

FONOLOGIA

Paulo Chagas de Souza
DL/FFLCH/USP

2024

CAPÍTULO 1

INTRODUÇÃO

“... I will try to follow a middle path
between inductive prudence and deductive presumption...”

“... Tentarei seguir um caminho intermediário
entre a prudência indutiva e a presunção dedutiva...”

Labov (1994: 13). Principles of Linguistic Change. Vol. 1

Se partirmos da divisão do signo proposta por Saussure em significado e significante, a qual pode ser estendida à linguagem em geral, a fonologia se ocupa essencialmente da face do significante, apesar de parte de suas análises resvalar também para o significado. Embora o significante possa ser sonoro, gestual ou escrito, a modalidade predominante historicamente tem sido e ainda hoje é a sonora. Parte considerável das línguas orais ainda não tem sistema de escrita associado a elas, e o número de falantes de línguas de sinais em termos mundiais é bem menor do que o de falantes de línguas orais.

Este livro trata apenas das línguas orais. Por dois motivos. O primeiro é que as línguas de sinais (ou línguas gestuais, como se diz em Portugal) têm semelhanças com as orais em termos de fonologia, mas o simples fato de usar outro meio, o visual, e não o auditivo, faz com que as línguas de sinais apresentem diferenças fundamentais com relação às orais. A substância distinta implica um tipo de sistema com diferenças enormes. Destaco uma possibilidade não encontrada nas línguas orais: o fato de podermos fazer gestos diferentes com cada mão. Isso abre uma possibilidade inexistente nas línguas orais. Dadas essas diferenças, se tornaria muito mais difícil a exposição e estenderia muito a extensão do livro. Acresce ainda que a teoria fonológica nas línguas de sinais ainda não se desenvolveu tanto

quanto a das línguas orais, e mesmo questões de base ainda são bastante controversas. Um último motivo é que, inclusive em virtude dessas diferenças, não tenho familiaridade suficiente para englobar as línguas de sinais neste livro.

Quanto à outra modalidade de língua, a escrita, ela surgiu para representar o que falamos. Normalmente a tentativa inicial, bem mais intuitiva, é a de representar o significado, mas com o passar dos séculos frequentemente se chega à conclusão que um caminho menos óbvio, representar o significante, facilita essa tarefa. Ocorre, no entanto, que à medida que o tempo passa, a escrita toma um percurso próprio, se desgarrando em parte do oral, e não necessariamente está muito próxima do significante atual. Uma grafia tradicional pode se manter, mesmo que esteja extremamente distante da língua falada. Um último ponto é que a escrita apresenta características bem distintas da língua oral, por exemplo, por ser visual e bidimensional. Por isso tratarei dela em outro livro que estou começando a redigir.

Por falar em língua oral e escrita, uma confusão tradicional, surpreendentemente ainda muito arraigada na concepção popular ou leiga, e presente inclusive em material didático de línguas, é a que não distingue claramente sons e letras, ou que erroneamente faz afirmações do tipo de “a letra *x* em *exame* é pronunciada como *z*”. Pode-se compreender que isso tenha uma função didática, mas o fato é que não pronunciamos letras, pronunciamos sons. Seria mais correto dizer que “a letra *x* em *exame* representa um *z* na pronúncia” ou “a letra *x* em *exame* corresponde a um *z* na pronúncia”. A escrita é imensamente mais recente do que as línguas orais, e surgiu com a intenção de representar o que era dito oralmente. As estimativas do surgimento das línguas orais variam enormemente: de 50.000 a 500.000 anos atrás. A escrita surgiu há cerca de 5.000 anos, tendo, portanto, uma história entre dez e cem vezes mais breve do que a das línguas orais. E mesmo com essa história milenar, ainda não se difundiu entre todos os povos ou para todas as línguas.

Para podermos evitar essa confusão entre sons e letras, e distinguir as duas coisas claramente, vamos usar a notação padronizada para cada um desses tipos de significante. Quando estivermos falando de letras, elas aparecerão entre parênteses angulares: < >. Se, por outro lado, estivermos falando de sons, eles aparecerão entre colchetes: []. Assim <*a*> representa a letra *a* e [*a*] representa o som vocálico emitido com a abertura máxima da cavidade oral, a vogal presente em palavras como *lá*. Quando for apresentada a grafia das palavras, e não sons isolados, ela aparecerá sem nada que a delimite, ou em itálico, se aparecer no corpo do texto.

Para representar a pronúncia, como convencionalmente se faz no mundo todo, adoto o Alfabeto Fonético Internacional, com uma ou duas modificações devidas a críticas a ele que compartilho com outros autores. Uma questão que pode surgir com relação a isso é se é de fato necessário adotar um alfabeto internacional padronizado. Para justificar essa

padronização, vejamos o que ocorre com algumas letras do alfabeto de origem romana. Se examinarmos outras línguas, veremos que muitas vezes a mesma letra não representa a mesma pronúncia em línguas diferentes. Embora alguns dos símbolos fonéticos não sejam óbvios, o essencial neste momento é que fique claro que a mesma letra corresponde a sons diferentes. Vejamos alguns exemplos:

- (1) a letra <a> em português: mar [a]
em húngaro: magyar [ɔ]
em inglês: hat [æ], war [ɔ]
- (2) a letra <z> em português: zero [z]
em alemão: Zeitung [ts]
em espanhol europeu: zapato [θ]
em espanhol da América: zapato [s]

Por alguns exemplos podemos ver que, inclusive na mesma língua, uma determinada letra pode representar (ou talvez um termo melhor seja estar associada) a várias pronúncias diferentes. É o que se verifica com a letra <a> nos dois exemplos do inglês, ou a letra <z> nas duas variedades do espanhol.

