



BiPhon-OT: Modelo Bidirecional de Fonologia e Fonética

<https://doi.org/10.47627/gradus.v7i1.178>

Autores

**Carmen Lúcia Barreto
Matzenauer**

<https://orcid.org/0000-0003-4505-7521>

Universidade Federal de Pelotas

**Roberta Quintanilha
Azevedo**

<https://orcid.org/0000-0001-6076-9768>

Instituto Federal Sul Rio-Grandense

Abstract

The focus of this paper is the characterization of BiPhon-OT, the Phonology and Phonetics model proposed by Paul Boersma. It consists of a grammar model, which is supported by Optimality Theory, and a processing model, which brings together linguistic comprehension and production, expressed by two modules in its architecture, giving the model the particularity of being bidirectional. To account for phonetic-phonological phenomena of linguistic data, it adopts three levels of representation: a phonetic level, expressed by [Phonetic Form], and two phonological levels, expressed by /Surface Form/ and by |Underlying Form|. Associated with Stochastic OT and, therefore, with a learning algorithm (Gradual Learning Algorithm – GLA), BiPhon articulates these representational levels through three types of constraints: Faithfulness, Structure and Cue ones. It aims to achieve explanatory adequacy through language simulations, thus accommodating gradient effects and probabilistic generalizations. Aiming to exemplify the BiPhon-OT functioning, this paper presents a theoretical exercise of analysis and formalization of the vowel epenthesis phenomenon in Brazilian Portuguese (BP), with simulations of this phonetic-phonological process in the language system, taking into account the Comprehension and Production modules, with data borrowed from Quintanilha-Azevedo. This theoretical exercise shows the formalization, in the BiPhon-OT model, of the epenthesis phenomenon in the heterosyllabic sequence 'p.t' in word-mid codas in BP.

Keywords: BiPhon-OT; Phonology and Phonetics; Representational Levels; Linguistic Comprehension and Production; Faithfulness, Structure and Cue Constraints.

Histórico do artigo

Recebido 2022-03-10

Aceito 2022-07-13

Publicado 2022-08-19

Resumo

O foco deste artigo está na caracterização do BiPhon-OT, modelo de Fonologia e Fonética proposto por Paul Boersma. Constitui-se em um modelo de gramática, que adota o suporte da Teoria da Otimalidade, e em um modelo de processamento, que congrega a compreensão e a produção linguísticas, expressas por dois módulos na sua arquitetura, o que lhe atribui a particularidade de ser bidirecional. Para dar conta de fenômenos fonético-fonológicos de dados linguísticos, adota três níveis de representação: um nível de natureza fonética, expresso pela [Forma Fonética], e dois níveis de natureza fonológica, expressos pela /Forma Fonológica de Superfície/ e pela |Forma Fonológica Subjacente|. Associado à OT Estocástica e, por conseguinte, a um algoritmo de aprendizagem (*Gradual Learning Algorithm* – GLA), o BiPhon articula esses níveis representacionais por meio de três tipos de restrições: Restrições de Fidelidade, Restrições de Estrutura e Restrições de Pista, e visa alcançar a adequação explicativa por meio de simulações da linguagem, acomodando, assim, efeitos gradientes e generalizações probabilísticas. Visando à exemplificação do funcionamento do BiPhon-OT, este artigo traz um exercício teórico de análise e formalização do fenômeno da epêntese vocálica, no Português Brasileiro (PB), com simulações desse processo fonético-fonológico no sistema da língua, levando em conta os módulos da Compreensão e da Produção, com dados emprestados de Quintanilha-Azevedo. O exercício teórico traz a formalização, no modelo BiPhon-OT, do fenômeno da epêntese na sequência heterossilábica 'p.t' em coda medial no PB.

Palavras-chave: BiPhon-OT; Fonologia e Fonética; Níveis Representacionais; Compreensão e Produção Linguísticas; Restrições de Fidelidade, de Estrutura e de Pista.

Introdução

Este artigo¹ visa a apresentar uma síntese do modelo intitulado *Parallel Bidirectional Phonology and Phonetics* (BiPhon), proposto por Boersma,² que se constitui em um modelo de gramática, seguindo os pressupostos da Teoria da Otimidade (*Optimality Theory* – OT), cujos fundamentos vêm de Prince e Smolensky.³ Boersma explica que esse modelo de gramática “pretende ser capaz de lidar com ‘toda’ a fonologia” e com os fenômenos a ela relacionados, desde as representações, os processos, a conexão com a semântica, a aquisição pela criança, a evolução ao longo das gerações e a tipologia de diferentes línguas, buscando atingir a adequação explicativa por meio de simulações da língua.⁴

No tocante às propriedades do BiPhon como um modelo de gramática, particularmente de gramática fonológica, é relevante destacar-se que reúne Fonética e Fonologia ao considerar que categorias fonológicas são representações constituídas, pelo falante/ouvinte (ou aprendiz de L1 e/ou de L2), a partir de pistas fonéticas. Também reúne compreensão e produção linguísticas, o que lhe acrescenta a qualidade de um modelo de processamento, bem como lhe atribui a característica de bidirecionalidade: é por representar a tarefa do indivíduo/aprendiz na condição de ouvinte e também na condição de falante que se configura como um modelo de processamento bidirecional.

O BiPhon expressa a noção de **representação** em diferentes níveis, conforme mostra a fig. 1; contém representações semânticas, representações fonológicas, representações fonéticas, que estão conectadas de forma interativa. Tratando de níveis de representação e também do armazenamento do conhecimento em um modelo que integra Fonologia e Fonética, Boersma, Benders e Seinhorst⁵ atribuem destaque ao nível fonológico e ao nível fonético, como se verifica na fig. 4.

A relação entre os níveis é articulada por meio de restrições nos moldes da OT; são as restrições o mecanismo de decisão escolhido pelo modelo. Na fig. 1, vêem-se referidas restrições articulatórias, restrições sensorio-motoras, restrições de pista, restrições estruturais, restrições de fidelidade, restrições lexicais e restrições semânticas. Embora em Boersma, Benders e Seinhorst⁶ tenha havido a introdução de um modelo de rede neural artificial para dar conta de fenômenos fonético-fonológicos de dados linguísticos (BiPhon-NN – *BiPhon neural-network*), o presente artigo atém-se ao BiPhon que utiliza o mecanismo da OT, também denominado pelo autor de BiPhon-OT.⁷

¹ O presente estudo sobre o BiPhon integra projeto apoiado pelo CNPq – Processo n° 306616/2018-1.

² BOERSMA, “Functional Phonology”: Formalizing the interactions between articulatory and perceptual drives (1998); BOERSMA, “Some listener-oriented accounts of h-aspiré in French” (2007); BOERSMA, “Cue constraints and their interactions in phonological perception and production” (2009); BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution” (2011).

³ PRINCE e SMOLENSKY, *Optimality Theory: Constraint Interaction in Generative Grammar* (2004 [1983]).

⁴ BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution” (2011).

⁵ BOERSMA et al., “Neural network models for phonology and phonetics” (2020).

⁶ BOERSMA et al., “Neural network models for phonology and phonetics” (2020).

⁷ Embora o BiPhon possa utilizar o modelo de avaliação do candidato ótimo nos moldes da Gramática Harmônica, no qual todas as restrições exercem papel na escolha do *output*, independentemente do seu peso, aqui será dado o tratamento às restrições segundo a OT clássica, com a propriedade estocástica, isto é, seguindo a noção de dominância estrita; as restrições com menores pesos/valores só se fazem relevantes nos casos em que as restrições mais altas na hierarquia, ou com maiores pesos/valores, não se mostram capazes de definir o candidato.

Seguindo Smolensky,⁸ Boersma⁹ afirma que as restrições são usadas bidirecionalmente, o que implica que o usuário de uma língua usa as mesmas restrições, organizadas no mesmo *ranking*, tanto na condição de ouvinte como na condição de falante, ou seja, os módulos da Compreensão e da Produção empregam os mesmos elementos gramaticais.

De modo a mostrar a capacidade explicativa do modelo, neste artigo será apresentado um exercício teórico sobre a epêntese vocálica, no português do Brasil (PB), que simulará o processo fonético-fonológico do fenômeno no sistema da língua, levando em conta o módulo da Compreensão e o da Produção. Para essa exposição, o artigo está dividido em duas grandes partes: caracterização do modelo BiPhon-OT e análise da epêntese vocálica no português, cujo suporte está no BiPhon-OT. Enquanto a primeira parte traz uma introdução ao modelo aqui apresentado e um resumo da arquitetura do BiPhon-OT, explicitando propriedades e conceitos que o caracterizam, a segunda parte reporta o estudo de Quintanilha-Azevedo:¹⁰ descreve o fenômeno da epêntese vocálica após segmentos plosivos no PB, as restrições pertinentes para dar conta do fenômeno e os dados e restrições em análise no exercício teórico, bem como apresenta a formalização do fenômeno no modelo BiPhon. As considerações finais estão expressas na última seção.

⁸ SMOLENSKY, “On the comprehension/production dilemma in child language” (1996).

⁹ BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution” (2011).

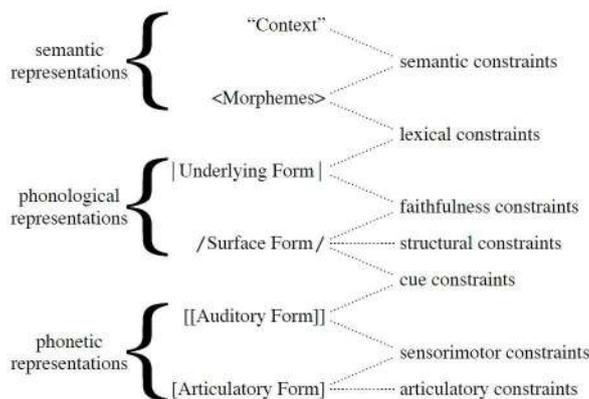
¹⁰ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

A arquitetura do BiPhon-OT

Para chegar-se à arquitetura do modelo, preliminarmente destaca-se que o BiPhon-OT é um modelo que faz a interface Fonética/Fonologia (é um modelo de Fonética e Fonologia), que assume diferentes níveis de representação e que inclui as propriedades de compreensão e de produção linguísticas. É um modelo de gramática e de processamento, desenvolvido a partir da corrente gerativa da Fonologia, que está associado à OT Estocástica e, por conseguinte, a um algoritmo de aprendizagem (*Gradual Learning Algorithm* – GLA). Essas características do BiPhon-OT são detalhadas a seguir, na apresentação da arquitetura do modelo.

A arquitetura do BiPhon em Boersma (2011)

A arquitetura do BiPhon-OT é apresentada por Boersma¹¹ com a estrutura retratada na fig. 1, contendo duas representações fonológicas, conectadas uma à outra e também conectadas a duas representações semânticas e a duas representações fonéticas.



As representações incluem dois níveis: as Representações Fonéticas incluem a [Forma Articulatória] e a [Forma Auditiva], as Representações Fonológicas incluem a /Forma Fonológica de Superfície/ e a |Forma Fonológica Subjacente|, e as Representações Semânticas incluem os < Morfemas > e o “Contexto”. O BiPhon-OT assume que as representações fonológicas são discretas e que as representações fonéticas são contínuas.

Boersma, Benders e Seinhorst¹² registram que os cinco níveis (desde < Morfemas > até a [Forma Articulatória]), formalizados na fig. 1, representam uma combinação simplificada do que fonólogos como, por exemplo, Trubetzkoy¹³ e Kiparsky¹⁴ propuseram em modelos de produção fonológica, e também do que psicolinguistas como, por exemplo, McClelland e Elman¹⁵ e Cutler *et al.*¹⁶ propuseram em modelos de compreensão, ou Levelt *et al.*¹⁷ propuseram em modelo de produção.

¹¹ BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution” (2011).

Figura 1: Modelo de Gramática (BiPhon). Fonte: BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution”, p. 2.

¹² BOERSMA et al., “Neural network models for phonology and phonetics” (2020).

