

**EFEITOS DA DURAÇÃO DO VOZEAMENTO
DA FRICATIVA [Z] NA IDENTIFICAÇÃO, POR
BRASILEIROS, DE PARES MÍNIMOS PRODUZIDOS
POR HISPÂNICOS: INSUMOS PARA A DISCUSSÃO
SOBRE INTELIGIBILIDADE DE FALA ESTRANGEIRA**
EFFECTS OF VOICING LENGTH IN THE PRODUCTION
OF [Z] BY L1-SPANISH SPEAKERS ON THE PERCEPTUAL
IDENTIFICATION OF MINIMAL PAIRS BY BRAZILIAN
LISTENERS: A DISCUSSION ON FOREIGN LANGUAGE
INTELLIGIBILITY

Ubiratã Kickhöfel Alves*
ukalves@gmail.com

Luciene Bassols Brisolará**
lucienebrisolará@furg.br

Leonardo Cláudio da Rosa***
rosa.leoc@gmail.com

Ana Carolina Signor Buske****
anabuske@gmail.com

Este trabalho propõe verificar os efeitos do grau de vozeamento da fricativa [z] produzida por falantes hispânicos para o estabelecimento da distinção entre as categorias 'surdo' e 'sonoro' por parte de ouvintes brasileiros. Para a realização do estudo, coletaram-se dados de fala de seis hispânicos que residiam no Brasil há, no máximo, doze meses. A partir das gravações e das manipulações de diferentes graus de vozeamento (referentes a 0%, 25%, 50%, 75% e 100% da duração total da fricativa), foi elaborada uma tarefa de identificação no *Software* TP (Rauber *et al.* 2012), que foi aplicada a 35 estudantes universitários brasileiros. Os resultados do experimento indicaram que a atribuição do *status* sonoro da fricativa prescinde de uma

O projeto a que este trabalho está vinculado teve o apoio financeiro da Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), por meio do Edital 02/2017 – Programa Pesquisador Gaúcho (PqG) – Processo 17/2551-0001000-0.

* Pesquisador CNPq - Universidade Federal do Rio Grande do Sul, Brasil.

** Universidade Federal do Rio Grande, Brasil.

*** Universidade Federal do Rio Grande do Sul, Brasil.

**** Bolsista PROBIC e FAPERGS - Universidade Federal do Rio Grande do Sul, Brasil.

vibração de pregas vocais que se estenda ao longo de toda a consoante. Além disso, foi possível observar que o padrão 25% de porção de vozeamento é mais dificultoso para a identificação dos aprendizes brasileiros. A análise de resultados inferenciais e descritivos individuais permitiu mostrar que não podemos considerar tal padrão como plenamente equivalente a não vozeamento, e que a transição entre as categorias ‘surda’ e ‘sonora’ varia entre os ouvintes, sem haver um ponto “limiar” comum a todos os participantes.

Palavras-chave: Português como L2. Duração do vozeamento de [z]. Intelligibilidade.

In this study, we aim to investigate the effects of the degree of voicing in the fricative [z] produced by L1 Spanish speakers on the distinction between the categories of ‘voiceless’ and ‘voiced’ consonants by Brazilian judges. Speech data were collected from six L1 Spanish speakers who had been living in Brazil for less than twelve months. From the recordings and manipulations of different degrees of voicing in the fricatives (0%, 25%, 50%, 75% and 100% of the total duration of the fricative), an identification task was built on TP Software (Rauber *et al.* 2012). Thirty-five Brazilian participants did this task. The results indicate that voicing the consonant all the way through was not a necessary condition for the identification of the fricative as voiced. It was also verified that the pattern with voicing throughout 25% of the fricative proved more difficult to identify. Both the inferential analysis and the verification of the data produced by each individual participant showed that this latter pattern cannot be considered to be voiceless in all cases. These results indicate that the amount of voicing necessary for the distinction between the categories of ‘voiced’ and ‘voiceless’ varies among participants, as there is no common threshold shared by all of them.

Keywords: L2 Portuguese. Length of voicing in [z]. Intelligibility.



1. Introdução

De acordo com uma visão tradicional de Fonologia, a distinção entre pares mínimos como ‘caça’ [s] e ‘casa’ [z], no português brasileiro (PB), deve-se ao valor positivo ou negativo do traço [voz], de modo que o segundo membro do par seja produzido com, e o primeiro sem vibrações das pregas vocais. Esse fato tem sido apontado pela literatura (Akerberg 2004; Oliveira 2016; Silveira & Souza 2011; Sobral, Nobre & Freitas 2006) como um fator de dificuldade para o aprendiz hispânico, pois, nesta língua, não há o fonema

/z/; há apenas o fonema /s/.¹ Por essa razão, considerando a sua dificuldade em vibrar as pregas vocais na produção da fricativa sonora, os membros de tais pares poderiam, em princípio, vir a ser indistinguíveis. A não vibração das pregas vocais acarretaria, portanto, uma não distinção entre os membros dos pares mínimos, o que teria implicações no que diz respeito à inteligibilidade local² de palavras do português produzidas por aprendizes hispânicos.

A observação acústica dos dados do português produzido por hispânicos advinda de nossa experiência como pesquisadores da área, em consonância com uma série de estudos sobre desenvolvimento fonético-fonológico de línguas adicionais (Alves & Zimmer 2015; Oliveira 2016; Zimmer & Alves 2012), sugere que a realização da fricativa sonora como em 'casa' vai muito além da mera dicotomia 'presença' vs. 'ausência' de vozeamento. De fato, ao nos dedicarmos à observação dos dados de fala de tais aprendizes, observamos que esses não mudam de um estágio em que simplesmente não vozeiam a fricativa a um estágio seguinte em que a fricativa é plenamente vozeada. O processo de desenvolvimento do aprendiz é tal que ele aprende a vozear gradativamente, de modo que a porção vozeada da fricativa acabe sendo cada vez maior, na medida em que o referido aprendiz vá adquirindo maior proficiência no português (cf. Oliveira 2016). Este fato fonético-fonológico pode ser entendido a partir de uma concepção dinâmica de desenvolvimento de segunda língua (De Bot *et al.* 2007; 2013; De Bot 2017) e de desenvolvimento fonológico de L2 (Alves 2018; Lowie 2011; Silva 2014; Zimmer & Alves 2012), de acordo com a qual a tarefa do aprendiz é aprender a temporalidade do vozeamento referente à fricativa. Em outras palavras, é preciso aprender a orquestrar a temporalidade (Zimmer & Alves 2012; Kupske & Alves 2017), explicável através de modelos de tempo intrínseco da Fonologia (cf. Fowler, 1980), tal como a Fonologia Acústico-Articulatória (Albano 2001), tomado como unidade primitiva tanto da produção (cf. Silva 2014) como da percepção (cf. Perozzo 2017) dos sons.

1 Conforme a Real Academia Española & Asociación de Academias de la Lengua Española (2011), no espanhol, há dois subsistemas, considerando as fricativas: um sistema denominado 'distinguidor', no qual aparecem os fonemas /s/ e /θ/, como nos exemplos 'casa', produzido como ['kasa] e 'caza', como ['kaθa] e o sistema 'seseante', no qual não existe /θ/ e ambos os vocábulos são pronunciados como ['kasa]. O primeiro subsistema inclui a Espanha, com exceção da Andaluzia e das Ilhas Canárias; o segundo, inclui a América Latina, Andaluzia e Ilhas Canárias.

2 Para uma definição do conceito de 'inteligibilidade local' (de uma só palavra) em oposição à 'inteligibilidade global' (cf. Munro & Derwing 2015), veja-se a seção de Contextualização Teórica deste trabalho.

Dado que, sob a perspectiva que rege o presente trabalho, o vozeamento não é uma questão de ‘tudo’ ou ‘nada’, questionamentos a respeito de uma possível zona limiar perceptual entre as categorias funcionais de ‘surdo’-‘sonoro’ fazem-se pertinentes. O presente trabalho, dessa forma, visa a verificar o grau de vozeamento da fricativa produzida por falantes hispânicos para o estabelecimento da distinção entre as categorias ‘surdo’ e ‘sonoro’ por ouvintes brasileiros. Para isso, utilizaram-se dados de seis falantes hispânicos (variedades latino-americanas de espanhol), que se encontravam residindo no Brasil entre três e doze meses. Uma vez que os padrões de vozeamento e de duração vocálica prévia à fricativa variavam tanto intra quanto inter-individualmente, optamos por manipular o grau de vozeamento dos dados. Assim, a partir de estímulos desvozeados, manipulamos as ondas acústicas de produções como ‘ca[s]a’, a fim de obtermos produções com 0%, 25%, 50%, 75% e 100% de vozeamento da consoante focalizada. Essa manipulação se deu pela sobreposição progressiva de trechos vozeados às produções desvozeadas, simulando o princípio de coarticulação de seguimentos, aplicando-se um efeito de *crossfade* entre os trechos. Estando a duração vocálica antecedente e a duração da fricativa controladas, fez-se possível, assim, analisar o papel da duração do vozeamento na fricativa.