Podemos constatar o mesmo com relação à vogal grafada <a> em português, em palavras como: *mar, loja, ano*. Ao <a> escrito normalmente correspondem três pronúncias diferentes, que podem ser transcritas, respectivamente [a], [ɐ] e [ẽ]. Mesmo dentro de uma única palavra temos exemplos de sons diferentes correspondentes ao que é grafado <a>, como no exemplo a seguir, que apresenta os três sons vocálicos citados:

- (3) bacana [ba'kẽnɐ]

O alfabeto fonético internacional foi criado e é gerido pela Associação Fonética Internacional (podemos abreviar os dois como AFI ou, como é mais comum, IPA, do inglês *International Phonetic Alphabet* e também *International Phonetic Association*). Ele se baseia nas letras latinas minúsculas para representar os sons mais comuns, sendo que cada uma delas representa um som que mais comumente essa letra representa nas línguas em que é usada ou o som que representava em latim. Por exemplo, a consoante <t> representa o som [t] que encontramos em palavras como *tempo*, e não o som que ele representa em muitas regiões do Brasil em palavras como *ótimo* (o qual podemos informalmente escrever como <tch>, nem o som que comumente ele tem no inglês norteamericano em palavras como *water*, o qual é semelhante ao som que nosso <r> representa em palavras como *hora*. O <z>, que faz parte do alfabeto de origem latina, mas era usado para representar o [dz] grego, representa um som que não existia no latim nem corresponde à pronúncia antiga do grego,

o [z]. Mas, como foi dito, esse é o som que ele mais comumente representa nas línguas que adotam o alfabeto latino.

Passando para o campo teórico, a partir de Trubetzkoy, se tornou tradicional distinguir entre duas disciplinas que tratam dos sons linguísticos: a fonética e a fonologia. Este livro trata principalmente da fonologia, mas considera imprescindível a contribuição da fonética para se avançar na fonologia. Vejo a fonética e fonologia como disciplinas complementares, que auxiliam uma à outra a se refinar. Entender como os sons são produzidos, transmitidos e percebidos, tarefa da fonética, ajuda a entender como eles funcionam gramaticalmente, tarefa da fonologia. Entender como eles funcionam gramaticalmente, por sua vez, nos leva a formular perguntas sobre a natureza concreta dos sons linguísticos. Na minha atuação na área, examino a fonologia, e isso me indica em que aspectos focalizar. Em seguida, olho a fonética para tentar entender melhor possíveis razões para os fenômenos fonológicos. Com isso, volto para a fonologia. E assim por diante. Espero que o livro reflita esse vaivém construtivo.

É importante já de início apresentar alguns termos técnicos e ressaltar a distinção entre eles. Parto do termo *som*. Quando utilizar o termo *som*, estarei me referindo a uma produção específica num determinado instante por uma pessoa determinada. Podemos medir e analisar as características específicas desse som, que pode ser por exemplo um [t], em que a parte da frente da língua faz contato com os dentes incisivos superiores ou com a região logo acima deles, os alvéolos. Se, no entanto, quisermos falar não só daquele [t] específico produzido por aquela pessoa naquele determinado momento mas do [t] como um tipo de som, fazendo então uma generalização e abstração, me referirei ao *fone* [t]. Podemos, então, descrever o fone [t] na língua portuguesa.

Fone é um termo utilizado principalmente na fonética. Um termo correspondente em fonologia é *segmento*. O uso deste termo se baseia na ideia de que podemos segmentar ou dividir a cadeia sonora em pedaços ou segmentos que se sucedem. É a concepção que está por trás da escrita alfabética. Há ainda o termo *fonema*, que leva não só a parte sonora em si do fone em consideração, ou o significante, mas sua relação com o significado, como veremos no capítulo 4.

Faço também neste capítulo algumas observações com relação ao texto e à notação utilizada. Em situações em que o símbolo fonético possa ficar difícil de identificar, se colocado imediatamente antes ou depois dos colchetes da transcrição fonética, deixarei um espaço entre ele e os colchetes ou barras. É o caso de símbolos como [ɲ] e [ɳ] no capítulo 2, que aparecerão assim: [ɲ] e [ɳ].

Em certos pontos, faço referência à pronúncia de São Paulo. É a pronúncia com que tenho mais familiaridade dos detalhes, principalmente porque é onde sempre morei no Brasil, de

modo que minha pronúncia é basicamente paulistana, embora com pinceladas nordestinas, por ser filho de piauienses: nasalizo, por exemplo, as vogais das duas primeiras sílabas de *banana*. Em termos de familiaridade com línguas variadas, minha diacronia fez com que ela fosse maior com relação a línguas da Europa, e em segundo lugar da Ásia, mas meu esforço foi o de incluir exemplos e exercícios de todos os continentes, e de famílias linguísticas o mais variadas possível.

Quanto ao conteúdo do livro, se uma característica fundamental de um pesquisador é ter um olhar crítico, tento ao máximo fazer isso com relação às teorias. Que aspectos parecem bem justificados? Que aspectos parecem manobras para os dados caberem na teoria? Que aspectos se converteram quase em dogmas e com o tempo se tornaram quase impossíveis de questionar?

Em relação ao conteúdo do livro, a cada passo é necessário fazer escolhas. Nos temas escolhidos, no tratamento dado ao tema, e assim por diante. Nos casos mais controversos, me esforço por deixar clara a motivação das minhas escolhas.

O livro é relativamente aberto, inclusive no sentido de tentar apresentar de forma ampla as teorias fonológicas, e não uma específica. Embora talvez minhas preferências fiquem claras, prefiro transitar entre a indução e a dedução, como se propõe Labov na epígrafe deste capítulo. A indução pura e simples pode levar a generalizações soltas que não compõem um todo coeso, mas a dedução pura e simples pode levar a equívocos de linguista de gabinete. Prefiro não errar por nenhum dos dois excessos. Erros sempre haverá, mas isso faz parte de ser humano.

CAPÍTULO 2

CONSOANTES

Neste capítulo examinamos as consoantes, o que as caracteriza como tais, e os tipos de consoantes existentes.