¹³ TRUBETZKOY, *Principles of Phonology* (1969 [1939]).

¹⁴ KIPARSKY, “From cyclic phonology to lexical phonology” (1982).

¹⁵ MCCLELLAND e ELMAN, “The TRACE model of speech perception” (1986).

¹⁶ CUTLER et al., “Phoneme identification and the lexicon” (1987).

¹⁷ LEVELT et al., “A theory of lexical access in speech production” (1999).

A par de ser um modelo de gramática, o BiPhon também é um modelo de processamento bidirecional, conforme já foi referido. Em assim sendo, a observação da estrutura mostrada na fig. 1 leva ao entendimento de que a tarefa do usuário/aprendiz da língua na condição de ouvinte implica um movimento ascendente, partindo da [Forma Auditiva] (o som) e terminando com uma mudança no “Contexto”, enquanto a tarefa na condição de falante ocorre partir de uma mudança pretendida no “Contexto”, terminando com uma [Forma Articulatória] (a pronúncia implementada pelos órgãos da fala). Nesse processo, os diferentes níveis de representação são articulados por meio de restrições em consonância com a OT.

A arquitetura simplificada do BiPhon em Boersma e Hamann (2009)

Uma estrutura do BiPhon apresentada por Boersma e Hamann,¹⁸ mostrada na fig. 2, explicita com clareza os movimentos efetuados pelo usuário/aprendiz de uma língua nas tarefas de compreensão e de produção linguísticas. Os autores explicam que é uma versão simplificada para expressar o processamento de uma primeira língua (L1).

¹⁸ BOERSMA e HAMANN, “Loanword adaptation as first-language phonological perception” (2009).

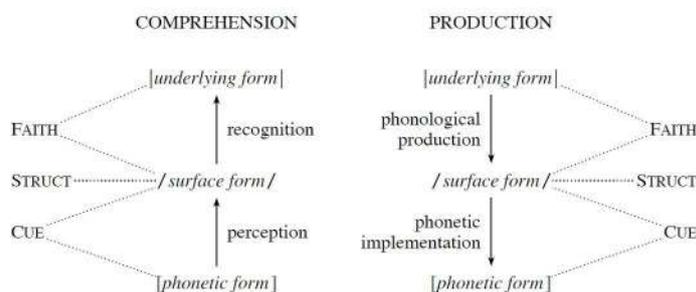


Figura 2: Modelo de Gramática e de Processamento de L1 (BiPhon). Fonte: BOERSMA e HAMANN, “Loanword adaptation as first-language phonological perception”, p. 1.

Ao se compararem as fig. 1 e 2, vê-se que, na fig. 2, não estão presentes as Representações Semânticas e que os dois níveis das Representações Fonéticas estão colapsados na chamada [Forma Fonética]. Assim, visando a expor o vínculo entre Fonética e Fonologia, a fig. 2 apresenta três níveis de representações – um nível de natureza fonética, expresso pela [Forma Fonética], e dois níveis de natureza fonológica, expressos pela /Forma Fonológica de Superfície/ e pela |Forma Fonológica Subjacente|; os níveis estão distribuídos em dois módulos: o da Compreensão e o da Produção linguísticas.

As características de modelo de gramática e de processamento podem ser reconhecidas na fig. 2: o processamento está formalizado por meio dos mecanismos ligados pelas flechas, enquanto a gramática tem a formalização expressa por meio das restrições.

Representando a tarefa do usuário/aprendiz da língua na condição de ouvinte, o Módulo da Compreensão, que implica o movimento ascendente na figura, a partir da [Forma Fonética], ex-

pressa que o ouvinte converte uma dada [Forma Auditiva] ([Forma Fonética] na fig. 2) em uma /Forma Fonológica de Superfície/, que é usada para encontrar uma |Forma Fonológica Subjacente| no léxico (*word recognition*). À direita da fig. 2, está representada a tarefa do usuário/aprendiz da língua na condição de falante por meio do Módulo da Produção, em um movimento descendente na figura, a partir da |Forma Fonológica Subjacente| até chegar a uma determinada [Forma Articulatória] ([Forma Fonética] na fig. 2).

A conexão entre os três níveis de representação evidencia a interface Fonética/Fonologia no processo de compreensão e de produção linguísticas. Das duas representações fonológicas, a |Forma Fonológica Subjacente| (*Underlying Form* – UF) constitui-se em uma sequência de estruturas fonológicas discretas associadas a morfemas no léxico, enquanto a /Forma Fonológica de Superfície/ (*Surface Form* – SF) se constitui em uma estrutura fonológica que consiste em elementos fonológicos abstratos, como traços, segmentos, sílabas e pés. A [Forma Fonética] (*Phonetic Form* – PF), na fig. 2, representa uma [Forma Auditiva] (*pitches*, formantes, silêncios e ruídos) nos módulos da Compreensão e da Produção, e representa uma [Forma Articulatória] (conjunto de comandos musculares: língua, lábios, velum, faringe, laringe e pulmões) no módulo da Produção.

No processo de comunicação, estando na condição de ouvinte, o usuário/aprendiz da língua opera no Módulo da Compreensão: capta a [Forma Auditiva], que é um fluxo contínuo de sons, e abstrai a informação acústica relevante mapeando-a em unidades fonológicas discretas, ou seja, mapeando as categorias fonológicas (esta é a chamada **percepção** na fig. 2); ocorre o chamado **reconhecimento lexical** na fig. 2, conforme explica Zhou,¹⁹ quando o ouvinte acessa o significado lexical pretendido ao combinar a forma percebida (/Forma Fonológica de Superfície/) com a representação fonológica armazenada na memória de longo prazo (|Forma Fonológica Subjacente|), que está vinculada à representação semântica. E essas duas etapas podem ocorrer não apenas sequencialmente, mas também de forma paralela, já que se evidencia efeito lexical na percepção.

Estando na condição de falante, o usuário/aprendiz da língua opera no Módulo da Produção. Zhou²⁰ detalha o que ocorre neste módulo explicando que, primeiramente, o falante recupera a |Forma Fonológica Subjacente| de uma entrada lexical que busca na memória de longo prazo e a mapeia em uma /Forma Fonológica de Superfície/ prosódica e segmentalmente detalhada, ou seja, em uma representação fonológica totalmente especificada. Esta representação fonológica é, então, convertida em uma forma-alvo auditiva que o falante pretende alcançar (que é a [Forma Auditiva]); esta forma, por sua vez, é transformada, pelo conhecimento sensorio-motor, em uma representação articulatório-motora (que é a [Forma Articulatória]).

¹⁹ ZHOU, “L2 speech learning of European Portuguese /l/ and /r/ by L1-Mandarin learners”: experimental evidence and theoretical modelling (2021), p. 14.

²⁰ ZHOU, “L2 speech learning of European Portuguese /l/ and /r/ by L1-Mandarin learners”: experimental evidence and theoretical modelling (2021), p. 15.

No BiPhon-OT, os níveis de representação estão articulados por meio de restrições,²¹ que, na fig. 2, são de três tipos: Restrições de Fidelidade, Restrições de Estrutura e Restrições de Pista (*cue constraints*). As **Restrições de Fidelidade** avaliam a relação entre as duas formas fonológicas no **reconhecimento** (módulo da Compreensão) e na **produção fonológica** (módulo da Produção), sendo a restrição IDENT_σ um exemplo, como registra Boersma;²² as **Restrições de Estrutura** avaliam restrições específicas de língua sobre a estrutura percebida (módulo da Compreensão) e sobre a estrutura produzida (módulo da Produção), como, por exemplo, as restrições NOCODA ; ALLFEETLEFT ; as **Restrições de Pista** avaliam pistas específicas de língua, integradas na percepção (módulo da Compreensão; um exemplo está na restrição: $*[x]_{\text{AudF}}/y/_{\text{SF}}$) e na produção (módulo da Produção; um exemplo está na restrição: $*/x/_{\text{SF}}[y]_{\text{ArtF}}$).²³

Na gramática fonológica de um falante adulto da língua, o *ranking* de restrições que articula os níveis representacionais é o mesmo nos dois módulos: no da Compreensão e no da Produção.

Boersma e Hamann²⁴ apresentam um exemplo da interação entre **Restrições de Estrutura** e **Restrições de Pista**, operando no Módulo da Compreensão, ao referirem que japoneses percebem a forma fonética $[ebzo]_{\text{Aud}}$ como $/e.bu.zo/_{\text{SF}}$, porque a percepção como $*/eb.zo/_{\text{SF}}$ violaria a restrição de estrutura CODACOND . Essa explicação expressa que a percepção, nesse caso, é determinada pela interação entre a **Restrição de Estrutura** CODACOND e a **Restrição de Pista** $[]_{\text{Aud}}/u/_{\text{SF}}$ (que milita contra a epêntese de um $/u/_{\text{SF}}$ para o qual não haja quaisquer pistas auditivas diretas); estas restrições estão estruturadas neste *ranking*: $\text{CODACOND} \gg * []_{\text{Aud}}/u/_{\text{SF}}$, ou seja, a **Restrição de Estrutura** domina a **Restrição de Pista**. Afirmam os autores que esse tipo de interação se constitui em uma das principais razões para “considerar-se a percepção como fonológica e, portanto, para se formularem Restrições de Pista como parte da Teoria da Otimidade”.

Ao considerar-se que o processamento linguístico se dá em paralelo nos módulos de Compreensão e de Produção, tem-se que esse fato é formalizado conforme mostra a fig. 3.

A fig. 3 também representa o processo de aquisição fonológica, tanto ao tratar-se de aquisição da L1, como da aquisição de uma L2.

A arquitetura do BiPhon retomada em Boersma, Benders e Seinhorst (2020)

Ainda merece destacar-se que Boersma, Benders e Seinhorst²⁵ expõem o BiPhon-OT como um modelo de representações e de conhecimento fonético-fonológico e o formalizam conforme

²¹ No modelo, a hierarquia das restrições propostas é estabelecida a partir dos pesos/valores obtidos com a utilização do Algoritmo de Aprendizagem Gradual – GLA (Boersma e Hayes, 2001).

²² BOERSMA, “Prototypicality Judgments as Inverted Perception” (2006).

²³ A leitura da restrição de pista obedece à seguinte estrutura: não percebe $[x]$ fonético como $/y/$ fonológico e não produza $/x/$ fonológico como $[y]$ fonético; a primeira leitura é pertinente para a percepção e a segunda, para a implementação fonética.

²⁴ BOERSMA e HAMANN, *Sibilant inventories in bidirectional phonology and phonetics* (2006), p. 2.

²⁵ BOERSMA et al., “Neural network models for phonology and phonetics” (2020).

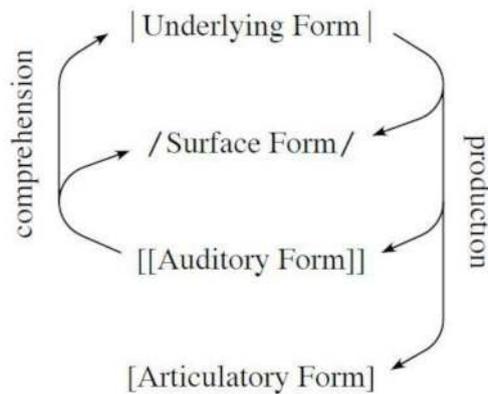


Figura 3: Processo Bidirecional de Compreensão e de Produção Fonético-Fonológica. Fonte: BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution”, p. 28.

mostra a fig. 4.

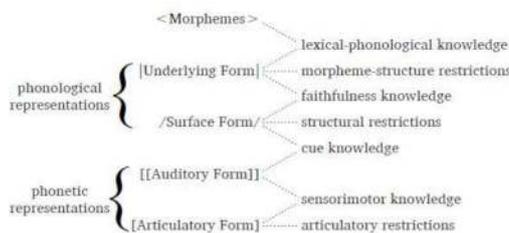
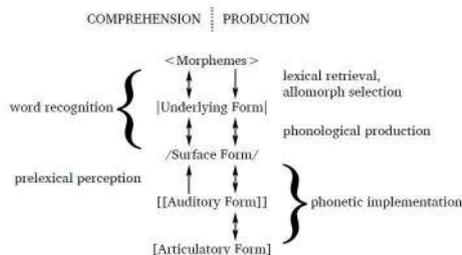


Figura 4: Níveis de representação e armazenamento do conhecimento em um modelo de Fonologia e Fonética. Fonte: BOERSMA et al., “Neural network models for phonology and phonetics”, p. 105.