Considerando o quadro exposto e tendo como base a visão dinâmica de desenvolvimento linguístico acima expressa, o presente trabalho, a partir do objetivo geral de verificar a identificação, por brasileiros, de fricativas sonoras do PB, em pares mínimos como ‘casa-caça’ ou ‘rosa-roça’, produzidos com diferentes graus de vozeamento nas fricativas dos membros sonoros desses pares, tem como objetivos específicos: (i) apontar os índices de atribuição de status sonoro às fricativas produzidas com diferentes graus de vozeamento (0%, 25%, 50%, 75%, 100% da duração total); (ii) discutir o impacto de tais resultados para o estabelecimento do construto de inteligibilidade de língua estrangeira e para o cenário de ensino de pronúncia do PB; (iii) promover, com base na descrição do material empírico, uma reflexão teórica acerca do caráter dinâmico da produção de sons.

Esperamos, com o presente trabalho, possibilitar a discussão acerca do papel da produção do vozeamento na distinção dos referidos pares mínimos do português brasileiro, de modo a possibilitar a reflexão sobre a necessidade ou não de vozeamento pleno da fricativa para tal distinção. Acreditamos que esta discussão se mostra de grande valia para o âmbito da pedagogia de ensino de pronúncia do PB, sobretudo considerando-se uma perspectiva de ensino dinâmica, voltada ao estabelecimento da inteligibilidade como objetivo maior (Alves 2015; Celce-Murcia *et al.* 2010; Derwing

& Munro 2015a; b; Kupske & Alves 2017; Levis 2005). Além disso, também a partir dos dados empíricos, esperamos fornecer argumentos adicionais para uma concepção de fonologia de caráter mais dinâmico, que, ao abandonar a dicotomia ‘presença’ vs. ‘ausência’ plena de vozeamento, considere a temporalidade dos gestos acústico-articulatórios para a caracterização dos aspectos funcionais das línguas.

2. Contextualização teórica

Nesta seção, discutiremos, sucintamente, as bases teóricas da investigação. Iniciaremos descrevendo as premissas básicas da visão de Língua como Sistema Dinâmico Complexo (De Bot *et al.* 2007; 2013; De Bot 2017), que guia a presente investigação. Na sequência, discutiremos como uma visão dinâmica de língua tem influência sobre a definição do primitivo fonético-fonológico, o que dará espaço para uma reflexão sobre inteligibilidade e ensino de pronúncia na aula de língua estrangeira.

Verspoor (2015) traz à luz as considerações de Weisstein (1999) para definir ‘sistema dinâmico’ como uma maneira de analisar e descrever como um estado (de um sistema) se desenvolve ou se transforma em outro estado ao longo do tempo. Por sua vez, dizer que um sistema é complexo implica afirmar que esse é composto por diversos elementos que interagem entre si e apresentam comportamento emergente, sendo que sua totalidade não corresponde à mera soma das partes (Beckner *et al.* 2009; Larsen-Freeman & Cameron 2008).³ Conforme apontam Lowie e Verspoor (2015, p. 72), o sistema é não somente não linear, mas também complexo; por esse motivo, a aprendizagem é resultado de uma série de aspectos, de modo que “muitos fatores sociais, psicológicos, físicos e ambientais estão envolvidos, e tais fatores estão interconectados”. De fato, conforme explicam De Bot, Lowie e Verspoor (2007, p. 8), é preciso considerar que “pequenas modificações em uma das variáveis pode levar a efeitos enormes, enquanto que perturbações maiores, por sua vez, podem ser absorvidas pelo sistema como um todo sem causar muitas modificações”. Em outras palavras, não há uma relação de linearidade entre o que poderia ser considerada a “alteração inicial” de

3 Dado a necessidade de delimitação deste estudo, na presente seção estaremos preocupados em descrever apenas os aspectos-chave da visão dinâmica que deverão ser aprofundados na análise do trabalho. Para uma visão introdutória sobre o modelo, aconselhamos a leitura de uma série de textos de apresentação da Teoria, como Larsen-Freeman (1997; 2014; 2015a; 2015b; 2017), Larsen-Freeman & Cameron (2008), De Bot, Lowie & Verspoor (2007; 2011) De Bot *et al.* (2013), Beckner *et al.* (2009).

uma das variáveis e o seu resultado no sistema inteiro. Essa concepção faz com que os pesquisadores e estudiosos busquem considerar as variáveis envolvidas no sistema, além das possíveis alterações que as mesmas estão sujeitas a sofrer e, por sua vez, acarretar.

Beckner *et al.* (2009), ao se referirem à linguagem como sistemas adaptativos complexos, afirmam que tais sistemas apresentam sete propriedades principais, tais como (a) controle distribuído e emergência coletiva; (b) diversidade intrínseca; (c) dinâmica perceptual; (d) adaptação através da amplificação e competição de fatores; (e) não linearidade e transições entre fases; (f) sensibilidade e dependência da estrutura de rede; (g) mudança de caráter local. Ao se pensar a linguagem como um Sistema Adaptativo Complexo, concebe-se que ela está inserida em um meio social, do qual sofre e sobre o qual exerce efeitos. Além disso, sendo um sistema complexo, a linguagem tem, também, caráter multimodal. Isso significa que o próprio desenvolvimento do componente fonético-fonológico não pode ser resumido, unicamente, à percepção e à produção dos sons da cadeia da fala. A comunicação se desenvolve não somente através do sistema de sons, mas, também, através dos gestos corporais, dos estímulos visuais, do próprio código escrito, enfim, de uma concatenação de fatores cujos efeitos individuais ainda se mostram difíceis, e muito provavelmente, sejam realmente impossíveis de determinar, uma vez que o desenvolvimento do conhecimento linguístico tem por base todas essas formas de estímulo.

Com o apogeu da visão complexa, tem sido amplamente defendida a relevância de dados individuais dos participantes, que podem vir a expressar padrões desenvolvimentais, no âmbito longitudinal, bastante diferentes daqueles verificáveis a partir da média do grupo. A partir de tais considerações, podemos encontrar, entre os proponentes do modelo (*cf.* Byrne 2013; De Bot, Lowie & Verspoor 2007; Lowie 2017; Lowie & Verspoor 2015; Verspoor 2015, dentre outros), posicionamentos bastante extremos, que visam a erradicar a descrição dos padrões de grupo das análises. Segundo Lowie e Verspoor (2015, p. 84),

o desenvolvimento linguístico é um processo inerentemente individual e dinâmico; por isso, não pode haver alguma expectativa lógica de que o padrão encontrado nas generalizações referentes ao nível do grupo possa refletir o verdadeiro desenvolvimento do aprendiz individual.

Esse argumento, de fato, constitui consenso entre os seguidores da teoria; muitos autores têm optado por abolir completamente qualquer trata-

mento empírico que conjugue as verificações dos dados do indivíduo com os do grande grupo.

Em meio a este debate acirrado, por outro lado, encontramos pesquisadores que apresentam uma posição mais conciliadora, tanto em seus textos teóricos quanto nas suas próprias análises empíricas. Em seu texto de 2017, Larsen-Freeman nos diz que “podemos fazer afirmações no nível do grupo, mas não podemos assumir que essas necessariamente se aplicam a indivíduos” (*op. cit.*, 27). Sob nossa leitura, tal afirmação simboliza um abrir de portas para a verificação justamente dos impactos de ambas as formas de descrição dos dados, ao termos sempre o cuidado de reconhecer que a verificação do grupo não apresenta um caráter generalizador dos indivíduos. Em termos de análise, trabalhos como os de De Leeuw, Mennen e Scobbie (2012), Nitta e Baba (2015) e Lima Jr. (2016a; b) têm abordado ambos os tipos de verificação para o entendimento do processo de desenvolvimento linguístico, contribuindo para a discussão sobre as semelhanças e diferenças entre estes dois tipos de descrição. Em meio ao debate instaurado, o presente trabalho visa a seguir esta mesma linha metodológico-analítica, ao promover ambos os tipos de descrição de dados.