2.1 DEFINIÇÃO

Entre os sons utilizados nas línguas do mundo, uma subdivisão muito familiar a todos nós é a que existe entre as consoantes e as vogais, embora seja mais comum pensarmos nas letras do que nos sons. O que distingue esses dois grupos de sons? Se pensarmos articulatoriamente, ou seja, em termos de como esses sons são produzidos, mesmo as vogais mais fechadas, como [i] e [u] são pronunciadas sem uma grande obstrução à passagem da corrente de ar, enquanto que nas consoantes há um obstáculo total ou muito acentuado a essa passagem pelo menos em algum ponto do trato vocal, como, por exemplo, na produção de um [b] ou de um [f], respectivamente.

O nome *consoante* traz em sua origem a ideia de que esse é um tipo de som que não costuma ocorrer sozinho, só acompanhando outro som que “soa” sozinho, a *vogal*, cujo nome provém da palavra *voz* em latim (*vox, vocis*, do qual se derivou o adjetivo *vocālis*, depois usado também como substantivo).

Em várias línguas, a denominação de consoantes e vogais alude a propriedades desse tipo. Um exemplo é o das línguas eslavas, entre as quais podemos citar o tcheco. Nele o termo vogal é *hláska* e consoante é *souhláska*, ambas derivadas da raiz de *hlas* ‘voz’. O termo que significa vogal, em tcheco, é um derivado do termo *voz*, num paralelo com *vocalis* e *vox, vocis*. O prefixo *sou-* em tcheco equivale ao nosso *com-*, indicando que a consoante tem que acompanhar um som ou uma “voz”. Provavelmente essa denominação é calcada no termo latino *consonans*.

Já em chinês, vogal é *mǔyīn* ou *yuányīn*. A tradução literal desses dois termos é, respectivamente, ‘som-mãe’ e ‘som-base’ ou ‘som fundamental’. Consoante em chinês é *zǐyīn*

ou *fǔyīn*. A tradução literal desses dois termos é, respectivamente, ‘som-filho’ e ‘som auxiliar’. Seguem os termos citados, com sua grafia em caracteres chineses e em *pinyin*, a transliteração oficial do mandarim:

(1) ‘vogal’ = 母音 mǔyīn ou 元音 yuányīn

(2) ‘consoante’ = 子音 zǐyīn ou 輔音 fǔyīn

Se isso que foi dito dá a impressão de que as consoantes são menos importantes do que as vogais, por que começar o livro falando delas? A principal razão é que mais fácil falar das consoantes, por motivos que ficarão mais claros quando chegarmos no próximo capítulo.

2.2 CLASSIFICAÇÃO

Como o que caracteriza as consoantes como tais é a presença de um obstáculo à passagem do ar, a classificação delas se baseia em grande parte justamente nas características do obstáculo imposto à passagem da corrente de ar pelo trato vocal. Duas dessas propriedades são o local onde é produzido esse obstáculo e o tipo de obstáculo que é produzido: total, parcial, etc.

O termo técnico utilizado para designar o local onde é produzido esse obstáculo é *ponto de articulação*, enquanto o utilizado para designar o tipo de obstáculo é *modo de articulação*. Para nos referirmos ao obstáculo em si usamos o termo *constricção*, que se origina na mesma raiz dos termos *estreito*, *estreitamento*, *restrito*, *restringir*, todos derivados da raiz do verbo *stringō*, ‘apertar; comprimir; estreitar’.

Para que seja produzido um obstáculo à passagem do ar, é necessário que dois órgãos estejam envolvidos, os quais entram em contato ou chegam perto disso. São os articuladores envolvidos na produção da constricção. Na maioria dos casos, um órgão se desloca em direção a outro estático, como, por exemplo, quando a ponta da língua encosta nos dentes. Em alguns casos, um órgão tem um deslocamento muito maior quando comparado a outro. É o caso dos sons bilabiais, que envolvem um deslocamento bem maior do lábio inferior do que do lábio superior. Aquele na realidade se movimenta mais por estar situado na mandíbula (ou maxilar inferior). Articulador ativo é, então, o único que se desloca ou aquele que tem um deslocamento muito maior que o outro envolvido na produção da consoante. O outro é o articulador passivo.

Em fonologia, muitas vezes é importante pensar no ponto de articulação específico de uma consoante, por exemplo, o [f] é labiodental, pois na sua produção o lábio inferior se aproxima dos dentes incisivos superiores, mas frequentemente pensamos numa

classificação menos detalhada, que leva em consideração primeiramente apenas o articulador ativo. No caso do [f], trata-se de uma consoante labial, já que é o lábio inferior que se desloca até quase encostar nos dentes incisivos superiores. Na parte inicial deste capítulo, focalizaremos esse tipo de classificação menos detalhada, a qual se baseia essencialmente apenas no articulador ativo, e examinaremos principalmente fones do português. Depois disso ampliaremos nosso campo de visão, incluindo línguas variadas, e veremos uma classificação mais detalhada das consoantes.

Todas as línguas do mundo usam como articuladores ativos órgãos localizados na cavidade oral. À primeira vista, poderíamos pensar em duas possibilidades: os lábios, que ficam no limite entre a cavidade oral e o ambiente externo ao organismo, e a língua, que é o único órgão móvel de fato contido na cavidade oral. Mas a língua é extremamente plástica, sendo capaz de um grande número de movimentos, e suas partes anterior e posterior se movimentam de maneira bastante independente uma da outra. Por isso, e por observar o que se encontra nas línguas do mundo todo, é mais adequado considerar que a língua, na cavidade oral, é composta de dois articuladores ativos: a coroa da língua, que pode ser dividida em ponta (ou ápice) e lâmina; e o dorso da língua. Lâmina é a parte superior da língua em sua porção anterior. A figura 1 apresenta essas regiões da língua.