Ao tecerem observações sobre a formalização do BiPhon-OT mostrada na fig. 4, os autores referem o conhecimento da gramática fonológica que os usuários/aprendizes de uma língua detêm, o que inclui o conhecimento sobre as restrições, que são organizadas em um *ranking* determinado, e também o conhecimento sobre as relações entre os níveis de representação, as quais, segundo registram, se limitam apenas aos níveis adjacentes.²⁶

Ainda apresentando considerações sobre o BiPhon-OT, Boersma, Benders e Seinhorst²⁷ expressam que um modelo abrangente deve levar em conta o comportamento do falante, do ouvinte e do aprendiz, e mostram, na fig. 5, as várias operações que ocorrem ao serem percorridos os níveis de representação presentes na fig. 4.



²⁶ Os autores referem que, em modelos de redes neurais, como o BiPhon-NN, esse conhecimento é representado como uma memória de longo prazo que consiste em pesos de conexão. BOERSMA et al., “Neural network models for phonology and phonetics” (2020), p. 106.

²⁷ BOERSMA et al., “Neural network models for phonology and phonetics” (2020), p. 106.

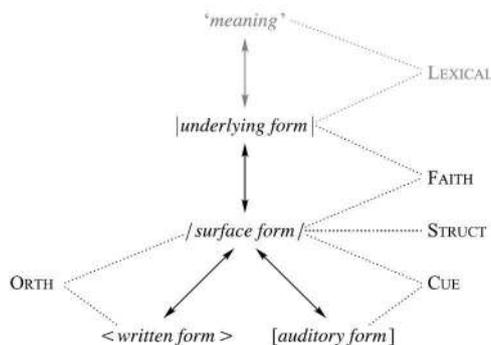
Figura 5: Operações em um modelo abrangente de Fonética e Fonologia. Fonte: BOERSMA et al., “Neural network models for phonology and phonetics”, p. 107.

As observações que fazem os autores sobre a formalização mostrada na fig. 5 são da mesma natureza daquelas já apresentadas, no presente artigo, nos parágrafos subsequentes à fig. 2, o que evidencia a semelhança entre as duas representações (fig. 2 e fig. 5) na modelagem das operações envolvidas nos processos de com-

preensão e produção linguísticas. Constatou-se que, na fig. 5, as flechas que registram o processamento entre os níveis de representações referem, por duas vezes, um mapeamento unidirecional ([Forma Auditiva] → /Forma Fonológica de Superfície/, no Módulo da Compreensão, e < Morfemas > → |Forma Fonológica Subjacente|, no Módulo da Produção), diferentemente do registro do que ocorre no mapeamento entre os outros níveis.

A arquitetura do BiPhon que inclui a influência ortográfica na aquisição – Hamann e Colombo (2017)

O entendimento de que, no processo de aquisição de uma língua, o aprendiz pode receber não apenas o *input* auditivo, mas também o *input* ortográfico por meio da leitura, conduziu à inclusão, na arquitetura do BiPhon, do nível <Forma Escrita> (ao lado de [Forma Auditiva]), como *input* para a construção das categorias fonológicas abstratas contidas na /Forma Fonológica de Superfície/, no Módulo da Compreensão – essa proposição está registrada em Hamann e Colombo.²⁸ A estrutura do BiPhon com a inclusão da <Forma Escrita> é mostrada na fig. 6.



²⁸ HAMANN e COLOMBO, "A formal account of the interaction of orthography and perception": English intervocalic consonants borrowed into Italian (2017).

Figura 6: Representação do BiPhon com o mapeamento da <Forma Escrita> em /Forma Fonológica de Superfície/ e vice-versa. Fonte: HAMANN e COLOMBO, "A formal account of the interaction of orthography and perception", p. 701.

Pela fig. 6, observa-se que o modelo passa a integrar a influência da ortografia na categorização de unidades fonológicas, prevendo um condicionamento de dupla face, uma vez que a <Forma Escrita> condiciona a /Forma Fonológica de Superfície/ e vice-versa, como também a [Forma Auditiva] condiciona a /Forma Fonológica de Superfície/ e vice-versa. Vê-se que, enquanto o mapeamento da [Forma Auditiva] na /Forma Fonológica de Superfície/ ocorre por meio de **Restrições de Pista** (CUE), o mapeamento da <Forma Escrita> na /Forma Fonológica de Superfície/ se dá por **Restrições Ortográficas** (ORTH). A previsão é que as **Restrições Ortográficas**, assim como as **Restrições de Pista**, interagem com **Restrições de Estrutura** na constituição da /Forma Fonológica de Superfície/ e com **Restrições de Fidelidade** na seleção da |Forma Fonológica Subjacente|. É relevante referir que, ao interagirem com as Restrições de Estrutura, há, de certa forma, uma competição entre Restrições Ortográficas e Restrições de Pista. Neste fato reside a explicação por que, ainda que estejam expostos ao *input*

oral da L2, muitos aprendizes “priorizam” a forma grafada (que pode levá-los a representações que se opõem àquelas encontradas no insumo linguístico).²⁹

A fig. 6 registra que a arquitetura do BiPhon modela o conhecimento da conversão grafema e fonema, expressando o reconhecimento de que pode haver efeito da ortografia nas representações fonológicas. Entende-se que o condicionamento da ortografia sobre as representações fonológicas possa ocorrer no processo de aquisição de uma L2, conforme registros de Gewehr-Borella; Zimmer e Alves³⁰ ou Cabañero e Alves,³¹ por exemplo, como também por uma reestruturação do conhecimento fonológico no processo de aquisição da escrita na L1, como defende Miranda.³² Zhou³³ descreve e formaliza, com o suporte do BiPhon-OT, a interação entre a percepção, a produção e a ortografia na aquisição da fonologia do Português Europeu como L2 por falantes nativos de Mandarim.

Em suma, o BiPhon-OT apresenta-se, pelas características aqui resumidamente expostas, como um modelo capaz de subsidiar uma descrição e uma análise de caráter fonético-fonológico bastante completa, tendo em vista que opera como um modelo de gramática e de processamento, que inclui a interação entre Fonética e Fonologia e que congrega a Compreensão e a Produção linguísticas. Um exemplo do poder explicativo que o BiPhon é capaz de oferecer relativamente ao funcionamento de fatos linguísticos é apresentado, nas seções subsequentes, por meio de uma análise e formalização do fenômeno de epêntese vocálica no português brasileiro, com dados emprestados de Quintanilha-Azevedo³⁴.

Subsídios teóricos à formalização do fenômeno da epêntese vocálica no PB

Esta seção objetiva apresentar os tópicos pertinentes para a compreensão do exercício teórico que será apresentado no modelo BiPhon-OT. Desta forma, considera-se relevante, primeiramente, situar o tema, apresentando o fenômeno da epêntese no PB. Na sequência, trazem-se os dados que serão utilizados para a análise e as restrições relevantes no processo para, ao final, mostrar-se parte da aplicação do modelo a partir do Algoritmo de Aprendizagem Gradual, já que a clareza de seu funcionamento é fundamental no entendimento do exercício teórico final.

A epêntese vocálica no Português Brasileiro

De forma sucinta, pode-se dizer que a epêntese vocálica refere um fenômeno no qual uma vogal é inserida em uma palavra com a função de reparar um *input* que não se enquadra nos requisitos

²⁹ Neste tópico, as autoras agradecem aos revisores pela perspicaz contribuição.

³⁰ GEWEHR-BORELLA et al., “Transferências grafo-fônico-fonológicas”: uma análise de dados de criança monolíngues (português) e bilíngues (hunsrückisch-português) (2011).

³¹ CABAÑERO e ALVES, “A transferência grafo-fônico-fonológica na produção de seqüências ortográficas ‘ng’ do inglês (L2)”: uma abordagem conexionista (2008).

³² MIRANDA, “A fonologia em dados de escrita inicial de crianças brasileiras” (2014).

³³ ZHOU, “L2 speech learning of European Portuguese /l/ and /r/ by L1-Mandarin learners”: experimental evidence and theoretical modelling (2021).

³⁴ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

estruturais de uma língua. Apesar de ser um tema amplamente discutido na fonologia do PB,³⁵ vê-se, no modelo BiPhon, a oportunidade de uma formalização que preserva aspectos pouco abordados na literatura, como a caracterização acústica desta vogal epentética ou a atenção ao processo de percepção que está integrado ao fenômeno da produção da epêntese.

Em Collischonn,³⁶ a partir de um aparato formal fundamentado em regras fonológicas, é proposto que, na silabação do PB, uma consoante não apta a ocupar a posição silábica de ataque ou coda não seria ligada a nenhum nó silábico, o que desencadearia a criação de uma outra sílaba, desprovida temporariamente de núcleo vocálico. Esse seria o contexto para a inserção de um núcleo vocálico, como é possível observar-se em palavras como *rap[i]to* ou *ap[i]neia*.

Essa noção do fenômeno, em um modelo gerativo de regras, aproxima-se de modelos fonológicos de produção com base em restrições, como a Teoria da Otimidade – OT³⁷ ou a Gramática Harmônica – GH,³⁸ quando se atenta para a representação em dois níveis: forma subjacente (*input* da produção) e forma de superfície (*output* da produção); a diferença está no fato de, na OT ou na GH, a relação entre o *input* e o *output* dar-se por meio de restrições universais e violáveis ao invés de regras.

Nesse sentido, para os modelos de restrições mencionados, conta-se com duas famílias de restrições: Restrições de Marcação, que vão avaliar as formas de *output*, e Restrições de Fidelidade, que vão avaliar a correspondência entre o *input* e o *output*. No modelo BiPhon, esta ideia é ampliada, já que é capaz de formalizar a qualidade acústica dos segmentos (no presente estudo, trata-se da vogal epentética), bem como delimitar o que é fonológico ou fonético, ao representar o processo em maior número de níveis do que os dois utilizados pela Fonologia Gerativa Clássica (a forma subjacente e a forma de superfície), assim como incluir o Módulo da Compreensão como parte da fonologia, conforme apresentado na primeira parte deste artigo, na seção que trata da arquitetura do BiPhon. Destaca-se que a inclusão, no BiPhon-OT, do nível de representação [Forma Fonética] implica a inclusão também de mais uma família de restrições: as Restrições de Pista.

Com a introdução, no BiPhon, do Módulo da Compreensão, o modelo passa a ser bidirecional e vai demandar outras informações (além dos dados de produção que já estavam previstos no modelo clássico da Teoria da Otimidade e em modelos fonológicos anteriores), e incluir outras restrições, que devem ser fornecidas pelo pesquisador. Este é o caso dos dados de percepção e da formalização das Restrições de Pista, que referem um tipo de restrição de fidelidade auditivo-fonológica e que participam do mapeamento entre a [Forma Fonética] e a /Forma Fonológica de Superfície/, na percepção, e entre a /Forma Fonológica de Superfície/ e a [Forma

³⁵ COLLISCHONN, “Um estudo da epêntese à luz da teoria da sílaba de Junko Itô (1986)” (1996); COLLISCHONN, “A epêntese vocálica no Português do Sul do Brasil”: análise variacionista e tratamento pela Teoria da Otimidade (2000); COLLISCHONN, “A Epêntese Vocálica no Português do Sul do Brasil” (2002); COLLISCHONN, “Epêntese vocálica no português do sul do Brasil”: variáveis extralinguísticas (2003); COLLISCHONN, “Epêntese vocálica e restrições de acento no português do sul do Brasil” (2004); KELLER, “O fenômeno da epêntese vocálica no português falado em Panambi e Blumenau” (1999); PARLATO-OLIVEIRA, “Diversidade, variabilidade e frequência em fonologia”: o caso da epêntese vocálica (2008); SILVEIRA e SEARA, “A vogal epentética em encontros consonantais heterossilábicos no português brasileiro”: um estudo experimental (2009).