No que concerne ao *status* da variação linguística, reafirmamos que, em uma perspectiva que focaliza o processo em vez do produto, a variabilidade dos dados, tanto no nível do indivíduo quanto entre os sujeitos, constitui material empírico revelador para o entendimento da caminhada desenvolvimental (De Bot, Lowie & Verspoor 2007; Lowie & Verspoor 2015). Ao pensarmos justamente na variação enquanto fenômeno que caracteriza o desenvolvimento linguístico, devemos, também, discutir os primitivos fonético-fonológicos condizentes não somente com a variabilidade, mas com a própria noção dinâmica de mudança ao longo do tempo. Isso implica trazer, indiscutivelmente, a temporalidade para dentro de tal primitivo. Tomando-se a fala como ação, a partir da incorporação das noções de tempo e movimento, tomamos o gesto articulatório (Browman & Goldstein 1986; 1992), ou ainda o gesto acústico-articulatório (Albano, 2001), como unidades primitivas da fonologia. Tais unidades são denominadas, por Fowler (1980), como unidades “de tempo intrínseco”, pelo fato de tais primitivos incorporarem a questão da temporalidade entre os articuladores envolvidos na produção da fala. No que diz respeito aos modelos de percepção e produção de L2, o *Perceptual Assimilation Model-L2*, doravante PAM-L2, (Best & Tyler 2007), a partir de um primitivo gestual de percepção, parte do pressuposto de que os sons da L2 são assimilados aos da L1. Perozzo (2017a; 2017b) propõe alterações à proposta do PAM-L2 a

partir das esferas cognitiva, linguística e filosófica. À luz da sua proposta, a interconexão entre essas três esferas impediria conflitos epistemológicos no modelo. O autor, dessa forma, investe em uma argumentação a favor de unidades representacionais, defendendo um primitivo acústico-articulatório (cf. Albano 2001), bem como uma visão Realista Indireta (Jackson 1977; 2010)⁴, em oposição à visão Realista Direta de Fowler (1986), Best (1995) e Best e Tyler (2007), para acomodar tais unidades representacionais dentro de uma escola filosófica que abarque tal representação.

No presente trabalho, palavras do português como L2 serão identificadas por ouvintes brasileiros como um ou outro membro de um par mínimo, de modo que possamos adotar a proposta de Perozzo (2017a; 2017b) para a percepção, uma vez que os sons produzidos pelos aprendizes hispânicos serão atribuídos a alguma categoria do PB. Esse “enquadramento” pode ser fundamental para o estabelecimento da inteligibilidade.⁵ Conforme apontam Alves e Silva (2016, p. 125), “desafios conceptuais futuros interessantes implicarão reflexões teóricas que conjuguem as noções de percepção e inteligibilidade, de modo que, ao mesmo tempo em que se esclareçam as diferenças entre esses dois entes teóricos, seja definida a relação intrínseca entre esses dois construtos”. De fato, ainda que a percepção individual dos sons contribua para o estabelecimento da inteligibilidade, essa constitui um construto muitíssimo maior do que a percepção das unidades acústico-articulatórias. A noção de inteligibilidade é estabelecida através de muitíssimos outros fatores, tais como o contexto da conversação, o ambiente da fala e diferenças individuais dos membros da comunicação (Alves 2015; Alves & Albuquerque 2017; Schwartzhaupt 2015), que exercem ação combinada entre si, e cujos efeitos devem ser considerados de forma integrada, como os elementos de um Sistema Adaptativo Complexo (Beckner *et al.* 2009; Larsen-Freeman & Cameron 2008), conforme discutiremos adiante.

Munro e Derwing (1995) definem ‘inteligibilidade’ como o grau de extensão através do qual a mensagem em fala estrangeira é entendida pelo

4 De acordo com a visão do Realismo Indireto, estamos em contato com a realidade através de objetos intermediários, tais como as representações mentais (Dancy 1985). Para maiores detalhes sobre essa escola, aconselhamos a leitura de Perozzo (2017a).

5 Não cabe a este trabalho a tarefa de refletir sobre as diferentes definições, encontradas na literatura, para o termo ‘inteligibilidade’, sobretudo por considerarmos que tal termo, pelo menos em sua caracterização tradicional, ainda se mostra desvinculado de uma concepção de língua que o sustente. Em sua Tese de Doutorado, Albuquerque (em elaboração) visa a caracterizar tal construto a partir de uma perspectiva dinâmica. Por ora, apresentamos a caracterização tradicional do termo expressa por Munro & Derwing (1995), ainda que reconheçamos as carências de tal caracterização.

ouvinte, ou seja, o grau de entendimento possibilitado pela fala em L2. De acordo com os autores (Munro & Derwing 2015), a inteligibilidade pode ser tanto ‘local’ (referente a unidades menores da fala, como palavras), ou ‘global’ (com unidades sintáticas maiores, envolvendo um alto grau de contexto informacional). Sobretudo considerando-se a dimensão global, as variáveis supracitadas devem ser consideradas, para que possamos dar o salto entre a percepção e a definição de tal construto. A inteligibilidade ‘global’ depende, indiscutivelmente, da ‘local’, que pode ser investigada, inclusive, através da tarefa de identificação proposta neste trabalho. Ao identificarem qual membro do par mínimo está sendo realizado, ao apontarem se o som ouvido é [s] e [z], os participantes brasileiros do presente estudo estarão verificando o grau de inteligibilidade da fala estrangeira no nível da palavra, em caráter local, o que poderá contribuir para um maior ou menor entendimento de toda a fala.

Finalmente, ao abordarmos a questão da inteligibilidade, não podemos ignorar o papel do componente fonético-fonológico no ensino de português como língua estrangeira. No que diz respeito ao cenário de ensino de pronúncia nas salas de aula de língua estrangeira, Levis (2005) explica que as práticas do professor podem ser divididas em dois diferentes tipos de postura: (i) o princípio da fala nativa, de acordo com o qual o aprendiz deveria soar como um falante “nativo” (o que constitui, diga-se de passagem, uma tarefa praticamente impossível; aqui, também, cabem todas as “ressalvas” ao que seria o tal “falante nativo”); (ii) o princípio da inteligibilidade, a partir do qual se espera que o aprendiz estrangeiro desenvolva padrões fonético-fonológicos que o permitam comunicar-se sem dificuldades, ainda que com um acento que apresente marcas de sua língua materna. Julgamos que o primeiro princípio se mostra inviável dentro de uma perspectiva complexa, uma vez que, ao considerarmos a diversidade linguística como um dos aspectos fundamentais dessa visão, a própria tarefa de determinar quem seria esse “falante nativo” já se mostra fragilizada. O objetivo do ensino de pronúncia não deve ser erradicar o acento em L1, mas, sim, contribuir para uma fala mais inteligível. Mais do que isso, o ensino do componente fonético-fonológico, a partir da noção dinâmica e complexa de língua, deve ser integrado aos outros componentes linguísticos.⁶ De fato, “se isolarmos os componentes artificialmente, perdemos a essência dos fenômenos que queremos descrever” (Larsen-Freeman 2017, p. 29). Tal

6 Para um exemplo de manual de pronúncia do português brasileiro que segue tal premissa, veja-se Alves, Brisolara e Perozzo (2017).

consideração se opõe à tradicional concepção de ensino de pronúncia, que tende a ser caracterizada, unicamente, pela mera tarefa de ‘ouvir’ e ‘repetir’ estruturas isoladas, de modo mecanicista.

Acreditamos ter deixado claras as premissas do presente trabalho. Com base em uma visão dinâmica de língua, pensamos em uma metodologia conjugada entre análise individual e de grupo, e em uma unidade primitiva de fonologia baseada na ação, no movimento e na temporalidade, tal como o gesto acústico-articulatório (Albano 2001). A partir desse primitivo dinâmico, podemos dar os primeiros passos para recharacterizar a relação entre percepção-produção e inteligibilidade em L2, e, inclusive, pensar em um ensino de pronúncia voltado mais para as funções comunicativas e pela integração com outros componentes linguísticos do que pela preconização do chamado “padrão nativo”.

3. Método

Construímos o *corpus* do estudo a partir de três etapas: captação de gravações, seleção de dados acusticamente semelhantes e, por fim, manipulação acústica dos dados selecionados. Na primeira etapa, recolhemos gravações em áudio da fala de seis indivíduos latino-americanos nativos do espanhol, falantes de inglês como L2, aprendizes de português como L3⁷ e com tempo de residência na cidade de Rio Grande – RS variando de três a doze meses. Todos começaram a ter contato com a língua portuguesa com sua chegada ao Brasil. Na Tabela 1, apresentam-se informações a respeito da origem e da experiência linguística dos participantes que serviram de locutores:

7 No presente trabalho, utilizamos os termos ‘L2’ e ‘L3’ para nos referirmos à ordem cronológica de aprendizado das línguas adicionais dos participantes. Considerando-se a teoria dinâmica que embasa o presente trabalho, esta distinção é importante, sobretudo por considerarmos que uma nova língua adicional (no caso, a L3) pode ser afetada não somente pela L1, mas, também, pela L2 (Pereyron 2017). Por tal motivo, tínhamos, inicialmente, a intenção de coletar dados de participantes que falassem português como L2, não como L3. Entretanto, uma vez que os participantes em questão eram alunos de cursos de pós-graduação, em intercâmbio no Brasil, verificamos que todos já apresentavam um nível pelo menos intermediário de língua inglesa, o que permitiu um controle no sentido de que todos os participantes falavam português brasileiro como L3. Ainda que os participantes hispânicos de nosso estudo falassem inglês, língua essa caracterizada pela presença dos fonemas /s/ e /z/ em seu inventário, os dados de produção em PB evidenciaram que a fricativa [z] da L3 era realizada praticamente sem vozeamento.