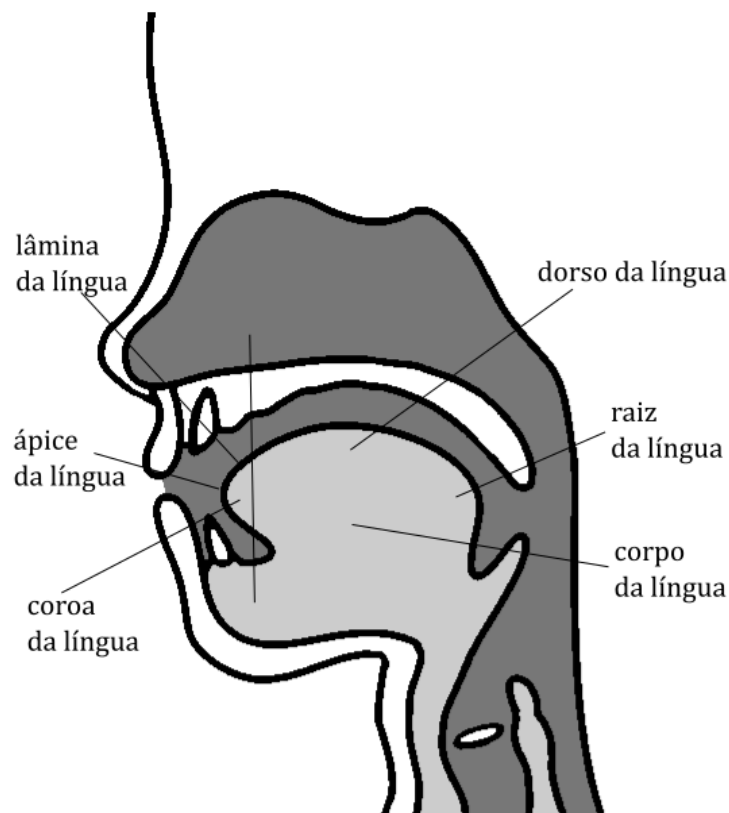


Figura 1. Partes da língua.

Quando produzimos algum fone que envolve o dorso da língua, na verdade movimentamos o corpo da língua, que é sua parte mais interna. Já a coroa da língua pode ser movimentada sem que necessariamente o corpo e o dorso da língua tenham que se deslocar significativamente. Os sons que envolvem essencialmente a movimentação do dorso da língua são chamados *dorsais*, e os que envolvem basicamente a movimentação da coroa da língua são denominados *coronais*. Além disso, há os sons *labiais*, que envolvem a movimentação dos lábios, principalmente o inferior. Portanto, os três articuladores ativos localizados na cavidade oral são: o lábio inferior, a coroa da língua, e seu dorso.

Passemos a examinar as consoantes, principalmente as encontradas no português, com mais detalhes. Vejamos agora quais os modos de articulação possíveis. Se as vogais não têm obstáculo à passagem do ar e as consoantes têm, num certo sentido as consoantes “mais consonantais” de todas são as que têm o maior obstáculo possível à passagem do ar. São as oclusivas, aquelas em que se interrompe completamente a passagem do ar, em geral com uma quantidade cada vez maior de ar se acumulando no trato vocal, até que a obstrução se desfaz subitamente com uma explosão. São sons semelhantes ao estouro de uma bexiga, instantâneos. O termo *oclusivas* tem a mesma raiz da palavra *close* do inglês, e portanto significa ‘produzidas com fechamento’ (da passagem do ar).

Provavelmente por elas serem as consoantes mais diferentes das vogais, ou as consoantes “mais consonantais” de todas, todas as línguas têm oclusivas. Trata-se de um universal absoluto. Como defendem Jakobson e Halle (1956: 37), ao comentar sílabas como /pa/ ditas por uma criança começando a adquirir a linguagem, essa polaridade entre um mínimo de energia, a oclusiva, e um máximo de energia, a vogal, representa um contraste entre a consoante ótima e a vogal ótima.

Além do ponto e do modo de articulação, a terceira característica que distingue uma consoante das outras está relacionada a como a corrente de ar passa pela nossa laringe, a estrutura cartilaginosa localizada na parte anterior do pescoço, também chamada de pomo de adão ou gogó, que é mais proeminente nos homens, e pode ser visto na fig. 2. Dentro dessa estrutura cartilaginosa há duas estruturas formadas de tecido muscular e mucosa, as pregas vocais, antigamente chamadas de cordas vocais. O termo *pregas* parece mais adequado porque o formato delas é de superfícies finas e largas, e não de um feixe alongado. A abertura entre as pregas vocais é denominada glote, e pode ser vista na fig. 3.



Figura 2.

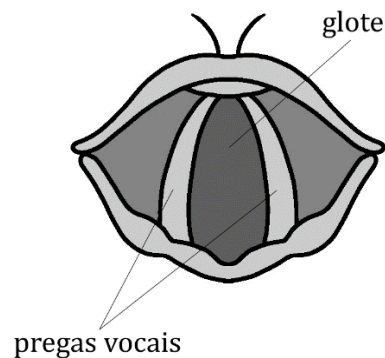


Figura 3.

Essa terceira característica importante na classificação das consoantes é o chamado *estado da glote*, ou seja, como se encontram nossas pregas vocais ao produzirmos cada consoante. As pregas vocais podem estar abduzidas, isto é, separadas, com uma abertura entre elas (do latim *abducere*, 'afastar', formado do prefixo *ab-*, que denota afastamento e *ducere*, 'trazer, puxar'). Podem também ser aduzidas ou aproximadas (do latim *adducere*, 'aproximar', formado do prefixo *ad-*, que denota aproximação, e *ducere*). Se não for um fechamento intenso demais, a corrente de ar vinda dos pulmões produz um fechamento intermitentemente. Nesse segundo caso, elas vibramos.

Há várias maneiras de detectarmos se as pregas vocais estão vibrando ou não. A mais óbvia e mais comum é colocar uma mão sobre a região da laringe enquanto pronunciamos um som. A vibração das pregas vocais comumente pode ser sentida pela mão colocada nessa região. Um par de sons muito bom para percebermos essa distinção são as consoantes [s] e [z], como presentes no início das palavras *sala* e *zero* em português. Se colocarmos a mão sobre a laringe quando pronunciamos um [s], não sentiremos vibração nenhuma, por mais energia que seja dispendida na produção desse som. Por outro lado, se fizermos o mesmo ao pronunciar um [z], mesmo sem fazer esforço vocal, sentiremos essa vibração.