³⁶ COLLISCHONN, “Análise Prosódica da Sílaba em Português” (1997).

³⁷ PRINCE e SMOLENSKY, *Optimality Theory: Constraint Interaction in Generative Grammar* (2004 [1983]).

³⁸ LEGENDRE et al., “Can Connectionism Contribute to Syntax? Harmonic Grammar, with an Application” (1992).

Fonética], na implementação da produção.

Os dados de análise da epêntese no PB

Na arquitetura simplificada do modelo BiPhon (vide fig. 2), a representação fonética é desmembrada da forma de superfície e reflete dados reais de fala. Desta forma, os dados são fundamentais na formalização do fragmento de gramática em foco. Com relação à representação subjacente, segmentos epentéticos estão, por definição, ausentes neste nível e são o resultado de uma hierarquia com Restrições de Estrutura altamente ranqueadas em detrimento das Restrições de Fidelidade. É nesse sentido que são apresentadas as possibilidades de *input* e *output* no modelo.

Para os dados reais de fala, foram utilizadas as coletas realizadas em Quintanilha-Azevedo, que trazem elementos para a formalização da percepção e da implementação fonética. Em Quintanilha-Azevedo³⁹, a análise instrumental dos dados de produção foi realizada com a utilização do *software* PRAAT⁴⁰, que permitiu a descrição das propriedades acústicas dos segmentos, como duração e frequências formânticas, o que possibilitou a caracterização diferenciada da vogal epentética do português em relação a uma vogal lexical. As palavras, listadas no Quadro 3, p. 82, obedeceram aos critérios assumidos em trabalho anterior de Quintanilha-Azevedo⁴¹, que tinham por objetivo padronizar a amostra: palavras de no máximo quatro sílabas, de apenas duas classes morfológicas (substantivos e adjetivos), e com estrutura morfológica sem prefixo, dando-se preferência a palavras primitivas. Em uma cabine com isolamento acústico, os informantes foram convidados a ler de forma espontânea o instrumento disposto em um modo de apresentação PPT (Microsoft POWER - POINT 2010), em um computador do tipo *laptop*. O instrumento continha frases do tipo: “Diga palavra novamente ou Diga palavra”, aleatorizadas eletronicamente uma única vez através do site <https://www.random.org/lists>. Foram obtidas 42 palavras, em duas frases veículos, totalizando 84 *tokens* analisados para cada informante. Os dados gravados tiveram os segmentos-alvo etiquetados manualmente, cuidadosamente, no *software* PRAAT. Participaram da coleta 8 (oito) informantes da cidade de Pelotas/RS, com idades entre 22 e 34 anos e que não falavam nenhuma língua estrangeira. O instrumento utilizado para a verificação da percepção da epêntese no português, o teste de identificação, foi elaborado a partir do *software* TP⁴² e aplicado aos mesmos informantes que participaram do teste de produção, mas em dias diferentes. Foram 15 (quinze) logatomas, listados nas p. 87 e 88, em Quintanilha-Azevedo⁴³, produzidos com epêntese vozeada, sem epêntese e com epêntese desvozeada, e 2 (duas) vozes (feminina e masculina). Ressalta-se, contudo, que este estudo vai lidar apenas com a plosiva em posição de coda para a sequência ‘p.t’.

³⁹ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

⁴⁰ <https://www.fon.hum.uva.nl/praat/>

⁴¹ QUINTANILHA-AZEVEDO, “A epêntese no Português Brasileiro (L2), em segmentos plosivos em codas mediais, por falantes nativos do Espanhol Colombiano (L1)”: uma análise via Teoria da Otimidade Estocástica e Gramática Harmônica (2011).

⁴² RAUBER et al., *TP – Testes de Percepção / Tarefas de Treinamento Perceptual* (2012).

⁴³ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

reduzindo-se para 3 (três) a quantidade de palavras na produção e na percepção.

Os dados coletados no estudo corroboram os trabalhos sobre produção de epêntese no PB no sentido de que o processo de epêntese após segmentos plosivos que ocupariam a posição de coda medial é variável, em palavras como *'rpto'*, *'cacto'*, *'acne'*, por exemplo. Acrescenta-se que este processo ocorre, também, na percepção, mesmo com um *input* fonético sem quaisquer pistas auditivas diretas para uma vogal após a consoante em coda. O BiPhon, portanto, incorpora à análise informações relativas a fatos fonéticos e perceptuais, conforme já foi explicitado nas seções precedentes deste artigo. Para o exercício aqui proposto, tomam-se, de Quintanilha-Azevedo,⁴⁴ apenas os dados relativos à sequência 'p.t', em palavras como *'rpto'*, por exemplo.

Apresentam-se, no quadro 1, os dados⁴⁵ que compõem este exercício teórico.

		Dados				Dados	
COMPREENSÃO				PRODUÇÃO			
Reconhecimento	↑	/Forma Fonológica Subjacente/		Produção Fonológica	↓	/Forma Fonológica Subjacente/	
		/Forma Fonológica de Superfície/				/Forma Fonológica de Superfície/	
1)	[apto] /ap.to/	[apto]	66,67%	1)	[apto] /ap.to/	[apto]	31,25%
2)	[apto] /api.to/	[apito]	33,33%	2)	[apto] /api.to/	[apito]	68,75%
3)	[apto] /a.to/	[a.to]	0,00%	3)	[apto] /a.to/	[a.to]	0,00%
Percepção	↑	/Forma Fonológica de Superfície/		Implementação	↓	/Forma Fonológica de Superfície/	
		[Forma Fonética]				[Forma Fonética]	
1)	[apto] /ap.to/	[apto]	66,67%	1)	[apto] /apto/	[apto]	31,25%
2)	[apto] /api.to/	[apito]	33,33%	2)	[apto] /apito/	[apito]	68,75%
3)	[apto] /a.to/	[a.to]	0,00%	3)	[apto] /apto/	[a.to]	0,00%

Considerando que o movimento do aprendiz da língua, na condição de ouvinte (Módulo da Compreensão), parte da percepção do som, inicia-se o movimento da apresentação dos dados neste nível. A partir do direcionamento dado pelo modelo BiPhon, fonologicamente, a tarefa de perceber é fazer um mapeamento da representação fonética contínua para uma estrutura de superfície fonológica discreta. Para Boersma,⁴⁶ esta tarefa corresponde à identificação, isto é, envolve a atribuição de um rótulo de categoria a um determinado estímulo sonoro. Este foi o motivo pelo qual Quintanilha-Azevedo,⁴⁷ para a coleta de dados de percepção, partiu de um teste de identificação, no qual o sujeito ouvia um logatoma e tinha de marcar uma entre três opções: se a palavra continha uma inserção vocálica, se estava sem a vogal ou se havia

⁴⁴ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

⁴⁵ Os dados extraídos de Quintanilha-Azevedo encontram-se nas páginas 128 e 130, para o Módulo da Produção e, na p. 149, para o Módulo da Compreensão. Os valores das frequências foram adaptados, já que não serão contemplados todos os *inputs* levantados pelo pesquisador. QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

Quadro 1: Dados de produção e percepção da epêntese no PB, a partir de QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português”.

⁴⁶ BOERSMA, “Cue constraints and their interactions in phonological perception and production” (2009).

⁴⁷ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

o apagamento da consoante na coda da sílaba.

Para o contexto 'p.t', de plosiva em coda medial de palavra, o *input* apresentado aos participantes contava com logatomas com esse contexto (ex.: *lpto*) e as possibilidades de *output* foram estabelecidas de acordo com as previsões da literatura sobre o português, na produção, conforme já referido acima: sem epêntese, com epêntese e com apagamento da coda.

Foi a partir deste instrumento que se chegou aos resultados apresentados no quadro 1, os quais mostram que a percepção de um *input* com 'p.t', como em '*corrupto*', tem a sua identificação com maior frequência, no *output*, com o candidato fiel, sem epêntese. Note-se que, no modelo BiPhon, são apresentados pares de *output*, já que os processos permitem a representação de diferentes *inputs* e *outputs* (vide candidatos no quadro 1).

Do outro lado, no Módulo da Produção, o movimento do aprendiz da língua, na condição de falante, parte da Produção Fonológica que, ao mapear a |Forma Fonológica Subjacente| na /Forma Fonológica de Superfície/, conta apenas com um *input*, sem epêntese, mas com três possibilidades no *output*, na /Forma Fonológica de Superfície/: com epêntese, sem epêntese e com apagamento da consoante na coda da sílaba.

Pelos dados do quadro 1, note-se que, diferentemente do que ocorre na percepção, os resultados mais frequentes nos *outputs* (no mapeamento da /Forma Fonológica de Superfície/ em [Forma Fonética]), ou seja, na produção, foram formas com a presença de epêntese, mostrando que a estrutura mal formada no PB, com plosiva na posição de coda, pode acontecer com epêntese já na percepção, mas é na implementação fonética que ela se mostra mais produtiva, para este contexto (comparem-se os dados do Módulo da Produção com os do Módulo da Compreensão). Estes aspectos devem estar representados na formalização deste fragmento de gramática e devem estar considerados pelas restrições, conforme apresenta a próxima subseção.

As restrições propostas para a formalização do fenômeno da epêntese no BiPhon

Pela ideia central de modelos de restrições, formas de *output* refletem uma resolução de conflitos entre restrições. Cada restrição representa uma exigência sobre algum aspecto gramatical dos candidatos a *output* ótimo; satisfazer uma restrição implica violar outra.

O modelo bidirecional, na arquitetura simplificada (vide fig. 2), ao congrega três níveis de representação, vale-se de três grupos de restrições atuando bidirecionalmente, já que as mesmas restrições responsáveis pela formalização da produção devem explicar

o lado da compreensão; são elas: Restrições de Fidelidade, Restrições de Estrutura e Restrições de Pista. Conforme já visto no início do texto, na seção que trata da arquitetura do BiPhon-OT, as Restrições de Fidelidade avaliam a similaridade entre a /Forma Fonológica de Superfície/ e a |Forma Fonológica Subjacente|. Já as Restrições de Estrutura avaliam a /Forma Fonológica de Superfície/ e interagem com as Restrições de Pista e as Restrições de Fidelidade, de forma a representar a interação entre níveis.

Para o exercício teórico aqui proposto, de formalização do fenômeno da epêntese na sequência heterossilábica /p.t/ (ex.: corrupto), em coda medial, no PB, a partir dos dados de Quintanilha-Azevedo,⁴⁸ serão utilizadas as restrições daquele estudo que sejam relevantes para representar este fragmento de gramática.

As Restrições de Fidelidade que vão compor a formalização do nível lexical, tanto para a produção quanto para a compreensão, são duas:

- DEP⁴⁹ – Segmentos do *output* têm correspondentes idênticos no *input* – não insira
- MAX⁵⁰ – Segmentos do *input* têm correspondentes idênticos no *output* – não apague

Estas restrições são responsáveis pela manutenção da forma presente no nível subjacente |pt|, tanto na produção, quanto no reconhecimento (módulo da Compreensão). Assim, a restrição DEP é violada por um candidato que, no nível /Fonológico de Superfície/, não tem correspondente segmental com o nível |Subjacente|. Por sua vez, a restrição MAX é violada por um candidato que, no nível |Fonológico Subjacente|, não tem correspondência com o nível /Fonológico de Superfície/.

Já as Restrições de Estrutura substituem, no modelo BiPhon, as Restrições de Marcação do modelo da OT. A diferença entre os dois tipos de restrições reside na estrutura que avaliam. Enquanto as Restrições de Marcação avaliam estruturas fonológicas como sílabas, pés, segmentos e, ao mesmo tempo, detalhes fonéticos de um dado segmento, as Restrições de Estrutura devem avaliar somente as propriedades estruturais das /Representações Fonológicas de Superfície/.