Tabela 1. Dados dos aprendizes hispânicos residentes no Brasil

	País de origem	Idade de início de estudo do PB	Meses de residência no Brasil
Falante 1	Venezuela	32	3 meses
Falante 2	Peru	24	12 meses
Falante 3	Honduras	23	3 meses
Falante 4	Colômbia	23	3 meses
Falante 5	Equador	27	9 meses
Falante 6	México	23	3 meses

O instrumento aplicado a tais informantes consistia na leitura em voz alta de frases contendo uma palavra na qual o vozeamento na consoante fricativa fosse decisivo para identificação lexical. Presentes no instrumento estavam pares tais como ‘asa/assa’, ‘casa/caça’, ‘rosa/roça’, ‘pesa/peça’, ‘rasa/raça’ e ‘tese/tece’. Com a verificação acústica destas produções, confirmamos a premissa, expressa na Introdução, de que o vozeamento da fricativa se dá de modo gradual, sem implicar uma mudança abrupta entre a total ausência ou vozeamento pleno ao longo do segmento consonantal.

A partir desses dados, realizamos a análise acústica das gravações obtidas. Utilizamos, para tanto, o *software* livre *Praat* (Boersma & Weenink 2018). Essa análise nos permitiu a execução da segunda etapa, a partir da qual selecionamos gravações semelhantes focalizando três aspectos acústicos: a duração da vogal anterior à fricativa⁸, a duração absoluta da consoante fricativa e a duração do vozeamento na consoante fricativa. Obtivemos, a partir dessa seleção, 12 gravações referentes aos pares supracitados, havendo um informante diferente para cada par.

De posse dessas gravações, partimos para a etapa de manipulação acústica. Utilizamos para tal o *software* livre *Audacity*. A manipulação se deu em dois procedimentos distintos: controle dos dois primeiros aspectos acústicos observados na etapa anterior (duração da vogal anterior à fricativa e duração absoluta da consoante fricativa) e variação da duração

8 Conforme será feito na seção que encerra este trabalho, reconhecemos como uma limitação do presente artigo a não-inclusão da verificação dos efeitos da duração vocálica precedente à fricativa no presente artigo. Tal não inclusão se justifica não somente pela necessidade de delimitação do número de páginas deste relatório de pesquisa, mas, também, porque a inclusão desta nova variável implicaria um aumento considerável na extensão da tarefa perceptual. Por esse motivo, decidimos verificar o papel da duração vocálica em um segundo experimento, posterior ao aqui retratado, cujos resultados serão expressos em artigo futuro.

do vozeamento na consoante fricativa. No que se refere ao controle, fixamos esses dois aspectos através da manipulação da contraparte sonora dos pares mínimos. Preservamos, para a manutenção da naturalidade dessas gravações, a duração média do período da onda sonora – o que resultou em durações ligeiramente distintas. Assim, obtivemos os estímulos descritos na Tabela 2:

Tabela 2. Dados acústicos dos estímulos resultantes do procedimento de controle

Identificador do estímulo	Duração da vogal anterior à fricativa (ms)	Duração absoluta da fricativa (ms)
informante1_assa	226,0	141,0
informante2_caça	222,0	146,0
informante3_peça	218,0	143,0
informante4_raça	219,0	143,0
informante5_roça	218,0	145,0
informante6_tece	220,0	145,0

No que se refere à variação da duração do vozeamento na consoante fricativa, escolhemos os valores de 0%, 25%, 50%, 75% e 100% de vozeamento referente à duração total da fricativa. Para gerar estímulos com essas características acústicas, manipulamos aqueles já descritos na Tabela 2 através do seguinte procedimento:

1. Silenciamos um trecho da consoante desvozeada, sendo o início deste trecho coincidente com o início da onda consonantal desvozeada e a duração do silêncio determinada pelo cálculo de 15% do valor final do vozeamento desejado.
2. Sobreponemos a esse silêncio o trecho sonoro vozeado da gravação da palavra correspondente, tendo este trecho a duração do valor final do vozeamento desejado mais 25% da vogal anterior.
3. Para fins de coarticulação, enfraquecemos progressivamente as bordas da sobreposição, de maneira a simular a transição acústico-articulatória referente à perda de vozeamento.

Geramos, para cada estímulo da Tabela 2, outros quatro, descritos em função do vozeamento absoluto da consoante fricativa na Tabela 3:

Tabela 3. Dados acústicos dos estímulos resultantes do procedimento de variação do vozeamento da consoante fricativa

Estímulos	Vozeamento absoluto da fricativa em milissegundos				
	0% (constantes na Tabela 2)	25%	50%	75%	100%
informante1_assa/asa	0,0	35,3	70,5	105,8	141,0
informante2_caça/casa	0,0	36,5	73,0	109,5	146,0
informante3_peça/pesa	0,0	35,8	71,5	107,3	143,0
informante4_raça/rasa	0,0	35,8	71,5	107,3	143,0
informante5_roça/rosa	0,0	36,3	72,5	108,8	145,0
informante6_tece/tese	0,0	36,3	72,5	108,8	145,0

A partir dos dados apresentados na Tabela 3, obtivemos 30 estímulos que compuseram nosso *corpus*, manipulados segundo o exemplo apresentado na Figura 1:

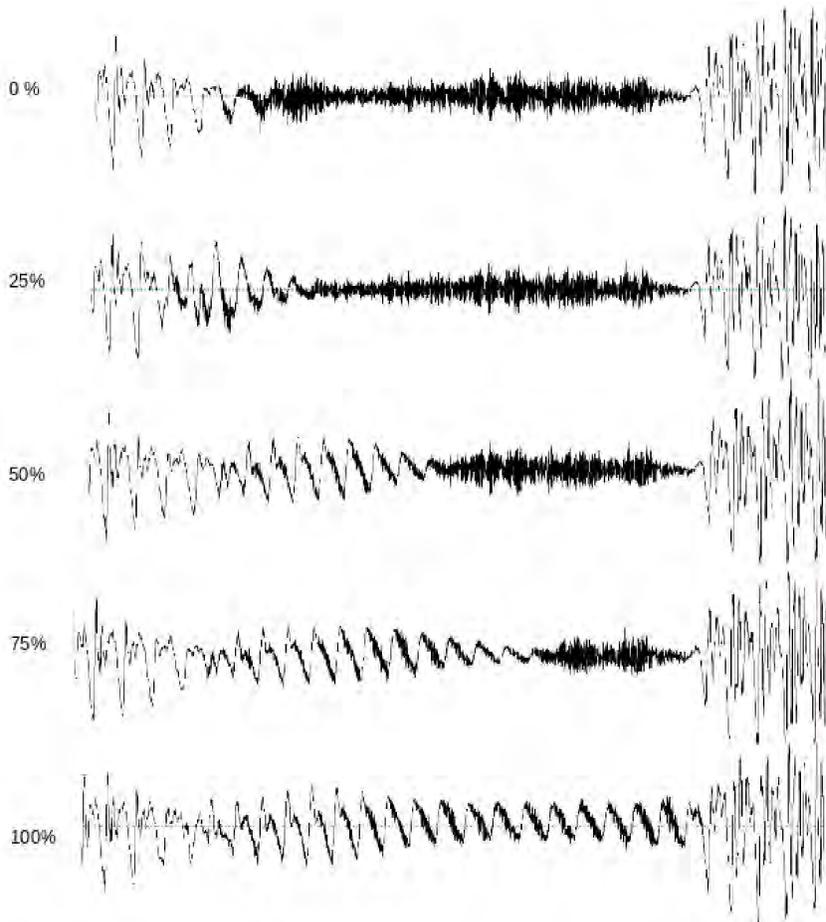


Figura 1. Formas de onda dos estímulos manipulados do informante 5 (roça/rosa)

De posse de tais estímulos, construímos um teste de identificação no *software* livre TP (Rauber *et al.* 2012). Os informantes brasileiros submetidos a este teste respondiam à seguinte pergunta: A palavra principal da frase “Diga _____ bem” tem o som de [s] ou de [z]? Cada estímulo foi repetido três vezes de forma aleatória ao longo do teste, totalizando 90 identificações, com um intervalo de descanso a cada 30. Cabe salientar que os participantes só puderam ouvir cada estímulo uma vez.

O teste de identificação foi aplicado a 35 nativos de português brasileiro, não monolíngues, sem experiência com a língua espanhola (con-

forme as informações de um questionário preenchido pelos participantes antes da coleta)⁹, com escolaridade universitária em andamento, idade universitária (entre 18 e 24 anos) e estabelecidos na região metropolitana de Porto Alegre – RS.