Uma outra forma de identificarmos se está havendo vibração das pregas vocais é pronunciar um som tampando os ouvidos. Mais uma vez, se pronunciamos um [s] tampando os ouvidos, não acontece nada digno de nota. Se fizermos o mesmo, no entanto, pronunciando um [z], sentiremos a cabeça inteira vibrando, podendo até sentir um pouco de tontura. Isso ocorre

porque, quando as pregas vocais vibram, sua vibração reverbera, repercute nos órgãos vizinhos.

Uma última maneira simples de percebermos se as pregas vocais estão vibrando ou não é tentar produzir um som com uma melodia; isso porque uma melodia é uma variação na frequência das notas produzidas através da vibração das pregas vocais. Se estas não vibrarem, é impossível produzirmos uma melodia. Como já deve ser possível imaginar, não conseguimos colocar uma melodia num [s], mas conseguimos se estivermos pronunciando um [z].

O português, como muitas línguas do mundo, tem oclusivas produzidas com cada um dos articuladores ativos: labiais, coronais e dorsais. Isso ocorre porque a tendência da fonologia das línguas é utilizar os recursos disponíveis para a articulação de sons que sejam mais facilmente percebidos como diferentes.

Existem no português duas oclusivas produzidas com cada um dos articuladores ativos, ou seja, labiais, coronais e dorsais, uma que é produzida sem vibração das pregas vocais e outra com essa vibração. São as labiais [p] e [b], as coronais [t] e [d], e as dorsais [k] e [g]. O último símbolo representa sempre o som grafado com <g> em palavras como *gato* e *grande*, nunca o som grafado com <g> em palavras como *gente* ou *lógico*.

Em cada um dos três pares de oclusivas citados, a primeira consoante é produzida sem vibração das pregas vocais, sem voz, sendo chamada de *surda* ou *desvozeada*. A segunda de cada par é produzida com essa vibração, sendo chamada de *sonora* ou *vozeada*. Assim, o [b] é uma oclusiva labial sonora (ou vozeada), e o [t] é uma oclusiva coronal surda (ou desvozeada, ou não vozeada). Uma desvantagem de usar os termos mais tradicionais em português, *surdo* e *sonoro*, é que eles não são opostos de maneira óbvia, ao contrário do par *vozeado* e *desvozeado*. Por outro lado, uma desvantagem de usar os termos *vozeado* e *desvozeado* é que, por eles serem participípios, podem ser interpretados como significando que as consoantes sofreram um processo de *vozeamento* ou *desvozeamento*. Por isso, também, pode-se usar o termo *não vozeado*. Alternarei entre os termos neste livro, para que leitoras e leitores se acostumem a encontrar todos eles.

Outro fone foneticamente classificado como oclusiva e que é muito comum nas línguas em geral, mas ausente linguisticamente no português, é o que é produzido com a interrupção total da passagem do ar justamente nas pregas vocais, através de uma adução com tensão suficiente para impedir que elas vibrem. Ela é denominada parada glotal ou golpe de glote, sendo transcrita com um sinal parecido com o ponto de interrogação, mas sem o ponto na parte inferior, ou seja, [?]. Ele ocorre na expressão paralinguística de negação que pode ser transcrita [m?m] ou [ʔm?m], em que produzimos um [m], interrompemos a passagem do ar e voltamos a produzir um [m]. Pode ocorrer a parada glotal no início dessa expressão

também. Essa “expressão” é classificada como paralinguística (termo formado com o prefixo grego *para-* ‘ao lado de’), por não ser uma palavra do português, embora tenha um significado.

A parada glotal também ocorre na expressão paralinguística *uh-oh* do inglês, entre suas duas porções vocálicas. Embora ela não ocorra necessariamente em palavra nenhuma do português, é muito comum que ela seja utilizada na música. Um ótimo exemplo do uso da parada glotal em música no português é a composição *Pescador de Ilusões* do conjunto *O Rappa*, nesse trecho: “Valeu a pena, ê ê.” Se prestarmos atenção, o trecho “ê ê” pode ser descrito como um “ê” que sofre uma interrupção e depois é retomado, podendo ser transcrito assim: [eʔe]. O mais curioso de tudo é que essa parada glotal nada mais é do que um silêncio mas, apesar disso, vale como um som linguístico ou uma posição na cadeia sonora. Como a parada glotal é produzida justamente pelo fechamento total momentâneo da glote, ela é um “som” surdo. Uma música que é pródiga em exemplos de parada glotal é *Varanda Suspensa*, da cantora e compositora Céu. É um exercício divertido tentar localizar nela todas as ocorrências de parada glotal.

Outro tipo de obstáculo muito comum na produção das consoantes é aquele em que se estreita muito a passagem do ar sem interrompê-la. Se esse estreitamento é acentuado o suficiente para produzir ruído, o tipo de consoante produzido se denomina *fricativa*. Esse nome tem a mesma origem de fricção, e remete ao tipo de ruído comum nesse tipo de consoante. Se para as oclusivas usei a analogia da bexiga explodindo, para as fricativas podemos usar a analogia de quando apertamos e esticamos o bico da bexiga de forma que o ar passa com bastante atrito, produzindo um ruído contínuo.

Com relação às fricativas, a situação já não é a mesma em todas as variedades do português. Mas a situação mais comum é aquela em que os falantes utilizam pelo menos três pares de fricativas, novamente uma surda e outra sonora. No entanto, diferentemente do que ocorre com as oclusivas, temos um par de fricativas labiais e dois pares de fricativas coronais. As labiais são [f] e [v]. Um dos dois pares de coronais é o formado pelo [s] e pelo [z]. A fricativa [s] é encontrada em palavras como *pensar*, *sempre*, grafada com <s>, mas também com diversas outras grafias, como em *certo*, *praça* e *massa*. A fricativa [z] é encontrada em *zebra* e *fazer*, mas também com outras grafias, como em *rosa* ou *exame*.