A Restrição de Estrutura suficiente para compor a análise, tanto para a produção quanto para a compreensão, segue a seguir:

- *PLOSIVA_{CODA}⁵¹ – plosivas não são licenciadas na posição de coda

Assim, quando as Restrições de Fidelidade DEP ou MAX estiverem acima da Restrição de Estrutura, na hierarquia, eliminando da disputa candidatos que não apresentam correspondência entre

⁴⁸ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

⁴⁹ MCCARTHY e PRINCE, *Faithfulness and reduplicative identity* (1995).

⁵⁰ MCCARTHY e PRINCE, *Faithfulness and reduplicative identity* (1995).

⁵¹ PRINCE e SMOLENSKY, *Optimality Theory: Constraint Interaction in Generative Grammar* (2004 [1983]).

input – output, isto é, candidatos com apagamento da coda ou com epêntese, isto quer dizer que o *output* ótimo será o candidato fiel, sem epêntese. Esta hierarquia deve estar presente tanto na produção fonológica quanto no reconhecimento, confirmando que a gramática é uma só, tanto para a Produção como para a Compreensão.

Por fim, são apresentadas as Restrições de Pista adotadas na análise. Na formalização das Restrições de Pista, o asterisco que apresenta a restrição não exerce função de proibição, tal como acontece com as Restrições de Estrutura, mas negação, para que possam ser lidas e utilizadas nos módulos da Compreensão e da Produção.

Dado que as Restrições de Pista representam a interface do contínuo para o discreto, todas as possibilidades de resultados fonéticos obtidos no levantamento de dados devem constar nas restrições, o que justifica a presença de cada uma delas. Considerando que para este exercício teórico não serão apresentados todos os detalhamentos fonéticos explicitados no trabalho de Quintanilha-Azevedo,⁵² propuseram-se, neste estudo, apenas cinco Restrições de Pista, listadas e definidas a seguir:

- */pi.t/[300Hz] – Não interpretar pistas acústicas de vogal epentética como uma vogal fonológica
- */p.t/[burst] – Não interpretar um burst auditivo como uma consoante em coda fonológica
- */p.t/[] – Não interpretar uma plosiva em posição de coda se não há quaisquer pistas auditivas diretas
- */pi.t/[burst] – Não interpretar um burst auditivo como uma vogal fonológica
- *//[burst] – Não interpretar um burst auditivo como apagamento da coda

Sobre as restrições acima, pode-se afirmar que:

- (a) a opção por utilizar segmentos ao invés de traços para representar a interação com o /nível fonológico/ não implica problemas para a restrição. O uso de rótulo com segmentos, como em */p.t/[burst], pode ser visto apenas como uma simplificação. A forma “(.)” que aparece no início da restrição, entre barras, se refere ao limite da sílaba na representação fonológica.
- (b) as restrições de pista listadas foram descritas considerando-se a percepção (movimento da [Forma Fonética] para a /Forma Fonológica de Superfície/), levando-se em conta a bidirecionalidade das restrições, dependendo do módulo de processamento em análise: se for produção, deve-se realizar a leitura das restrições de forma a preservar o movimento contrário à percepção,

⁵² QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

isto é, da /Forma Fonológica de Superfície/ para a [Forma Fonética] e, assim, trocar a notação de “não interpretar” para “não produzir”.

- (c) com relação às restrições */p.t/[burst], */pi.t/[burst] e *//[burst], considerando-se que no nível fonético a representação se dá pelas pistas, optou-se por identificar a plosiva em coda como [burst]. O “burst” é a explosão de ar obtida com a soltura dos articuladores, durante a realização da oclusiva e, de acordo com Quintanilha-Azevedo⁵³, quase 100% das produções de palavras com a sequência medial ‘p.t’ apresentaram soltura audível, justificando o uso da pista burst.
- (d) com relação à restrição */pi.t/[300Hz], que no nível fonético apresenta uma vogal epentética, a pista utilizada para a sua identificação acústica é o valor da frequência formântica 1 (F1), que representa a altura da vogal. Este valor foi, também, extraído do estudo de Quintanilha-Azevedo⁵⁴.

O esquema apresentado por meio da fig. 7, a seguir, resume a posição de cada restrição na arquitetura simplificada do modelo BiPhon e suas interfaces nos diferentes níveis de representação:



É com base nestas informações acerca de dados e restrições que se dará a análise no modelo BiPhon. Reitera-se que não apenas os dados, mas o levantamento das restrições pertinentes para a formalização do fragmento de gramática em estudo é de responsabilidade do pesquisador. Além disso, um exercício teórico não vai apresentar todas as restrições presentes na gramática do português, mas a hierarquia de restrições que represente a amostra de dados apresentada. Note-se que, por exemplo, o *input* para a percepção poderia contar com formas com e sem epêntese, mas, por questão de definição metodológica, optou-se por demonstrar apenas o resultado do teste de identificação para as formas de *input* sem epêntese [pt], apesar de a literatura e mesmo a coleta de dados de Quintanilha-Azevedo⁵⁵ mostrar que o *input* da percepção possui candidatos com e sem epêntese. Esta decisão metodológica implicará um conjunto mais restrito de Restrições de Pista para a percepção.⁵⁶

Finalmente, para que seja possível a formalização dos dados e a demonstração da capacidade explicativa do modelo BiPhon, é importante estar clara a função do algoritmo de aprendizagem

⁵³ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

⁵⁴ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016), 118, tabela 3.

Figura 7: Arquitetura simplificada do modelo BiPhon, com as restrições e frequências de ocorrências das formas de *input* e *output*.

⁵⁵ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

⁵⁶ Neste ponto, optou-se por resumir os *inputs* da percepção para candidatos sem epêntese ([pt]), pois cada possibilidade lançada ([pit] ou []) incorreria na necessidade de mais Restrições de Pista. Assim como, na implementação fonética, seriam excluídas das possibilidades pares como /pi.t/[pt] ou /p.t/[pit], tendo em vista que não fazem sentido na produção. Esta decisão terá implicação na bidirecionalidade do conjunto de restrições. Acredita-se, apesar disto, que não haverá prejuízo na análise proposta. Maior detalhamento descritivo será dado na Seção 4, na formalização dos dados.

junto ao modelo. Para tanto, a próxima subseção apresentará o algoritmo, relacionando a sua participação na organização das restrições no *ranking*, bem como destacando as suas propriedades e capacidade na demonstração não apenas de *outputs* em variação, como na possibilidade de informar quais *outputs* são mais frequentes.

O Algoritmo de Aprendizagem Gradual – GLA e os dados de epêntese

Uma versão estocástica da Teoria da Otimidade, tal qual será utilizada nas simulações, no BiPhon, para a avaliação do candidato a *output* ótimo, requer um algoritmo de aprendizagem. É o algoritmo que tem a missão de dar conta da movimentação das restrições, de maneira lógica e pré-estabelecida, até a formação de uma hierarquia de restrições que dê conta da formalização de uma gramática que apresente o candidato mais harmônico, seja ele um candidato fiel (exemplo no nível da percepção: [apto] → /apto/) ou um candidato menos marcado (exemplo no nível da implementação fonética.: /api.to/ → [apito]).

O algoritmo de aprendizagem, no presente estudo o **Algoritmo de Aprendizagem Gradual – GLA**,⁵⁷ proposto pelo mesmo autor do modelo BiPhon,⁵⁸ é o que guiará o aprendiz para a hierarquia da língua. É o que vai garantir certa flexibilidade ao modelo para que forneça uma resposta mais próxima aos dados reais, isto é, vai permitir que sejam demonstradas as variações que ocorrem não apenas na produção, mas na compreensão, conforme acontece nos dados de análise apresentados no quadro 1.

Em contraste com a Teoria da Otimidade Clássica, a Teoria da Otimidade Estocástica, como o próprio nome traduz, permite eventos aleatórios e, neste caso, duas são as possibilidades a *output* ótimo: um candidato a *output* ótimo categórico e um candidato a *output* ótimo variável. E, para tanto, o algoritmo, que descreve o passo a passo que as restrições percorrem até chegar à hierarquia de restrições ideal, atribui valores ou pesos numéricos às restrições a partir de um estágio inicial.

Note-se que não apenas a definição das restrições, mas os pesos iniciais das restrições são teorizados pelo pesquisador,⁵⁹ e o modelo deve ser alimentado manualmente, via *scripts*, para que possa dar início à simulação esperada. O termo *Script*, na informática, significa um conjunto de instruções para que determinada tarefa seja executada em um determinado aplicativo. Neste caso, a tarefa é chegar à hierarquia de restrições que represente o fenômeno em análise e o aplicativo é o sistema PRAAT, desenvolvido por Paul Boersma e David Weenink, do Instituto de Fonética da Universidade de Amsterdã. As instruções referem-se, em um *script*, às restrições utilizadas, aos candidatos a *input* e *output* e às violações

⁵⁷ O GLA é implementado no PRAAT e cada rodada do algoritmo representa uma simulação de uso da língua.

⁵⁸ BOERSMA e HAYES, “Empirical tests of the gradual learning algorithm” (2001).

⁵⁹ Será considerada uma gramática em um estágio inicial em que todas as restrições têm os mesmos pesos, arbitrariamente classificados em 100, tal como o fazem Boersma e Hayes (2001).

incurridas por cada candidato e, em outro *script*, aos candidatos a *input* e *output* e às frequências de ocorrência de cada um deles. Estas informações são cruzadas e pesos são fornecidos pelo algoritmo a cada simulação linguística. A partir de então, passam a ser atributo do algoritmo as alterações dos pesos iniciais das restrições, respeitando-se os 4 (quatro) passos para ele estabelecidos (Dados, Geração, Comparação e Ajuste) até o estágio final.

Assim, cada restrição receberá, em um estágio inicial, 2 (dois) valores numéricos ou pesos: um valor de ranqueamento ou valor central (*ranking value*) e um valor de seleção (*disharmony*), com funções diferentes; cabe ao **valor central** estabelecer se as restrições podem alternar a sua colocação na hierarquia e cabe ao **valor de seleção** estabelecer a ordem das restrições no *ranking* naquele momento de produção: quanto maior o valor de seleção, mais alta a restrição estará no *ranking*.

Na prática, as restrições deixam de ser um ponto em uma hierarquia e passam a ser uma curva de distribuição normal, conforme é apresentado na fig. 8.

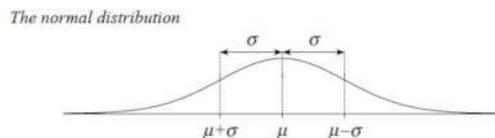


Figura 8: Exemplo de curva de distribuição normal.

Dessa forma, com o uso do algoritmo, cada uma das restrições que compõem a representação do fragmento de gramática em análise segue a operacionalização da curva da distribuição normal. Uma distribuição é descrita por sua média ' μ ', que é o valor central da restrição, e pelo desvio padrão ' σ ', que descreve a largura da curva (ver fig. 8). Além disso, 68% (sessenta e oito por cento) dos valores estão contidos em uma distribuição normal, dentro de um desvio padrão a partir da média (' $\mu - \sigma$ ' e ' $\mu + \sigma$ '). Isto quer dizer que o **valor central** da restrição é um valor fixo e o **valor de seleção** vai mudar a cada simulação, isto é, a cada momento de produção linguística, sendo que os **valores de seleção** têm maior probabilidade de ocorrer mais próximo ao valor central.

Com isso, todas as restrições da análise estarão representadas por uma curva de mesmo tamanho, com 10 (dez) pontos; se a média ou valor central da restrição for, por exemplo, igual a 5 (cinco), o valor de ponto de seleção pode variar de 0 (zero) a 10 (dez).

Os *tableaux* exemplificativos, a seguir, apresentam candidatos a *output* categóricos e variáveis, na versão de avaliação estocástica.