Utilizamos para a coleta dois computadores diferentes: um computador de mesa *Dell OptiPlex 3040* com sistema operacional *Windows 7* e um *notebook Samsung Expert X41* com sistema operacional *Windows 10*. Os equipamentos de reprodução de áudio foram um *headset HyperX Cloud Stinger* e um *headset C3TECH Raptor MI-2870*. A tarefa foi aplicada individualmente aos 35 participantes do estudo.

4. Descrição dos resultados

Ao considerarmos uma visão dinâmica de língua, uma discussão bastante pertinente diz respeito à verificação dos dados a partir de análises individuais vs. de grupo, bem como das análises descritivas vs. transversais. No presente trabalho, desenvolvido a partir de tal perspectiva, adotamos uma postura conciliadora de descrição de dados, de acordo com a qual os dados individuais e do grande grupo, bem como os cuidados estatísticos descritivos e inferenciais, são considerados como complementares em uma abordagem de dados. Por este motivo, iniciaremos com a descrição dos dados por participantes e, logo em seguida, complementaremos a análise individual com dados de estatística descritiva e inferencial acerca de todo o grupo.

Cabe salientar que cada uma das categorias de vozeamento manipuladas (0%, 25%, 50%, 75% e 100%) contava com 18 *tokens* a serem identificados por cada participante. Partindo-se da premissa de que, em uma concepção dinâmica, as diferenças individuais de cada aprendiz devem ser levadas em consideração (*cf.* Larsen-Freeman 2017; Lowie 2017; Lowie & Verspoor 2015), na Tabela 4, apresentamos os dados de identificação de cada um dos 35 participantes do estudo, organizados de acordo com o grau de vozeamento na fricativa:

⁹ Conforme informações obtidas neste questionário, a grande maioria dos participantes considerou-se ser falante de inglês como L2. Considerando que o inglês se caracteriza por uma língua que também apresenta os fonemas /s/ e /z/ em seu inventário, acreditamos que tal fato pode ter representado uma variável a ser considerada em estudos futuros, uma vez que a experiência bilingue pode vir a exercer efeito sobre os resultados. Destacamos, outrossim, a dificuldade de se encontrarem participantes monolíngues em ambientes universitários, o que justifica a inclusão de participantes bilíngues.

Tabela 4. Dados de identificação da fricativa como [z] por parte dos brasileiros

Participantes	z_0	z_25	z_50	z_75	z_100
1	2	6	16	18	18
2	3	14	18	17	18
3	0	11	17	16	18
4	0	5	16	18	18
5	1	5	15	18	18
6	1	4	14	18	18
7	2	12	18	18	18
8	2	6	14	17	18
9	4	8	18	18	18
10	2	7	16	18	18
11	2	9	14	17	17
12	0	2	16	17	18
13	1	9	17	17	18
14	2	4	17	17	17
15	2	9	13	18	18
16	6	8	14	15	17
17	0	0	16	18	18
18	4	9	14	17	18
19	2	7	15	18	18
20	5	8	13	14	14
21	3	11	17	18	18
22	1	8	14	18	18
23	3	8	17	18	18
24	3	5	12	18	18
25	1	5	16	17	18
26	1	4	13	16	18
27	2	11	16	18	18
28	1	4	15	18	18
29	3	6	15	18	17
30	4	11	18	18	18
31	5	7	12	14	15

32	5	5	16	18	17
33	2	2	11	17	17
34	0	3	12	18	18
35	0	5	14	13	18

Conforme aponta a Tabela 4, o grau de atribuição do status sonoro cresce gradativamente à medida que a proporção de vozeamento na fricativa vai aumentando, sendo que, com exceção dos participantes 29 e 32 (em que as fricativas com 75% de vozeamento foram identificadas como sonoras em 100% dos casos), o índice de atribuição de vozeamento tende a ser maior nos estímulos com 100% de vozeamento do que nas instâncias de estímulos que exibem outros padrões. Além disso, dos 35 participantes, 27 atribuíram status sonoro a todos os 18 estímulos que exibiam vozeamento pleno ao longo da fricativa.

Estímulos com 75% também foram identificados predominantemente como sonoros. De fato, 20 dos 35 participantes também atribuíram *status* sonoro a todas as instâncias de estímulo com tal grau de vozeamento, e nove outros participantes atribuíram status sonoro a 17 dos 18 *tokens* escutados. Apenas quatro participantes, portanto, identificaram como surdas mais de duas consoantes com este padrão de vozeamento, e, salvo os participantes 2, 3 e 35, a grande maioria dos participantes atribuiu índices maiores de identificação de [z] a este padrão do que o que apresentava a fricativa com 50% de vozeamento.

No que diz respeito ao outro extremo, referente às fricativas com 0% de vozeamento, verificamos baixíssimos índices de identificação da fricativa como sonora. De fato, seis dos participantes não atribuíram *status* de sonoro a nenhum dos estímulos escutados com este padrão, sete identificaram como sonora apenas um dos estímulos e dez atribuíram status sonoro a duas instâncias. Apenas um participante (participante 16) atribuiu *status* sonoro a mais de cinco *tokens*; ainda assim, o grau de identificação das consoantes como sonoras, por tal participante, vai crescendo gradativamente na medida em que o grau de vozeamento ao longo da fricativa é aumentado.

O grau de variabilidade maior no que diz respeito à identificação das consoantes como surda ou sonora ocorre, justamente, nas consoantes que exibem 25% e 50% de vozeamento ao longo da fricativa. No que diz respeito às consoantes com 50% de vozeamento, quatro participantes já identificam todos os 18 exemplares da consoante com tal padrão de vozeamento como [z]. Além disso, o número mínimo de identificações de tal padrão como

sonoro foi de 11 (participante 33), o que sugere que, no que diz respeito a este padrão, mais da metade dos estímulos já são identificados como sonoros. Tal quadro, por sua vez, não se mostra muito claro no que diz respeito às consoantes com 25% de vozeamento: conforme apontam as células preenchidas em negrito na Tabela 4, dez dos 35 participantes atribuíram status sonoro a pelo menos nove dos 18 estímulos deste padrão; 21 julgaram ser sonoros entre quatro e oito estímulos; três consideraram entre dois e três estímulos sonoros, e apenas o participante 17 atribuiu status surdo aos 18 estímulos.

Conforme aponta a Tabela 4, o padrão com 25% foi o que apresentou maiores índices de variabilidade na atribuição do status surdo-sonoro. De fato, encontramos desde 0% de identificação da consoante como sonora (participante 17) até o índice de 14 identificações de tal consoante como [z] (participante 2). Tal variabilidade fica evidente na Tabela 5, referente aos valores de estatística descritiva do grupo para cada padrão:

Tabela 5. Estatística descritiva referente aos participantes brasileiros

	N	Mínimo	Máximo	Média	Mediana	Desvio Padrão
Z_0%	35	,00	6,00	2,1429	2	1,62956
Z_25%	35	,00	14,00	6,8000	7	3,14175
Z_50%	35	11,00	18,00	15,1143	15	1,92157
Z_75%	35	13,00	18,00	17,1714	18	1,31699
Z_100%	35	14,00	18,00	17,6286	18	,87735
N válido	35					

Conforme podemos ver na Tabela 5, o padrão com manipulação de 25% de vozeamento ao longo da fricativa é aquele que apresenta os maiores índices de desvio-padrão. Tal constatação deve ser vista como bastante interessante, pois, sobretudo a partir de uma concepção dinâmica de língua, como a que rege o presente estudo, a variabilidade é inerente aos sistemas linguísticos (De Bot 2015; Larsen-Freeman 2017; Lowie 2017; Lowie & Verspoor 2017). Notamos, dessa forma, que, se a partir do padrão de 50% de vozeamento a fricativa já tende a ser identificada como sonora, na duração de 25% ela é predominante sonora para alguns dos participantes, e predominantemente surda para outros. Trata-se, justamente, de uma região nebulosa em termos de percepção. Possivelmente, na atribuição de *status* surdo/sonoro frente a este padrão, os ouvintes estejam lançando mão de outras pistas acústico-articulatórias tomadas em conjunto com a dura-

ção do vozeamento, usando tais pistas de modo conjugado para a tomada de decisão. A esse respeito discutiremos na seção seguinte.