Se pronunciamos as palavras *lance* e *lanche*, provavelmente conseguimos perceber que, na consoante grafada <c> em *lance* e na consoante grafada <ch> em *lanche*, a língua não fica na mesma posição. Em *lance*, praticamente só erguemos a língua em direção ao céu da boca, enquanto que em *lanche* ela recua. A consoante grafada <ch> em *lanche* e <x> em *luxo* é transcrita [ʃ]. Essa é uma consoante surda. Sua correspondente sonora é a consoante grafada <g> em *gente* e <j> em *junto*. Ela é transcrita [ʒ]. Observe que os símbolos dessas

fricativas coronais recuadas, [ʃ] e [ʒ], são modificações dos símbolos das fricativas coronais não recuadas, [s] e [z].

Foneticamente podem ser também classificados como fricativas dois fones glotais: a fricativa surda, comumente chamada de aspiração, o [h], e sua correspondente “sonora”, embora as pregas vocais vibrem não com voz modal, mas murmurada, na sua produção, o [ɦ] (v. seção 5.12). Ambos são usados, por exemplo, na pronúncia de Belo Horizonte.

Tendo visto as oclusivas e as fricativas, podemos falar de outro tipo de consoante, o qual não ocorre na pronúncia de todos os falantes do português. São as *africadas*, consoantes complexas, compostas de uma fase oclusiva seguida por outra fase fricativa. O som transcrito [t] é uma oclusiva. Mas o som grafado <t> em *tia* não é pronunciado [t] por todos os falantes do português no Brasil. Na verdade, menos falantes pronunciam essa consoante como uma oclusiva nessa palavra, no Brasil.

Se grafarmos essa palavra como <tchia>, estaremos tentando representar a pronúncia africada dessa consoante. O mesmo ocorre com a palavra *dia*, se ela for grafada <djia>. Essa pronúncia africada é transcrita [tʃ] no caso da surda e [dʒ] no caso da sonora. Essas duas transcrições refletem o fato de que essas duas consoantes começam como oclusivas e terminam como fricativas. Se quisermos deixar claro que apesar de aparentemente termos dois “sons” em sequência, se trata de uma única consoante com duas fases distintas, podemos utilizar a ligadura, um sinal auxiliar acima das duas letras de cada par: [t͡ʃ] e [d͡ʒ]. Os sinais auxiliares como a ligadura são denominados *diacríticos* (termo que significa *distintivo*). Na transcrição do português, a utilização desse diacrítico pode até ser considerada um tipo de preciosismo, já que uma africada [t͡ʃ] não se distingue de uma oclusiva [t] seguida da fricativa [ʃ]. Já numa língua como o inglês, pode ser importante fazer essa distinção, deixando claro que em *night shift* temos uma oclusiva seguida de uma fricativa, transcritas [tʃ], ao passo que em *watch* temos uma africada, transcrita [t͡ʃ]. Contudo, essa distinção pode ser representada de outra maneira indicando por exemplo a separação das sílabas com um ponto. A diferença de transcrição ficaria sendo, então, [t.ʃ] em *night shift* e [t͡ʃ] em *watch*. Podemos ainda indicar que em *night shift* há dois morfemas, enquanto que em *watch* só há um.

Como foi dito acima, na produção de uma fricativa, ocorre um estreitamento muito acentuado da passagem do ar, o suficiente para produzir ruído. Se, no entanto, a aproximação do articulador ativo com o passivo não chegar a ser acentuada o suficiente para produzir ruído, o tipo de consoante produzido não será uma fricativa, mas sim uma aproximante. Duas aproximantes que todo falante do português utiliza são as semivogais que ocorrem nas palavras *praia* e *Piauí*. A que é grafada <i> em *praia* é transcrita como [j]. É comum a associação dessa letra com esse som, assim como na grafia do alemão *ja* ‘sim’. A

semivogal que é grafada <u> em *Piauí* é transcrita como [w], como no inglês *we*. As duas são sonoras. O [j] é uma aproximante coronal, e o [w] é uma aproximante labiodorsal, pois os lábios se arredondam e o dorso da língua se ergue em direção ao céu da boca.

Há uma outra aproximante utilizada por parte dos falantes do PB, a qual podemos considerar que ocorre em duas versões. É a pronúncia do <r> em *porta*, comum em uma região ampla, que abrange o interior de São Paulo, parte do Paraná, o sul de Minas Gerais, Mato Grosso e Goiás. É comum que essa consoante seja chamada de *r* caipira, mas o termo técnico para descrever esse tipo de consoante é *retroflexo*, já que é um tipo de consoante em que a língua se curva (se *flexiona*) para trás (*retro*).

Podemos considerar que ele ocorre em duas versões: uma com recuo moderado e outra com recuo acentuado da língua. Esta segunda é a consoante retroflexa tradicional do interior paulista, por exemplo. Se formos distinguir os dois tipos de retroflexo, o do interior paulista é transcrito [ɭ], e o retroflexo suave é transcrito [ɻ], e não aparece na coluna de consoantes retroflexas na tabela do AFI, sendo definido como aproximante alveolar. Há um exercício sobre essa questão ao final do capítulo.

Embora os retroflexos sejam considerados como típicos do interior no estado de São Paulo, tem se difundido a partir da periferia, o retroflexo mais suave, com menor recuo na capital. Em Oushiro e Mendes (2011: 81), vemos que a pronúncia retroflexa já se encontra bem presente na cidade de São Paulo. Há, no entanto, uma assimetria. No levantamento feito, o retroflexo aparece em 42,8 % das ocorrências na periferia, mas só em 21 % nas regiões centrais, uma forte indicação de que ele tem se difundido a partir da periferia. Soriano (2016: 45) sugere que parte dos paulistanos parece ter sensibilidade a esses graus de retroflexão, percebendo a mais suave como parte do falar paulistano e a mais acentuada como mais típica do interior. Ela se refere ao [ɻ] como aproximante alveolar, e ao [ɭ] como aproximante retroflexa, a mesma terminologia usada na tabela do alfabeto fonético internacional, como veremos mais adiante neste capítulo.

Se não formos necessário chegar a esse nível de detalhe, distinguindo dois graus de retroflexão na transcrição, podemos transcrever qualquer retroflexa do português do Brasil com [ɻ] ou com [ɭ].