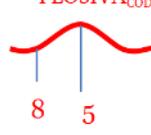
Input →	/kaktó/	 <p>DEP</p> <p>Valor Central = 20 Valor de Seleção = 17</p>	 <p>*PLOSIVA_{CODA}</p> <p>Valor Central = 5 Valor de Seleção = 8</p>
Candidato a <i>output</i> 1	☺ [kaktó]		*
Candidato a <i>output</i> 2	[kakitó]	*!	

Tableau 1: Tableau exemplificativo de um candidato categórico.

Neste *tableau 1* ficou exemplificado um candidato categórico que, a partir de um *input* na /Forma Fonológica de Superfície/ sem epêntese (/kaktó/), mostra um *output* sem epêntese na [Forma Fonética] ([kaktó]). Para a formalização de dados com apenas uma possibilidade de *output*, o algoritmo vai estipular, para as restrições decisivas na disputa, valores centrais com distância superior a 10 (dez) pontos; no presente caso DEP = 20 e *PLOSIVA_{CODA} = 8 (diferença de 12 pontos). Assim, não seria possível que o valor de seleção da restrição de marcação *PLOSIVA_{CODA}, que se encontra abaixo, no *ranking*, pudesse em outro momento estar acima, conforme mostrado no exemplo a seguir:

Input →	/kaktó/	 <p>*PLOSIVA_{CODA}</p> <p>Valor Central = 5 Valor de Seleção = 6</p>	<p>DEP</p> <p>Valor Central = 7 Valor de Seleção = 4</p>
Candidato a <i>output</i> 1	[kaktó]	*!	
Candidato a <i>output</i> 2	☺ [kakitó]		*

Tableau 2: Tableau exemplificativo de um candidato variável.

Observe-se que, neste segundo caso, os valores centrais das restrições têm diferença inferior a 10 pontos: DEP tem valor central igual a 7 e *PLOSIVA_{CODA}, igual a 5 (diferença de 2 pontos). Como são os valores de seleção que organizam a hierarquia das restrições e, neste momento, o valor de seleção da Restrição de Estrutura (*PLOSIVA_{CODA} = 6) é maior do que o valor de seleção da Restrição de Fidelidade (DEP = 4), isto quer dizer que, neste momento, o candidato menos marcado, com epêntese, é o vencedor, mas poderia, em outro momento, o candidato fiel voltar a ser o vencedor, em razão de uma inversão no *ranking* das restrições.

Outra capacidade importante do algoritmo é a possibilidade de mostrar o candidato que ocorre de forma mais frequente no *out-*

put. A restrição que recebe o peso central maior claramente estará por mais vezes acima, no *ranking*, definindo o candidato que, de forma mais frequente, será o ótimo. No *tableau* exemplificativo 2, o candidato mais frequente é o candidato fiel, sem epêntese, pois, apesar de, neste momento, DEP estar abaixo, na hierarquia, por ter recebido um valor de seleção menor (4 pontos), seu valor central é maior do que o da outra restrição.

Apresentadas as propriedades do algoritmo GLA utilizado nesta versão do BiPhon, ressalta-se a validade de se trabalhar com um algoritmo de aprendizagem acoplado a um modelo fonológico. Entretanto, observa-se que o diferencial do modelo não está apenas no uso do algoritmo, mas no seu caráter bidirecional e multinível, conforme já explanado nas seções anteriores e conforme será apresentado no exercício teórico, na próxima seção.

A formalização da epêntese vocálica à luz do BiPhon-OT

Nesta Seção mostrar-se-á como acontecem as interações entre os níveis fonético e fonológico, nos módulos da Compreensão e da Produção, na formalização do fenômeno da epêntese no PB, na sequência heterossilábica ‘p.t’, em coda medial, no modelo BiPhon-OT. Para aplicar com clareza os conceitos de cada interface, optou-se por formalizar o processamento dos dois módulos em separado, começando-se pela Percepção, no módulo da Compreensão (vide fig. 2).

Percepção (Módulo da Compreensão)

Seguindo-se a concepção de que a percepção é, também, fonológica, ela deve ser altamente influenciada pela gramática da língua. No presente caso, tal fato é demonstrado a partir dos resultados obtidos no teste de identificação aplicado aos informantes: quando confrontados com uma forma sendo produzida com a plosiva em coda medial (‘aptidão’), de forma variável, percebiam-na como forma com epêntese e sem epêntese.

Tal fato se deve à restrição, no português brasileiro, à plosiva em posição de coda. Assim, o modelo deve dar conta da possibilidade de *outputs* – /Forma Fonológica de Superfície/ (SF) – como /p.t/ e /pi.t/ a partir de um *input* – [Forma Fonética] (AudF) – igual à forma [pt], sem epêntese.

No presente estudo, por decisão metodológica de recorte de dados, foram considerados apenas [*inputs*] sem epêntese, os quais estão denominados como [pt]AudF, conforme demonstrado na fig. 7 e formalizado no *tableau* 3, a seguir.

	ranking value	disharmony	plasticity
*[burst] não / /	107.244	109.108	1.000000
*[burst] não /pi.t/	97.499	95.907	1.000000
*{plosiva}coda	95.257	95.457	1.000000
*[burst] não /p.t/	95.257	93.518	1.000000

[pt]AudF	*[burst] não / /	*[burst] não /pi.t/	*{plosiva}coda	*[burst] não /p.t/
[pt]AudF /api.to/SF		*!		
[pt]AudF /ap.to/SF			*	*
[pt]AudF /a.to/SF	*!			

Tableau 3: Formalização do nível da Percepção no modelo BiPhon – Primeira simulação (mapeamento da [Forma Fonética] em /Forma Fonológica de Superfície/).

Para um *input* sem epêntese, a frequência de ocorrência das possibilidades de cada candidato a *output* foi esta:

- 33,33% dos informantes ouviram a plosiva em coda como se houvesse uma vogal epentética
- 66,67% dos informantes ouviram a forma fiel ao *input*.

Isto quer dizer que foram fundamentais, na escolha dos candidatos, a Restrição de Estrutura (*PLOSIVACODA) e a Restrição de Pista que milita contra a epêntese na /Forma Fonológica de Superfície/, para a qual não há qualquer pista auditiva direta ([burst] não /pi.t/) e não há a percepção de [burst] como plosiva em posição de coda /p.t/ ([burst] não /p.t/).

Sabendo-se, assim, que duas são as possibilidades a *output* ótimo, a hierarquia de restrições deve estar disposta de forma que a restrição violada pela forma com epêntese no *output* fonológico (*[burst] não /pi.t/) e as restrições violadas pela forma sem epêntese (*[burst] não /p.t/ e *PLOSIVACODA) possam mudar de posição na hierarquia para que sejam possíveis várias formas de saída.

Conforme apresentado na Subseção que tratou do algoritmo de aprendizagem, é graças aos pesos centrais (*ranking value*) e de seleção (*disharmony*) que as restrições podem alterar sua posição no *ranking*. Se a cada momento de produção linguística (simulação no PRAAT) o valor de seleção é alterado, dentro de uma faixa de cinco pontos acima e cinco pontos abaixo do valor central, isto quer dizer que o candidato ótimo sem epêntese (vide *tableau 3*) pode não vencer a disputa em outro momento, conforme apresenta o *tableau* a seguir.

	ranking	value	disharmony	plasticity
*[burst] não / /	107.244	104.482	1.000000	
*{plosiva}coda	95.257	96.986	1.000000	
*[burst] não /p.t/	95.257	96.539	1.000000	
*[burst] não /pi.t/	97.499	95.899	1.000000	

[pt]AudF	*[burst] não / /	*{plosiva}coda	*[burst] não /p.t/	*[burst] não /pi.t/
[pt]AudF /api.to/SF				*
[pt]AudF /ap.to/SF		*!	*	
[pt]AudF /a.to/SF	*!			

Tableau 4: Formalização do nível da Percepção no modelo BiPhon – Segunda simulação (mapeamento da [Forma Fonética] em /Forma Fonológica de Superfície/).

Note-se que o valor central (*ranking value*) das restrições não mudou. Logo, para que seja possível identificar os candidatos mais frequentes, as restrições com pesos centrais mais altos estarão mais vezes acima, no *ranking*. Mesmo que sejam possíveis duas formas de *output* ótimo, é necessário identificar que uma delas, no caso a forma sem epêntese (/ap.to/), aparece de forma mais frequente, em 66,67% das ocorrências.

Em resumo, o fato de a restrição *[burst] não /pi.t/ ter recebido o peso central igual a 97.499, superando o peso das restrições *PLOSIVA_CODA (95.257) e *[burst] não /p.t/ (95.257), informa que o candidato com epêntese [pit]AudF incorre em violação fatal mais vezes do que o candidato sem epêntese, mas o fato de o peso central destas restrições ter diferença inferior a 10 pontos aponta que estão em variação.

Com relação ao terceiro candidato a *output* incluído na análise, destaca-se que esse foi proposto com apagamento da coda na /Forma Fonológica de Superfície/ ([pt]AudF /a.to/SF) para demonstrar que também há, na gramática, uma forma de dominância entre as restrições, sem possibilidade de variação. Assim, o peso central da restrição violada pelo candidato com apagamento (107.244), posicionada acima das outras restrições, além de demonstrar que o candidato a *output* com apagamento é o menos frequente, também mostra que não altera de posição com outra restrição, pois o seu peso tem diferença de mais de 10 pontos em relação às demais restrições.

Diante do exposto, a hierarquia das restrições mais os pesos centrais evidenciam e formalizam que o *output* fiel, sem epêntese, aparece como ótimo mais vezes nas simulações em relação ao candidato menos marcado, com epêntese. Sendo assim, a hierarquia vai apresentar as Restrições de Pista, que agem como se fossem Restrições de Fidelidade, comparando a [Forma Auditivo-Fonética] com a /Forma Fonológica de Superfície/, acima no *ranking* mais vezes, dominando a Restrição de Estrutura. Ora, se uma sequência heterossilábica, tal como a que está em análise neste estudo ('p.t'), que apresenta um platô de sonoridade em termos de contato silábico, expressa maior identificação da plosiva em posição de coda,

em detrimento de uma forma com a inserção de uma vogal, por suas características acústicas, isto quer dizer que a questão deve estar prevista no nível fonético.

Em consideração ao exposto, para o nível da percepção, destacam-se duas inovações relevantes do modelo bidirecional com relação aos modelos tradicionais de fonologia:

- (a) a percepção está incluída na fonologia
- (b) a fonética está separada da fonologia na /Forma Fonológica de Superfície/.

Tais inovações, para o presente estudo, foram relevantes, pois com a inclusão deste nível de representação foi possível explicar a possibilidade de identificação, no português brasileiro, de uma vogal “ilusória” em um *input* que não possui vogal: uma vogal que foi percebida graças a uma restrição a uma estrutura que não se enquadra nos requisitos da língua, confirmando que a percepção é fonológica. Além disto, a /Forma Fonológica de Superfície/, sendo um nível intermediário entre a |Forma Fonológica Subjacente|, forma mais abstrata, e a [Forma Fonética], permite que seja possível a identificação de formas distintas no *output*.

Reconhecimento (Módulo da Compreensão)

Frise-se que no nível do reconhecimento, nível meramente fonológico, as restrições que articulam a /Forma Fonológica de Superfície/ com a |Forma Fonológica Subjacente| são diferentes das restrições envolvidas no nível da percepção. No reconhecimento, a disputa passa a ser entre restrições de Fidelidade e Estrutura.

Os *outputs* do nível da percepção, na /Forma Fonológica de Superfície/, passam a ser o *input* do reconhecimento até a |Forma Fonológica Subjacente|. Logo, são dois os candidatos possíveis: com epêntese /pi.t/ e sem epêntese /p.t/, que devem, neste caso, chegar à forma de *output* sem epêntese (|pt|).

Da mesma forma tratada no nível da percepção, foi incluída como candidato a *output* uma opção com apagamento da coda apenas para demonstrar que, de forma categórica, este candidato não é uma possibilidade. Esta situação está presente nos dados apresentados no quadro 1.