A partir dos resultados descritivos expressos na Tabela 5, conduzimos uma análise inferencial, através do *Software SPSS-Version 21* (IBM Corp. 2012). Para verificarmos diferenças no que diz respeito ao grau de identificação de [z] entre os padrões testados (0%, 25%, 50%, 75% e 100%), realizamos o Teste de Friedman¹⁰, que apontou uma diferença significativa ($X^2(4) = 131,498$, $p = ,000$) entre as identificações de [z] atribuídas aos diferentes graus de vozeamento. Os testes *post-hoc* de Dunn-Bonferroni¹¹ apontaram uma diferença significativa nos índices de identificação de [z] atribuídos aos padrões de 25% e 50% ($p = ,008$), mas não entre 0% e 25% ($p = ,156$), entre 75% e 100% ($p = ,091$) e entre 75% e 100% ($p = 1,00$). Verificamos, através deste resultado, que realmente o padrão 25% se diferencia do 50%: se, com a metade de vozeamento ao longo da fricativa, a grande maioria dos estímulos tende a ser atribuída à consoante sonora, no caso de 25%, não encontramos um quadro plenamente definido. Ainda que muitos ouvintes tendam a identificar tal padrão como surdo, dez aprendizes o identificam predominante como sonoro, o que configura a alta variabilidade descrita na Tabela 5, em que foram apresentados os desvios-padrão para cada grau de vozeamento.

5. Discussão e considerações finais

O presente estudo partiu de nossa experiência com dados de desenvolvimento de L2, sobretudo referentes à produção de fricativas sonoras por hispânicos. Em tais dados, verificávamos que o processo de desenvolvimento referente ao vozeamento da fricativa apresentava caráter gradual, de modo que a produção de vozeamento fosse crescendo gradativamente ao longo da duração total da fricativa. Com base nestes dados, indagávamo-nos sobre o caráter funcional deste “vozeamento parcial” entre os ouvintes brasileiros.

Os resultados do experimento aqui desenvolvido nos mostram que a atribuição do *status* sonoro da fricativa prescinde de uma vibração de pregas vocais que se estenda ao longo de toda a fricativa. Permitem-nos,

10 O Teste de Friedman é um teste não paramétrico que permite a análise de medidas repetidas. Através deste teste podemos verificar se há diferenças em uma variável dependente ordinal considerando “três ou mais momentos temporais ou condições distintas” (Martins 2011, p. 197).

11 A Correção de Dunn-Bonferroni é uma correção que permite ajustar o nível de significância, controlando a “*taxa de erro do Tipo I global* quando vários testes de significância são executados” (Field 2009, p. 644).

portanto, contemplar o primeiro objetivo deste estudo, que visava justamente a verificar como se dava a atribuição do status surdo/sonoro a fricativas cujos vozeamentos se estendiam ao longo de diferentes proporções da duração total (0%, 25%, 50%, 75%, 100%). Os resultados apresentados na seção anterior nos mostram que a presença de vozeamento ao longo de 75% da consoante leva, também, a um alto grau de identificação de tal segmento como [z]. De fato, inclusive a proporção de 50% da fricativa levou, em todos os participantes do estudo, a um número maior de atribuições de tal segmento à categoria de [z] do que [s] (sendo 11/18 o menor índice de atribuição de tal padrão a [z], por parte do participante 33).

É justamente no padrão de 25% que começamos a encontrar resultados que diferem entre participantes: enquanto 25 dos participantes preferem atribuir a maioria dos *tokens* ouvidos à categoria de [s], dez dos participantes preferem atribuir status de sonoro à maioria dos 18 estímulos ouvidos. Frente a este resultado descritivo, conseguimos verificar que este padrão se mostra mais dificultoso para os aprendizes. A partir de tais índices, cabem, ainda, considerações a respeito da metodologia de análise de dados empregada. Com base nos resultados dos testes *post-hoc* da análise estatística inferencial, verificamos não haver diferença estatisticamente significativa entre os padrões 0% e 25%. Entretanto, a partir da visão complexa aqui sustentada, defendemos uma verificação de dados híbrida, que promova uma complementação entre resultados inferenciais e descritivos individuais. Tal decisão metodológica se mostrou fundamental para o entendimento do caráter “nebuloso” do padrão de 25% de vozeamento: a partir da observação dos dados individuais, verificamos que dez dos 35 participantes tendiam a atribuir status sonoro também a esse padrão de manipulação. Dessa forma, a análise integrada aqui realizada permite mostrar que não podemos considerar tal padrão como equivalente a não vozeamento, e que a transição entre as categorias ‘surda’ e ‘sonora’ varia entre os aprendizes, sem haver um ponto “limiar” comum a todos os participantes. Mais do que isso, considerando-se a gradiência da noção dinâmica e dos próprios primitivos de fonologia aqui defendidos, a transição entre as categorias funcionais parece, no nível individual, ser também gradiente: parte-se da atribuição ao *status* de surdo e, na medida em que o grau de vozeamento for aumentando, o ouvinte adentra regiões de alto índice de variabilidade, de modo a ir em direção à categoria de sonoro na medida em que o vozeamento for

ainda mais aumentado.¹² Tal região nebulosa, verificável, em nossos dados, na manipulação de vozeamento ao longo de 25% da fricativa, só se faz possível justamente a partir das decisões metodológicas aqui tomadas, fundamentadas na noção dinâmica de língua: os dados dos participantes nos mostraram que a categoria 25% é, efetivamente, uma região de dúvida para os participantes. Esse fato deve ser ressaltado, sobretudo considerando-se a concepção aqui advogada, uma vez que não podemos afirmar que tal padrão corresponderia a 0% ou 50%, padrões de manipulação esses cujas identificações já se mostram um pouco mais estáveis entre os participantes. Finalmente, salientamos, novamente, que a atribuição praticamente categórica do status de [z] ocorre muito antes de o estímulo apresentar 100% de vozeamento. Em estudos futuros, pretendemos aumentar o número de padrões de manipulação de vozeamento, para expressar de forma ainda mais detalhada essa gradiência.

Os resultados de identificação apontados pelos falantes nativos de PB, ao corresponderem ao que Munro e Derwing (2015) denominam de ‘inteligibilidade local’, exercem efeitos, também, à inteligibilidade global, e têm consequências na própria metodologia de ensino de línguas, o que nos leva a abordar o segundo objetivo deste estudo. Partindo-se de uma perspectiva de ensino de pronúncia voltada à inteligibilidade (*‘intelligibility principle’*, cf. Levis 2005), podemos concluir que, ainda que o ensino de pares mínimos como ‘casa-caça’ se faça necessário, em termos de desenvolvimento fonético-fonológico, o aprendiz passa por uma gradualidade no que diz respeito ao vozeamento da fricativa, e que tal gradualidade se manifesta, inclusive, na identificação das categorias funcionais de ‘surdo’ e ‘sonoro’ nas percepções dos ouvintes nativos de português. Ainda que os efeitos de tal gradualidade na produção precisem ser verificados no que diz respeito aos índices de compreensibilidade¹³ de tais padrões, o que corresponde, também, a um interessante tema de trabalho futuro, os dados aqui apresentados já exigem do professor de fonética do PB uma postura mais compreensiva referente à trajetória de desenvolvimento de duração de vozeamento do aluno, pois, afinal, baixos índices de vozeamento já parecem ser capazes

12 Conforme já mencionado na seção anterior, é possível que, na alta variabilidade referente ao padrão 25% de vozeamento, os participantes estejam fazendo uso de outras pistas acústico-articulatórias, tais como a duração do segmento vocálico que antecede a fricativa. Essa questão será abordada ainda na presente seção, ao nos referirmos às limitações do trabalho e a sugestões de estudos futuros.

13 Derwing & Munro (2015b) definem ‘compreensibilidade’ como a facilidade ou a dificuldade que um ouvinte apresenta ao ouvir um enunciado. A verificação de tal construto se dá através de respostas dos ouvintes por meio de escalas de *Likert*.

de promover a identificação das fricativas como sonoras, pelo menos por parte de ouvintes brasileiros. Nesse sentido, em termos de prática de ensino de fonética, incentivamos a instrução referente a tal aspecto fonético-fonológico de uma forma contextualizada, e, de modo mais importante, sem necessariamente exigir do aprendiz precisão nativa no que diz respeito a este vozeamento. Em outras palavras, situações de *feedback* corretivo referentes a este aspecto devem ser repensadas pelos professores de língua, que deverão fazer uso dessa prática somente quando realmente a produção de tais aspectos incorrer em problemas de inteligibilidade. Considerando-se, ainda, que a inteligibilidade é um construto compartilhado entre falante e ouvinte, e que o parâmetro de inteligibilidade não deva residir apenas no falante nativo, ressaltamos, ainda, a pertinência de replicarmos tal estudo com ouvintes de outras línguas-maternas, como o crioulo haitiano e o chinês, por exemplo, para verificarmos os efeitos de tais padrões na interação do português brasileiro em um contexto de língua-franca.