Em textos de fonologia, é comum encontrar a utilização de alguns símbolos fonéticos sem fazer a distinção claramente entre fricativas e aproximantes. De certa forma, podemos considerar que o som sublinhado na palavra inglesa *mother* e na palavra espanhola *nada* são iguais, transcrevendo ambos com o símbolo fonético [ð], que é um <d> com a haste vertical colocada na diagonal e cortada. Todavia, se atentarmos bem, perceberemos que o som da consoante inglesa é produzido com um certo ruído, enquanto que a consoante do

espanhol não apresenta ruído algum. Em outras palavras, a consoante do inglês é uma fricativa, mas a do espanhol é uma aproximante.

Tratando desse detalhe, um ponto importante a destacar, é que a transcrição fonética de uma palavra não é única. Como no retrato desenhado ou pintado de uma pessoa, podemos representar mais detalhes ou menos da pronúncia, de acordo com a conveniência e a necessidade. As transcrições fonéticas de modo geral podem ser classificadas em dois grupos. Podemos ter uma transcrição fonética *lata* ou uma transcrição *estrita*. São os dois adjetivos que se opõem em expressões como: *em sentido lato*, ou a expressão original latina *lato sensu*, e, de outro lado, *em sentido estrito*, ou a expressão original latina, *stricto sensu*. Algo estrito é mais preciso, mais rigoroso, ao contrário de algo lato. Os termos correspondentes em inglês são, respectivamente, *broad transcription* e *narrow transcription*. Se a *narrow transcription* traz mais detalhes, *transcrição estreita* não parece uma boa tradução, embora seja encontrada na literatura.

Numa transcrição lata, a diferença da pronúncia entre a fricativa e a aproximante dentais mencionadas, do inglês e do espanhol, não precisaria aparecer de forma distinta. Se for importante distinguirmos na transcrição, fazendo uma transcrição mais estrita, podemos usar um diacrítico embaixo do símbolo fonético para representar a consoante do espanhol: [ð̞]. Talvez não seja muito fácil visualizar o diacrítico [̞], mas ele é um pequeno T colocado abaixo do símbolo fonético. A ponta do T representa de maneira icônica uma abertura maior da cavidade oral. Há diacríticos correspondentes para indicar um som mais fechado (⌞), mais anterior (→), ou mais posterior (←) do que o som normalmente representado pelo respectivo símbolo fonético sem esses diacríticos. Se lembrarmos que posterior e anterior aparecem na posição padrão de ilustração do aparelho fonador, em que a pessoa está de perfil, com o rosto no lado esquerdo da ilustração, respectivamente, à direita e à esquerda da ilustração, como nas figuras 1 e 2, deve facilitar imaginar uma seta com a ponta colocada na extremidade da base do “T”, isto é, pensar nesses diacríticos como se fossem, respectivamente, as seguintes setas: ⌞, ⌞̂, ← e →.

Um outro tipo de consoantes, talvez também usado por todos os falantes do português, é o que é produzido elevando a coroa da língua até ela encostar no céu da boca uma única vez muito brevemente. O nome técnico utilizado para se referir a esse tipo de consoante é uma forma aportuguesada de um termo inglês. A palavra inglesa é *tap*, e foi aportuguesada como *tepe*. É o som grafado com <r> nas palavras *Brasil* e *fora*. Ele é transcrito como [ɾ]. Proponho um termo vernáculo para denominar essa consoante na seção 5.8.8.

Outro tipo de consoante encontrado no português, mas que nem todos os falantes utilizam, é uma consoante parecida com o tepe, mas produzida com mais energia, a qual envolve um fechamento intermitente da passagem do ar na cavidade oral, com vários contatos da coroa

da língua com o céu da boca. Esse tipo de consoante se denomina *vibrante*. É a pronúncia do som grafado <r> ou <rr> encontrada no espanhol e no italiano, e comum em imigrantes italianos e seus descendentes quando pronunciam palavras como *rua* e *carro*. É a pronúncia do locutor de tv que diz *R-R-R-R-R-Ronaldo!* ou do apresentador de tv que diz *Errrrrrrou!* A vibrante coronal é transcrita com [r].

Um outro grupo de consoantes é o das que são produzidas fechando a cavidade oral e deixando o ar passar pela cavidade nasal: as consoantes nasais. Há três no português: o [m] presente em *mãe* e *cama*, o [n] presente em *nada* e *ano*, e o [ɲ] presente em *unha*. Elas se distinguem pelo ponto em que realizada a oclusão da cavidade oral. O [m] é produzido fechando a passagem do ar nos lábios, sendo labial. O [n] é produzido com a coroa da língua fechando a passagem na frente da cavidade oral, sendo coronal. O [ɲ] é produzido com o corpo da língua fechando a passagem no palato duro, a parte da frente do céu da boca. É controverso na literatura se o mais adequado é considerar essa consoante como coronal. Ela pode ser também considerada ao mesmo tempo coronal e dorsal, justamente por envolver o meio da língua. Há na verdade outra pronúncia possível desta última nasal, a qual será discutida na seção 3.6.

Outra consoante nasal bastante comum, a qual até ocorre na pronúncia de alguns falantes do português em determinadas posições, é a nasal dorsal [ŋ]. Ela pode ocorrer na pronúncia de parte dos falantes em palavras como *som*, pronunciada ['sõŋ]. Ela é comum nas línguas germânicas, sendo grafada <ng>. Exemplos: o inglês *long* ['lɒŋ] e o alemão *lang* ['laŋ], ambos significando 'comprido'.

Na realidade, em algumas palavras inglesas essa grafia representa dois sons em sequência, o [ŋ] e o [g], mas em muitas ele é um dígrafo que representa a nasal dorsal. Assim como nosso <nh> representa um único som e não dois em sequência, o <ng> em palavras como *sing* ou *song* representa uma única consoante nasal. Uma maneira de produzir esse som é se preparar para pronunciar o <g> no final de *sing* ou *song* mas parar logo antes de pronunciar o [g] propriamente dito. Há grandes chances de produzirmos uma nasal dorsal se fizermos assim.