De acordo com o levantamento realizado em Quintanilha-Azevedo⁶⁰, o apagamento da plosiva em posição de coda não foi um recurso utilizado pelos participantes da pesquisa nem na percepção, nem na produção, motivo pelo qual o candidato com apagamento sempre incorrerá em violação fatal neste exercício teórico.

⁶⁰ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

12. OTGrammar Reconhecimento_fonológica
File Edit Query Font

	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
MAX	107.800	107.785	1.000000
DEP	96.694	96.015	1.000000
*plosiva{coda}	95.506	94.636	1.000000

/apito, apto, ato/SF	MAX	DEP	*plosiva{coda}
/api.to/SF apto UF		*!	
☞ /ap.to/SF apto UF			*
/a.to/SF apto UF	*!		

Tableau 5: Formalização do nível do Reconhecimento no modelo BiPhon (mapeamento da /Forma Fonológica de Superfície/ em |Forma Fonológica Subjacente|).

Conforme foi mostrado no quadro 1, e replicado na fig. 7, os candidatos a *output* não têm opção no reconhecimento: devem chegar à representação subjacente |pt| nas frequências abaixo:

- “/pi.t/SF |pt|UF” 33.33%
- “/p.t/SF |pt|UF” 66.67%

Isto quer dizer que, para o caso de plosiva em coda medial, no português, a representação subjacente nem sempre é a própria representação fonética. A plosiva em posição de coda desrespeita os requisitos estruturais do português, demandando uma alternativa para a correção. Em sendo assim, cabe ao modelo demonstrar não apenas que o *output* é variável, mas que ora é escolhido um candidato fiel, ora é escolhido um candidato com uma correção para satisfazer as restrições da língua.

Na simulação apresentada no *tableau 5* vê-se um momento no qual o ouvinte percebeu o som da plosiva em coda, ou seja, categorizou-o na /Forma Fonológica de Superfície/ /pt/ e vai categorizá-lo na |Forma Fonológica Subjacente|. Logo, a Restrição de Fidelidade DEP ficou acima, na hierarquia, mostrando que, neste momento, o *output* ótimo foi um candidato fiel, sem epêntese.

Cumpramos ressaltar, novamente, que o modelo, com o uso do algoritmo de aprendizagem, vai além de demonstrar um *output* em variação; ele refere também qual é o candidato mais frequente. Retomando os pesos centrais das restrições, estabelecidos no *tableau 5*, tem-se que DEP (96.694) e *PLOSIVACODA (95.506) apresentam os pesos centrais próximos, mas o peso central de DEP é maior, fazendo com que esteja mais vezes acima, no *ranking* e, assim, fazendo com que o candidato com epêntese seja eliminado da disputa mais vezes em detrimento do candidato sem epêntese.

Note-se que a Restrição de Fidelidade MAX permanece acima, na hierarquia, com peso central (107.800) com distância superior a 10 pontos das demais restrições, o que evidencia que, categori-

camente, o candidato com apagamento nunca será ótimo, pois sempre receberá uma violação fatal.

Ressalta-se, no entanto, que as informações explicativas dadas pelo algoritmo, que permitem verificar candidatos em variação e o candidato mais frequente, não são exclusivas do BiPhon. Um diferencial do modelo está em seus três níveis de representação, para demonstrar características específicas da língua com relação a detalhes fonéticos e, conseqüentemente, na inclusão de restrições que não são, portanto, universais; outro diferencial está na inclusão do módulo da Compreensão na análise fonológica.

Produção Fonológica (Módulo da Produção)

Conforme foi apresentado, o modelo BiPhon é bidirecional, isto é, prevê que, com as mesmas restrições e a mesma hierarquia, é possível formalizar-se: reconhecimento/produção fonológica e percepção/implementação fonética.

Assim, para a formalização do nível da produção fonológica, três são as restrições interagindo de modo a apresentar a hierarquia que represente o fenômeno (DEP, MAX e $*_{\text{PLOSIVA}}\text{CODA}$), conforme o *Tableau 6* a seguir:

	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
MAX	106.727	105.595	1.000000
DEP	95.970	96.687	1.000000
*{plosiva}codas	97.304	94.558	1.000000

apto UF	MAX	DEP	*{plosiva}codas
apto UF /api.to/SF		*!	
apto UF /ap.to/SF			*
apto UF /a.to/SF	*!		

Tableau 6: Formalização do nível da Produção Fonológica no modelo BiPhon (mapeamento da |Forma Fonológica Subjacente| em /Forma Fonológica de Superfície/)

Comparando-se o *tableau 6*, da formalização da produção fonológica, com o *tableau 5*, da formalização do reconhecimento, vê-se que ambos têm a mesma hierarquia de restrições, para o mesmo *output* ótimo, isto é, o candidato sem a correção por epêntese. E, como o *output* é variável, a diferença entre a Restrição de Fidelidade, que coíbe a epêntese, e a de Estrutura, que proíbe a plosiva em posição de coda, é menor do que 10 pontos, indicando que podem estar em variação, permitindo mais de um candidato ótimo. Então, essa formalização permite observar-se que a competência de um informante adulto, ao tratar-se da relação entre os módulos da Produção e da Compreensão, é a mesma.

O nível da produção fonológica, nível meramente fonológico, é

um nível mais familiar às análises em fonologia, especialmente familiar às análises realizadas com a Teoria da Otimidade ou a Gramática Harmônica, mas diferencia-se destas no que se refere às restrições. Os detalhes fonéticos de um dado segmento, no modelo BiPhon, devem ser avaliados apenas nos níveis inferiores, no nível da implementação fonética (no Módulo da Produção) ou no nível da percepção (no Módulo da Compreensão). É por esse motivo que o modelo adota a nomenclatura de Restrições de Estrutura e não mais Restrições de Marcação, pois as Restrições de Estrutura devem avaliar somente as propriedades estruturais das representações fonológicas da /Forma Fonológica de Superfície/.

O último *tableau*, apresentado a seguir, trata do nível da Implementação Fonética ([Forma Fonética]). Considerando-se a novidade deste último nível aos modelos fonológicos, a análise, naturalmente, é um pouco mais exaustiva. Cumpre ressaltar, novamente, que, em se tratando de um modelo bidirecional, as mesmas restrições utilizadas na análise do nível da percepção compõem esta próxima análise.

Implementação Fonética (Módulo da Produção)

Para este nível têm-se quatro restrições compondo a análise: são três Restrições de Pista e uma Restrição de Estrutura:

- */pi.t/ não [300Hz]
- */p.t/ não [*burst*]
- */p.t/ não []
- *PLOSIVA_{CODA}

Entretanto, comparando-se este conjunto de restrições com as restrições propostas para o nível da percepção, observa-se que a restrição */pi.t/ não [*burst*], utilizada na percepção, não será utilizada nesta análise, pois não foi previsto candidato a *output* sem epêntese a partir de uma forma fonológica com epêntese, no nível da implementação fonética.

Se na /Forma Fonológica de Superfície/ o informante já se utiliza do recurso da epêntese para corrigir a plosiva em posição de coda, é pouco provável que na implementação ele retorne à forma fiel. De qualquer modo, esta forma (|apto|UF|/api.to/SF [pt]AudF) poderia ter sido prevista como candidato a *output*. A diferença da formalização entre implementação fonética e percepção, neste caso, é a de que a restrição */pi.t/ não [*burst*], que seria violada pelo candidato |apto|UF| /api.to/SF [pt]AudF, na implementação fonética, estará sempre acima, no *ranking*, com peso central com distância superior a 10 pontos em relação às demais restrições, tal

como ocorreu, nos outros níveis, com os candidatos com apagamento, e essa posição da restrição no *ranking* tiraria o candidato da disputa.

No caso da percepção, esta restrição */pi.t/ não [burst] recebeu um peso central do algoritmo com distância inferior a dez pontos da Restrição de Estrutura, indicando que é um candidato possível.

Do mesmo modo, a restrição */pi.t/ não [300Hz], presente nesta análise, não aparece nas restrições do nível da percepção. Como foi considerado apenas o *input* sem epêntese, do teste de identificação realizado em Quintanilha-Azevedo,⁶¹ esta restrição não teria funcionalidade naquele nível. Por outro lado, na implementação fonética foram observados 68,75% deste *output* com epêntese.

Esta é uma questão de recorte metodológico. No modelo BiPhon, quanto mais candidatos são previstos na análise, mais Restrições de Pista são necessárias. A formalização destas restrições constitui-se em uma inovação que se choca com a noção gerativista de economia do modelo. Entretanto, a ideia do modelo BiPhon é a de que o contínuo capte cada uma das possibilidades acústicas relevantes no processamento da relação entre o contínuo e o categórico. O funcionamento da gramática incorpora a interface entre o fonético e o fonológico.

O limitador do conjunto de Restrições de Pista não está no conjunto de restrições universais (CON), tal como ocorre com as Restrições de Estrutura e Fidelidade. No entanto, as Restrições de Pista também parecem ter o seu limitador: esse seria o espaço acústico ocupado pela categoria prototípica do fenômeno em análise.

Essa é uma situação interessante, pois corrobora a relevância da existência de dois módulos (de Compreensão e de Produção) em um modelo fonológico. Com o mesmo conjunto de restrições foi possível ver que, na percepção do PB, os informantes ouviram uma vogal ilusória em um *input* sem epêntese, mas na produção esta situação não faria sentido.

A observação interposta poderia levar ao entendimento de que grande parte das restrições seria dispensável na análise, por não estarem sendo violadas, mas deve ser salientado que as mesmas Restrições de Pista devem lidar tanto com a implementação fonética, como com a percepção. Sendo assim, todas elas são essenciais na análise completa do fenômeno, pois vão desempenhar um papel na produção ou na compreensão; o mesmo conjunto de restrições deve oferecer a base para que uma única gramática seja implementada. Neste artigo tomou-se a liberdade de extrair as restrições que não receberiam violações, pois as análises foram realizadas em separado, por nível, permitindo esta decisão.

⁶¹ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

O *tableau* a seguir apresenta um momento de produção no qual o candidato ótimo, na [Forma Fonética] (Forma Fonética – AudF), é sem epêntese:

	ranking	value	disharmony	plasticity
*p.t/ não []	106.871		109.989	1.000000
*/pi.t/ não [300Hz]	99.329		100.010	1.000000
*{plosiva}coda	100.671		99.665	1.000000
*p.t/ não [burst]	93.800		91.226	1.000000

apto UF	*/p.t/ não []	*/pi.t/ não [300Hz]	*{plosiva}coda	*p.t/ não [burst]
apto UF /api.to/SF [pit]AudF		*!		
apto UF /ap.to/SF [pt]AudF			*	*
apto UF /ap.to/SF []AudF	*!		*	

Tableau 7: Formalização do nível da Implementação Fonética no modelo BiPhon.

Na implementação fonética, a frequência de ocorrência dos candidatos foi de:

- 68,75% para o candidato com epêntese na forma fonética (|apto|UF /api.to/SF [pit]AudF) e
- 31,25% para o candidato sem epêntese na forma fonética (|apto|UF /ap.to/SF [pt]AudF)

Nesta simulação (*tableau 7*) tem-se como candidato ótimo a forma fiel |pt|UF /p.t/SF [pt]AudF. Isto quer dizer que as restrições que estão impedindo os candidatos a *output* com a correção por epêntese (|pt|UF /pi.t/SF [pit]AudF) e por apagamento (|pt|UF /p.t/SF []AudF) de vencer a disputa encontram-se acima, na hierarquia, demonstrando que, neste momento de produção, o candidato ótimo é fiel, isto é, sem epêntese. Entretanto, o candidato mais frequente é aquele com a epêntese. Tal fato está demonstrado no *tableau 7*, no peso central que recebe a Restrição de Estrutura que proíbe a plosiva em posição de coda (100.671).

O avanço do entendimento da inter-relação entre fatos fonéticos e fonológicos tem levado à documentação de fatos fonéticos na formalização fonológica. Assim, aponta-se a vantagem de se olhar para o dado fonético e de incorporá-lo aos modelos de análise fonológica, tal como faz o modelo eleito para o estudo aqui apresentado. Com a formalização do fenômeno da epêntese em multiníveis, fica mais fácil representar a diferença que há entre a percepção e a produção do fenômeno aqui estudado no português.