Finalmente, os dados apresentados no presente trabalho, em consonância com a visão dinâmica aqui advogada, permitem uma discussão a respeito dos primitivos fonético-fonológicos empregados no processo de identificação das categorias funcionais ‘surdo’-‘sonoro’ no PB, discussão essa que constitui o terceiro objetivo específico do presente trabalho. Primeiramente, apontamos que tanto a produção quanto a própria identificação das categorias fonológicas se dá através de uma gradiência, e não através da ausência ou presença de uma categoria em sua plenitude. Dessa forma, e tendo por base a visão de língua aqui sustentada, argumentamos contra a adoção de um primitivo contra o traço primitivo e sua caracterização binária. De fato, ainda que sejam duas as categorias funcionais, [s] e [z], a escolha por uma ou outra categoria vai além da mera ‘presença’ ou ‘ausência’ de vibração das cordas vocais.

Como contra argumento, seria possível sugerir que a adoção do primitivo binário poderia ser mantida, desde que repensássemos o “limiar” acústico para a troca de um valor de traço para outro. Em outras palavras, estaríamos alterando os limiares fonéticos para a atribuição fonológica, caso concebêssemos uma visão dicotômica entre as duas ciências. Uma possível solução analítica, nesse sentido, seria atribuir como ‘sonoro’ todo o segmento com mais de 50% de vozeamento, e como ‘surdo’ todo o segmento com um grau de vozeamento abaixo desse. A observação dos dados deste estudo, entretanto, constitui argumento contra esta possível solução teórica, ao apontar que a atribuição das categorias de surdo e sonoro não somente apresenta uma gradiência, mas varia entre participantes. A par-

tir desses dados, e considerando-se o arcabouço teórico que sustenta este trabalho, podemos pensar no gesto articulatório (Browman & Goldstein 1986; 1992) ou no gesto acústico-articulatório (Albano 2001) como primitivos que expressam tal gradiência. Modelos perceptuais como o *Perceptual Assimilation Model-L2* (Best & Tyler 2007), com base no primitivo de Browman e Goldstein, bem como a versão revista de tal modelo, proposta por Perozzo (2017), com base no primitivo acústico-articulatório (esse último o qual adotamos neste trabalho), não somente permitem a formalização do processo perceptual, mas também expressam o processo desenvolvimental de uma maneira mais ecologicamente plausível, através da qual o aprendiz vai “aprender fazendo” (cf. Albano 2001). Através de sua experiência, o indivíduo vai aprender os efeitos que os gestos por ele produzidos podem ter no processo de inteligibilidade, bem como delimitar as próprias ‘bordas’ perceptuais a partir da interação com os outros falantes. Estamos, assim, advogando a favor de um primitivo que, ao mesmo tempo em que se mostre epistemologicamente coerente com a argumentação aqui feita, venha a condizer com a leitura dos dados apresentados na seção anterior.

Ao considerarmos o próprio primitivo aqui discutido e a noção de temporalidade que se mostra intrínseca a ele, não podemos, entretanto, desconsiderar a possibilidade que a identificação entre as categorias ‘surdo’-‘sonoro’ vá muito além da fricativa *per se*. De fato, ao levarmos em consideração que, na variedade de português brasileiro falada no sul do Brasil, vogais que antecedem segmentos sonoros tendem a ser significativamente mais longas do que aquelas que antecedem consoantes surdas (M. Alves 2016; Ribeiro 2017), não podemos de maneira alguma descartar a possibilidade de ser essa uma pista acústico-articulatória levada em consideração pelos ouvintes. No presente estudo, uma vez que os estímulos foram manipulados a partir de fricativas plenamente vozeadas, as vogais tiveram suas durações controladas, equivalendo à duração de uma vogal que antecede um segmento sonoro, independentemente do grau de vozeamento da fricativa. De certa forma, talvez esse seja um fator que consiga prover um maior entendimento acerca do alto grau de variabilidade encontrado no padrão de 25% de vozeamento ao longo da fricativa: não podemos descartar a possibilidade de que alguns participantes, ao atribuírem o *status* de sonoro à fricativa, tenham se valido das informações referentes à duração vocálica. Dado que não testamos a duração de tal vogal no presente experimento, encontramos, neste ponto, uma limitação do estudo que abre caminhos para investigações futuras, através da verificação dos efeitos da duração vocálica conjugada com o grau de vozeamento. Essa verificação é

sobretudo interessante ao observarmos o padrão de 25%, pois poderá dizer o quanto o ouvinte faz uso da pista acústico-articulatória de vozeamento para apoiar sua identificação categórica quando a vibração das pregas vocais é produzida em uma porção curta da fricativa. A partir dos resultados de estudo futuro, acreditamos que estaremos, também, colaborando com o próprio ensino de pronúncia do PB, uma vez que serão fornecidos insumos à discussão acerca da necessidade ou não de se abordar a duração vocálica como pista decisiva para a categorização das distinções em pares como 'casa' – 'caça'.

As questões aqui levantadas abrem caminhos para uma série de investigações, que no momento já começam a ser traçadas por nosso grupo de investigação. Apesar do caráter sucinto do presente trabalho e de suas limitações, acreditamos ter provido contribuições bastante importantes, tanto no que diz respeito à definição do primitivo fonético-fonológico para a produção e percepção fônica, quanto para o próprio contexto de ensino de pronúncia, ao argumentarmos a favor de uma postura mais flexível do professor, sobretudo no que diz respeito ao *feedback* corretivo frente a tal fenômeno. Esperamos, com este trabalho e seus decorrentes estudos, promover uma reflexão que se mostre pertinente tanto para as áreas da Teoria e Análise Linguística quanto para a Linguística Aplicada ao Ensino de Línguas.

Referências

- Akerberg, M. A. (2004). A importância da palavra escrita para a pronúncia. In A. R. M. Simões, A. M. Carvalho & L. Weidemann (Eds.), *Português para falantes de espanhol: Artigos selecionados escritos em português e inglês* (pp. 115-124). Campinas, Brasil: Pontes Editores.
- Albano, E. C. (2001). *O gesto e suas bordas - esboço da Fonologia Acústico-Articulatória para o português brasileiro*. Campinas, Brasil: FAPESP/Mercado de Letras.
- Albuquerque, J. I. A. *Inteligibilidade e compreensibilidade de línguas adicionais: Um olhar via sistema dinâmico complexo para os dados de haitianos, aprendizes de português brasileiro como la* (Tese de doutoramento - em elaboração, Universidade Federal do Rio Grande do Sul).
- Alves, M. A. (2015). *Estudo dos parâmetros acústicos relacionados à produção das plosivas do português brasileiro na fala adulta: Análise acústico-quantitativa* (Tese de doutoramento, Universidade Federal de Santa Catarina).

- Alves, U. K. (2015). Ensino de pronúncia na sala de aula de língua estrangeira: Questões de discussão a partir de uma concepção de língua como sistema adaptativo e complexo. *Versalete*, 3(5), 374-396.
- Alves, U. K. (2018). Teoria dos Sistemas Dinâmicos e desenvolvimento fonético-fonológico em uma nova língua. In E. Ortiz-Preuss & I. Finger (Eds.), *Um conceito e duas línguas: a dinâmica do processamento bilingue* (pp. 81-128). Campinas, Brasil: Pontes Editores.
- Alves, U. K. & Albuquerque, J. I. A. (2017). Compreensibilidade em L2: uma discussão sobre o efeito da experiência do ouvinte e do tipo de meio em excertos do português brasileiro produzidos por um falante haitiano. *Revista X*, 12(2), 43-64.
- Alves, U. K.; Brisolará, L. B. & Perozzo, R. V. (2017). *Curtindo os sons do Brasil: Fonética do português do Brasil para hispanofalantes*. Lisboa, Portugal: LIDEL Edições Técnicas.
- Alves, U. K. & Silva, A. H. P. (2016). Implicações de uma perspectiva Realista Direta para o PAM-L2: Desafios teórico-metodológicos. *Revista do Gel*, 13(1), 107-131.
- Alves, U. K. & Zimmer, M. C. (2015). Perception and production of English VOT patterns by Brazilian learners: The role of multiple acoustic cues in a DST perspective. *Alfa Revista de Linguística*, 59(1), 155-175.
- Audacity. (2015). Software disponível em: <<http://www.audacityteam.org/>>, versão 2.1.2.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W. & Schoenemann, T. (2009). Language is a complex adaptive system - position paper. *Language Learning*, 59(1), 1-26.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: issues in cross-language research* (pp. 171-204). Timonium, Estados Unidos da América: York Press.
- Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception: commonalities and complementarities. In O-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning. In honor of James Emil Flege* (pp. 13-34). Amsterdão, Holanda: John Benjamins Publishing Company.
- Boersma, P. & Weenink, D. (2008). *Praat. Doing Phonetics by Computer (versão 5.3.84)*. Disponível em: <<http://www.fon.hum.uva.nl/praat/>>.
- Browman, C. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3(1), 219-252.
- Browman, C. & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(1), 155-180.
- Byrne, D. (2013). Evaluating complex social interventions in a complex world. *Evaluation*, 19(3), 217-228.
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide*. Cambridge, Inglaterra: Cambridge University Press.
- Dancy, J. (1985). *Epistemologia contemporânea*. Lisboa, Portugal: Edições 70.

