
- _____. 1976. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão.
- _____. 1977. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Padrão.
- CHOMSKY, N. & M. Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- COLEMAN, J. & J. Pierrehumbert. 1997. Stochastic phonological grammars and acceptability. Comunicação ao III Meeting of the Association for Computational Linguistics, Somerset, New Jersey, julho de 1997.
- FOWLER, C. 1996. Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*, 99: 1730-1741.
- _____; J. Brown & V. Mann. 1999. Compensation for coarticulation in audiovisual speech perception. *Proceedings ICPhS '99*, vol. 1, 639-642.
- KELSO, J.; E. Saltzman; B. Tuller. 1986. The dynamical perspective on speech production: data and theory. *Journal of Phonetics*, 14: 29-59.
- KINGSTON, J. 1999. Failing to parse Hz in perceiving [voice] and [continuant]. *Proceedings ICPhS '99*, vol. 2, 869-872.
- _____. & R. Diehl. 1994. Phonetic knowlege. *Language*, 70: 419-474.
- _____. 1995. Intermediate properties in the perception of distinctive feature values. In: B. Connell & A. Arvaniti. (orgs.). *Phonology and phonetic evidence: papers in laboratory phonology IV*. Cambridge: Cambridge University Press, pp. 7-27.
- KLATT, D. 1979. Speech perception: a model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7: 279-312.
- LEITE, Y. 1974. Portuguese stress and related rules. Tese de doutorado inédita: University of Texas at Austin.
- LIBERMAN, A.; F. Cooper; D. Shankweiler; M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review*, 74: 431-461.
- LOPEZ, B. 1979. The sound pattern of Brazilian Portuguese (Cariocan dialect). Tese de doutorado inédita: University of California, Los Angeles.
- MARSLÉN-WILSON, W. 1989. Access and integration: projecting sound onto meaning. In: W. Marslen-Wilson (org.) *Lexical representation and process*. Cambridge: MIT Press, 1989, pp. 6-24.
- PRINCE, A. e P. Smolensky. 1993. Optimality theory: constraint interaction in generative grammar. Manuscrito inédito: Rutgers University e University of Colorado.
- SAPIR, E. 1969 [1925]. Os padrões sônicos na linguagem. In: ____ *Lingüística como ciência*, tradução de Joaquim Mattoso Câmara Jr. Rio de Janeiro: Acadêmica.
- _____. 1981 [1933]. A realidade psicológica dos fonemas. In: M. Dascal. org. *Fundamentos metodológicos da Lingüística*, vol. II, *Fonologia e Sintaxe*. Campinas: Edições do Autor.

- STEVENS, K. 1985. Evidence of the role of acoustic boundaries in the perception of speech sounds. In: V. Fromkin (org.). *Phonetic Linguistics: essays in honor of Peter Ladefoged*. Nova lorque: Academic Press, pp. 243-255.
- WARREN, R. 1976. Auditory illusions and perceptual processes. In: N. Lass (org.) *Contemporary issues in experimental phonetics*. Nova lorque: Academic, pp. 389-417.
- WHALEN, D. 1999. Three lines of evidence for direct links between production and perception. *Proceedings ICPhS '99*, vol. 2, 1257-1260.