Com o exercício teórico proposto, utilizando-se o fenômeno da epêntese no português, acredita-se ter sido possível apresentar a ideia do modelo BiPhon como modelo de gramática fonológica.

Considerações finais

Com o objetivo de apresentar uma síntese do Modelo Bidirecional de Fonologia e Fonética, proposto por Boersma,⁶² este artigo procurou trazer os fundamentos do BiPhon-OT, explicitando as notações utilizadas nas representações fonética e fonológica, bem como a relação entre compreensão e produção introduzidas no modelo. Reforçando a noção de que percepção não é o inverso da produção, mas que é parte do Módulo da Compreensão e que é, portanto, parte da gramática, buscou-se mostrar que, com o BiPhon, a visão da percepção passa a ter uma perspectiva fonológica, que integra os mecanismos de formalização da gramática.

Sem a intenção de comparar o BiPhon a modelos anteriores para defender alguma vantagem ou desvantagem, mas com a ideia de esclarecer as bases da proposta teórica, refletindo-se sobre a sua arquitetura, destacaram-se as características que determinam a denominação de “Modelo Bidirecional de Fonologia e Fonética” em suas dimensões teóricas e aplicadas.

Com o exemplo de aplicação do modelo por meio da simulação dos dados de epêntese no PB, emprestados de Quintanilha-Azevedo,⁶³ e da representação proposta na fig. 7, foi possível observar-se:

- a) como os diferentes níveis ([Forma Fonética], /Forma Fonológica de Superfície/ e |Forma Fonológica Subjacente|) são articulados via restrições;
- b) quais conjuntos de restrições (Estrutura, Fidelidade e Pista) estão presentes na interação entre os níveis ([Forma Fonética], /Forma Fonológica de Superfície/ e |Forma Fonológica Subjacente|);
- c) como é representado cada nível;
- d) como atuam os módulos da Compreensão e da Produção, considerando-se a bidirecionalidade das restrições envolvidas na análise;
- e) como são formalizadas as pistas acústicas, pensando-se que os detalhes fonéticos são importantes para alguns processos fonológicos;
- f) como a fonética e a fonologia são reunidas e interagem entre si.

Com as noções que dão suporte ao modelo e com a arquitetura que oferece para a análise e a formalização de dados de natureza fonético-fonológica, o BiPhon-OT parece cumprir o papel de modelo de Fonologia e Fonética, tanto na condição de modelo gramática com também na condição de modelo de processamento.

⁶² BOERSMA, “Functional Phonology”: Formalizing the interactions between articulatory and perceptual drives (1998); BOERSMA, “Some listener-oriented accounts of h-aspiré in French” (2007); BOERSMA, “Cue constraints and their interactions in phonological perception and production” (2009); BOERSMA, “A programme for bidirectional phonology and phonetics and their acquisition and evolution” (2011).

⁶³ QUINTANILHA-AZEVEDO, “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português” (2016).

Referências

- BOERSMA, Paul (1998). "Functional Phonology: Formalizing the interactions between articulatory and perceptual drives". PhD dissertation. Amsterdam: Universiteit van Amsterdam, p. 327.
- BOERSMA, Paul (2006). "Prototypicality Judgements as Inverted Perception". In: *Gradience in Grammar: Generative Perspectives*. Editado por Gisbert FANSELOW, Caroline FÉRY, Ralf VOGEL e Matthias SCHLESEWSKY. Oxford: Oxford University Press, pp. 167–184. ISBN: 9780199274796.
- BOERSMA, Paul (2007). "Some listener-oriented accounts of h-aspiré in French". *Lingua* 117.12, pp. 1989–2054. DOI: 10.1016/j.lingua.2006.11.004.
- BOERSMA, Paul (2009). "Cue constraints and their interactions in phonological perception and production". In: *Phonology in Perception*. Editado por Paul BOERSMA e Silke HAMANN. Berlin: Mouton de Gruyter, pp. 55–110. ISBN: 9783110219227.
- BOERSMA, Paul (2011). "A programme for bidirectional phonology and phonetics and their acquisition and evolution". In: *Bidirectional Optimality Theory*. Editado por Anton BENZ e Jason MATTAUSCH. Amsterdam: John Benjamins Publishing Company, pp. 33–72. ISBN: 9789027284525.
- BOERSMA, Paul, Titia BENDERS e Klaas SEINHORST (2020). "Neural network models for phonology and phonetics". *Journal of Language Modelling* 8.1, pp. 103–177. DOI: 10.15398/jlm.v8i1.224.
- BOERSMA, Paul e Silke HAMANN (2006). *Sibilant inventories in bidirectional phonology and phonetics*. Third Old World Conference in Phonology. handout. URL: <https://fon.hum.uva.nl/paul/presentations/BoersmaHamannHandout.pdf>.
- BOERSMA, Paul e Silke HAMANN (2009). "Loanword adaptation as first-language phonological perception". In: *Loan Phonology*. Editado por Andrea CALABRESE e W. Leo WETZELS. Amsterdam: John Benjamins Publishing Company, pp. 11–58. ISBN: 9789027288967.
- BOERSMA, Paul e Bruce HAYES (2001). "Empirical tests of the gradual learning algorithm". *Linguistic Inquiry* 32.1, pp. 45–86. DOI: 10.1162/002438901554586.
- CABAÑERO, Márcia Bueno e Ubiratã Kickhöfel ALVES (2008). "A transferência grafo-fônico-fonológica na produção de seqüências ortográficas 'ng' do inglês (L2): uma abordagem conexcionista". *Revista Virtual de Estudos da Linguagem - ReVEL* 6.11.
- COLLISCHONN, Gisela (1996). "Um estudo da epêntese à luz da teoria da sílaba de Junko Itô (1986)". *Letras de Hoje* 31.2. URL: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/15600>.
- COLLISCHONN, Gisela (1997). "Análise Prosódica da Sílaba em Português". Tese de doutoramento. Porto Alegre: PUCRS.
- COLLISCHONN, Gisela (2000). "A epêntese vocálica no Português do Sul do Brasil: análise variacionista e tratamento pela Teoria da Otimalidade". *Letras de Hoje* 35.1, pp. 285–318. URL: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/14769>.
- COLLISCHONN, Gisela (2002). "A Epêntese Vocálica no Português do Sul do Brasil". In: *Fonologia e variação: recortes do Português Brasileiro*. Editado por Leda BISOL e Cláudia BRESCANCINI. Porto Alegre: EDIPUCRS, pp. 205–230. ISBN: 9788574303000.
- COLLISCHONN, Gisela (2003). "Epêntese vocálica no português do sul do Brasil: variáveis extralinguísticas". *Revista Letras* 61, pp. 285–297. DOI: 10.5380/re1.v61i0.2892. URL: <https://revistas.ufpr.br/letras/article/view/2892>.
- COLLISCHONN, Gisela (2004). "Epêntese vocálica e restrições de acento no português do sul do Brasil". *Signum: estudos da linguagem* 7.1, pp. 61–78.
- CUTLER, Anne, Jacques MEHLER, Dennis NORRIS e Juan SEGUI (1987). "Phoneme identification and the lexicon". *Cognitive Psychology* 19.2, pp. 141–177. DOI: 10.1016/0010-0285(87)90010-7.

- GEWEHR-BORELLA, Sabrina, Márcia Cristina ZIMMER e Ubiratã Kickhöfel ALVES (2011). “Transferências grafo-fônico-fonológicas: uma análise de dados de criança monolíngues (português) e bilíngues (hunsrückisch-português)”. *Gragoatá* 16.30. DOI: 10.22409/gragoata.v16i30.32931.
- HAMANN, Silke e Ilaria E. COLOMBO (2017). “A formal account of the interaction of orthography and perception: English intervocalic consonants borrowed into Italian”. 35, pp. 683–714. ISSN: 0167-806X. DOI: 10.1007/s11049-017-9362-3.
- KELLER, Tatiana (1999). “O fenômeno da epêntese vocálica no português falado em Panambi e Blumenau”. *Cadernos do IL*.
- KIPARSKY, Paul (1982). “From cyclic phonology to lexical phonology”. In: *The Structure of Phonological Representations (part 1)*. Editado por Harry VAN DER HULST e Norval SMITH. Dordrecht: Foris. ISBN: 9783112328071.
- LEGENDRE, Geraldine, Yoshiro MIYATA e Paul SMOLENSKY (1992). “Can Connectionism Contribute to Syntax? Harmonic Grammar, with an Application”. *CLS 26: Papers from the 26th Regional Meeting of the Chicago Linguistic Society*. 26th Regional Meeting of the Chicago Linguistic Society. Editado por Michael ZIOLKOWSKI, Manuela NOSKE e Karen DEATON. Chicago: The Chicago Linguistic Society.
- LEVELT, Willem J. M., Ardi ROELOFS e Antje S. MEYER (1999). “A theory of lexical access in speech production”. *Behavioral and Brain Sciences* 22.1, pp. 1–38. DOI: 10.1017/S0140525X99001776.
- MCCARTHY, John J. e Alan S. PRINCE (1995). *Faithfulness and reduplicative identity*. URL: https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1009&context=linguist_faculty_pubs.
- MCCLELLAND, James L. e Jeffrey L. ELMAN (1986). “The TRACE model of speech perception”. *Cognitive Psychology* 18.1, pp. 1–86. ISSN: 0010-0285. DOI: [https://doi.org/10.1016/0010-0285\(86\)90015-0](https://doi.org/10.1016/0010-0285(86)90015-0).
- MIRANDA, Ana Ruth Moresco (2014). “A fonologia em dados de escrita inicial de crianças brasileiras”. *Linguística* 30, pp. 45–80. URL: http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S2079-312X2014000200004&nrm=iso.
- PARLATO-OLIVEIRA, Erika Maria (2008). “Diversidade, variabilidade e frequência em fonologia: o caso da epêntese vocálica”. *Letras de Hoje* 42.3. URL: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/2797>.
- PRINCE, Alan S. e Paul SMOLENSKY (2004 [1983]). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Wiley. ISBN: 9781405119337.
- QUINTANILHA-AZEVEDO, Roberta (2011). “A epêntese no Português Brasileiro (L2), em segmentos plosivos em codas mediais, por falantes nativos do Espanhol Colombiano (L1): uma análise via Teoria da Otimidade Estocástica e Gramática Harmônica”. Dissertação de mestrado. Pelotas: Universidade Católica de Pelotas.
- QUINTANILHA-AZEVEDO, Roberta (2016). “Formalização fonético-fonológica da interação de restrições na produção e percepção da epêntese em variedades do Português”. Tese de doutorado. Pelotas: Universidade Católica de Pelotas.
- RAUBER, Andréia, Anabela RATO, Giane SANTOS, Denise KLUGE e Marcos FIGUEIREDO (2012). *TP – Testes de Percepção / Tarefas de Treinamento Perceptual*. URL: http://www.worken.com.br/tp/tp_instala.html (acedido em 19 de setembro de 2016).
- SILVEIRA, Francine e Izabel Christine SEARA (2009). “A vogal epentética em encontros consonantais heterossilábicos no português brasileiro: um estudo experimental”. *Revista do GEL* 6.2, pp. 9–35. URL: <https://revistas.gel.org.br/rg/article/view/94>.
- SMOLENSKY, Paul (1996). “On the comprehension/production dilemma in child language”. *Linguistic Inquiry* 27.4, pp. 720–731.
- TRUBETZKOY, Nikolai Sergeievich (1969 [1939]). *Principles of Phonology*. Traduzido do alemão por Christiane A. M. BALTAXE. Berkeley: University of California Press.

ZHOU, Chao (2021). "L2 speech learning of European Portuguese /l/ and /r/ by L1-Mandarin learners: experimental evidence and theoretical modelling". Tese de doutoramento. Lisboa: Universidade de Lisboa.