- De Bot, K. (2015). Rates of change: Timescales in second language development. In Z. Dörnyei, P. D. Macintyre & A. Henry (Eds.), *Motivational Dynamics in Language Learning* (pp. 29-37). Bristol, Inglaterra: Multilingual Matters.
- De Bot, K. (2017). Complexity theory and dynamic systems theory: Same or different? In L. Ortega & Z. Han (Eds.), *Complexity Theory and Language Development: in celebration of Diane Larsen-Freeman* (pp. 51-58). Amsterdão, Holanda: John Benjamins Publishing Company.
- De Bot, K., Lowie, W., Thorne, S. & Verspoor, M. (2013). Dynamic systems theory as a comprehensive theory of second language development. In M. P. G. Mayo, M. J. G. Mangado & M. M. Adrián (Eds.), *Contemporary Approaches to Second Language Acquisition* (pp. 199-220). Amsterdão, Holanda: John Benjamins Publishing Company.
- De Bot, K., Lowie, W., Thorne, S. & Verspoor, M. (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language & Cognition*, 10(1), 7-21.
- De Leeuw, E., Mennen, I. & Scobbie, J. (2012). Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism*, 17(6), 683-700.
- Derwing, T. M. & Munro, M. J. (2015a). The interface of teaching and research: what type of pronunciation instruction should L2 learners expect? In P. L. Luchini, M. A. García Jurado & U. K. Alves (Eds.), *Fonética y Fonología: articulación entre enseñanza e investigación* (pp. 14-26). Mar del Plata, Argentina: Universidad Nacional de Mar del Plata.
- Derwing, T. M. & Munro, M. J. (2015b). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdão, Holanda: John Benjamins Publishing Company.
- Field, A. (2009). *Descobriendo a estatística usando o SPSS*. Porto Alegre, Brasil: Artmed.
- Fowler, C. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 8(1), 113-133.
- Fowler, C. (1986). An event approach to a theory of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14(1), 3-28.
- IBM Corp. (2012). *IBM SPSS Statistics for Windows – Version 21.0*. Armonk, NY: IBM Corp.
- Jackson, F. (1977). *Perception: A representative theory*. Nova Iorque, Estados Unidos da América: Cambridge University Press.
- Jackson, F. (2010). Representative realism. In J. Dancy, E. Sosa & M. Steup (Eds.), *A companion to epistemology* (pp. 702-705). (2ª ed.) Malden, Estados Unidos da América: Blackwell.
- Kupske, F. F. & Alves, U. K. (2017). Orquestrando o caos: O ensino de pronúncia de língua estrangeira à luz do paradigma da complexidade. *Fórum linguístico, Florianópolis*, 14(4), 2771-2784.

- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141-165.
- Larsen-Freeman, D. (2015a). It's about time. *The Modern Language Journal*, 98(2), 665-666.
- Larsen-Freeman, D. (2015b). Complexity theory. In B. Van Patten & J. Williams (Eds.), *Theories in Second Language Acquisition* (pp. 227-244). (2^a ed.) Londres, Inglaterra: Routledge.
- Larsen-Freeman, D. (2015c). Saying what we mean: Making a case for 'language acquisition' to become 'language development'. *Language Teaching*, 48(4), 491-505.
- Larsen-Freeman, D. (2017). Complexity Theory: the lessons continue. In L. Ortega & Z. Han (Eds.), *Complexity Theory and Language Development: In celebration of Diane Larsen-Freeman* (pp. 11-50). Amsterdão, Holanda: John Benjamins Publishing Company.
- Larsen-Freeman, D. & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford, Inglaterra: Oxford University Press.
- Levis, J. M. (2005). Changing concepts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369-377.
- Lima Jr., R. M. (2016a). A necessidade de dados individuais e longitudinais para análise do desenvolvimento fonológico de L2 como sistema complexo. *Revista Virtual de Estudos da Linguagem*, 14(27), 203-225.
- Lima Jr., R. M. (2016b). Análise longitudinal de vogais do inglês-L2 de brasileiros. *Gradus: Revista Brasileira de Fonologia de Laboratório*, 1(1), 145-176.
- Lowie, W. (2011). Early L2 Phonology: a dynamic approach. In M. Wrembel, M. Kul & K. Dziubalska-Kolaczyk (Eds.), *Achievements and perspectives in SLA of speech: New Sounds*, 2(1), 159-170.
- Lowie, W. (2017). Lost in state space? Methodological considerations in Complex Dynamic Theory approaches to second language development research. In L. Ortega, & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 123-141). Amsterdão, Holanda: John Benjamins Publishers.
- Lowie, W. & Verspoor, M. (2015). Variability and variation in Second Language Acquisition orders: a dynamic reevaluation. *Language Learning*, 65(1), 63-88.
- Martins, C. (2011). *Manual de análise de dados quantitativos com recurso ao IBM SPSS: saber decidir, fazer, interpretar e redigir*. Braga, Portugal: Psiquilíbrios Edições.
- Munro, M. J. & Derwing, T. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.
- Munro, M. J. & Derwing, T. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 377-396). Malden, Estados Unidos da América: Wiley Blackwell.
- Nitta, R. & Baba, K. (2015). Self-regulation in the evolution of the ideal L2 self: A complex dynamic systems approach to the L2 motivational self system. In Z. Dörnyei, P. D.

- MacIntyre & A. Henry (Eds.), *Motivational Dynamics in Language Learning* (pp. 367-396). Bristol, Inglaterra: Multilingual Matters.
- Oliveira, R. A. (2016). *A influência da instrução fonética explícita na produção e na percepção dos fricativos sonoros por hispanofalantes aprendizes de português como língua estrangeira – PLE* (Dissertação de mestrado, Universidade Federal de Pernambuco).
- Pereyron, L. (2017). *A produção vocálica por falantes de espanhol (L1), inglês (L2) e português (L3): uma perspectiva dinâmica na (multi) direcionalidade da transferência linguística* (Tese de doutoramento, Universidade Federal do Rio Grande do Sul).
- Perozzo, R. V. (2017a). *Sobre as esferas cognitiva, acústico-articulatória e realista indireta da percepção fônica não nativa: para além do PAM-L2* (Tese de doutoramento, Universidade Federal do Rio Grande do Sul).
- Perozzo, R. V. (2017b). Percepção fônica de línguas não nativas no arcabouço da cognição e do realismo indireto: Complementaridade entre aspectos cognitivos e filosóficos a partir do PAM-L2. *Gradus*, 2(1), 52-72.
- Rauber, A. S., Rato, A., Kluge, D. & Santos, G. (2012). TP: perception tests and perceptual training with immediate feedback, versão 2.0. Disponível em: <http://www.worken.com.br/tp_regfree.php>. Acedido em: 15 out. 2012.
- Real Academia Española, Asociación de Academias de la Lengua Española. (2011). *Nueva gramática de la lengua española. Fonética y fonología*. (1ª ed.) Barcelona, Espanha: Espasa libros.
- Ribeiro, R. S. (2017). *Duração de vogais tônicas antecedentes a consoantes plosivas no português brasileiro* (Trabalho de Conclusão de Curso, Universidade Federal do Rio Grande do Sul).
- Schwartzhaupt, B. M., Alves, U. K. & Fontes, A. B. A. L. (2015). The role of L1 knowledge on L2 speech perception: investigating how native speakers and Brazilian learners categorize different VOT patterns in English. *Revista de Estudos da Linguagem*, 23(2), 311-334.
- Silva, A. H. P. (2014). A variável tempo nos estudos em aquisição. In A. Brawerman-Albini & M. L. C. Gomes (Eds.), *O jeitinho brasileiro de falar inglês: Pesquisas sobre a pronúncia do inglês por falantes brasileiros* (pp. 33-48). São Paulo, Brasil: Pontes Editores.
- Silveira, R. & Souza, T. T. (2011). A percepção e a produção das fricativas alveolares da língua portuguesa por hispano-falantes. *Revista de Estudos da Linguagem*, 19(2), 167-184.
- Sobral, C. S., Nobre, M. M. R. & Freitas, M. A. (2006). Relação fone-fonemagrafema na produção oral de aprendizes de PLE. *Portuguese Language Journal*, 1, 1-18.
- Verspoor, M. (2015). Initial conditions. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 38-46). Bristol, Inglaterra: Multilingual Matters.

Zimmer, M. C. & Alves, U. K. (2012). Uma visão dinâmica da produção da fala em L2: o caso da Dessonorização Terminal. *Revista da Abralín*, 11(1), 221- 272.

[recebido em 14 de setembro de 2018 e aceite para publicação em 27 de novembro de 2018]