O último tipo de consoante encontrado no português é aquele em que a língua entra em contato com o céu da boca, mas não interrompe a passagem do ar, porque ele continua passando pelo lado. São as consoantes laterais, ambas coronais. A que todos os falantes do português utilizam é o [l], presente em palavras como *lado* e *bola*. A outra consoante lateral não é utilizada por todos os falantes do português, embora talvez seja utilizada pela maioria. Ela também é coronal, mas é mais recuada. É a consoante grafada <lh>, transcrita [ʎ], presente em palavras como *filho* e *falha*. Eu disse que nem todos os falantes utilizam essa

consoante na pronúncia, porque há falantes que sistematicamente produzem a semivogal [j]. A palavra *falha* pronunciada “*faia*” seria um exemplo.

Olhando a tabela do Alfabeto Fonético Internacional, vemos que ela contém uma linha de consoantes inexistentes no português: as fricativas laterais. Discutirei essa linha da tabela quando chegarmos no capítulo 5, em que tratarei da fonética, mais especificamente na seção 5.4.

2.3 DETALHANDO E EXPANDINDO

Nesta seção exploraremos algumas consoantes inexistentes ou incomuns no português, pelo menos na maior parte de suas variedades, e veremos uma classificação mais detalhada das consoantes, especialmente no que se refere ao ponto de articulação. Dependendo do interesse de cada leitor(a), esta sessão pode ser pulada neste momento e lida só num momento posterior.

É conveniente examinarmos essas consoantes partindo de seu ponto de articulação. Como vamos nos aprofundar um pouco daqui para a frente, é importante consultarmos a tabela de consoantes pulmonares ou pulmônicas, que destaquei da tabela completa do alfabeto fonético internacional. Na tabela, o termo plosiva pode ser considerado equivalente a oclusiva. A rigor, pode ser feita uma distinção, segundo a qual plosivas são apenas as oclusivas pulmonares. Veremos outros tipos de oclusivas na seção 2.7.

CONSOANTES (PULMÔNICAS) © 2019 IPA

	Bilabial	Labiodental	Dental	Alveolar	Pós-alveolar	Retroflexo	Palatal	Velar	Uvular	Faringal	Glotal
Plosiva	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Vibrante				r					ʀ		
Tap ou flap		ɸ		ɾ		ɽ					
Fricativa	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Fricativa lateral				ɭ ɮ							
Aproximante		ʋ		ɹ		ɻ	j	ɰ			
Aproximante lateral				ɭ		ɮ	ʎ	ʟ			

Os símbolos à direita de uma célula são vozeados, à esquerda são não vozeados. Áreas sombreadas denotam articulações julgadas como impossíveis.

Quadro 1. As consoantes pulmônicas ou pulmonares.

As colunas da tabela mostram os pontos de articulação e as linhas os modos de articulação. Em cada célula, os símbolos à esquerda representam consoantes surdas e os da direita, consoantes sonoras. Localizando o símbolo ʒ, por exemplo, vemos que ele representa uma

fricativa pós-alveolar sonora. Assim podemos identificar a classificação de qualquer uma dessas consoantes consultando a tabela.

Dos 11 pontos de articulação presentes na tabela, os dois primeiros, bilabial e labiodental, contêm consoantes labiais. Os 5 seguintes, dental, alveolar, pós-alveolar, retroflexo e palatal, são coronais. As consoantes palatais, como já foi dito, podem ser consideradas ao mesmo tempo coronais e dorsais. Os dois pontos de articulação seguintes, velar e uvular, são dorsais. Faringal e glotal não se situam na cavidade oral, abrangendo, respectivamente, os sons produzidos na faringe (garganta) e nas pregas vocais.

LABIAIS

Dentre os pontos de articulação utilizados no português, veremos primeiro as bilabiais. Além de [p], [b] e [m], três consoantes não existentes no português pertencem a esse grupo. Começamos pelas fricativas. Nossas fricativas labiais [f] e [v] são labiodentais. São menos comuns nas línguas em geral, mas existem fricativas realizadas não com a aproximação do lábio inferior e dos dentes superiores, mas simplesmente com a aproximação dos lábios inferior e superior. Provavelmente as fricativas bilabiais são menos comuns por serem menos perceptíveis do que as labiodentais. Os símbolos fonéticos correspondentes são as letras gregas fi minúscula para a fricativa surda (ϕ) e beta minúsculo para a sonora (β). Elas podem ser surdas, como no japonês *tōfu* [to:ϕu] ‘tofu’, ou sonoras, como no espanhol *nuevo* [ˈnweβo] ‘novo’. Se se considerar importante fazer a distinção entre a fricativa e a aproximante, pode-se usar o diacrítico [ɸ], visto quando discuti a diferença de pronúncia entre o [ð] do inglês e o do espanhol. Dessa forma, a fricativa bilabial sonora fica sendo transcrita [β] e a aproximante correspondente como [ɸ]. Numa transcrição mais estrita, o exemplo do espanhol, que contém uma aproximante, seria transcrito, portanto, [ˈnweɸo].

Uma das pouquíssimas línguas que distinguem fricativas labiodentais e bilabiais é o evé (Gana). Nele temos exemplos como [vò] ‘terminar’ e [βò] ‘porta’, além de [fè] ‘dívida’ e [ɸè] ‘ano’. É muito comum, nas línguas em geral, que a fricativa bilabial sonora varie com uma aproximante bilabial sonora ou se torne aproximante.

Uma outra consoante bilabial, muito rara, é a vibrante bilabial, cujo símbolo é um em versalete, ou seja, tem o formato da maiúscula, mas é do tamanho de uma letra minúscula (ɂ). Ela ocorre no pirahã (Amazonas) e em algumas línguas esparsas em todos os continentes, frequentemente precedida de nasal bilabial. Um exemplo precedido de nasal é [mɂ́] ‘cachorro’ em medumba (Camarões). Um exemplo não precedido de nasal é a palavra [ʔààɂóí] do pirahã, que ocorre em variação com [ʔààbóí] como a pronúncia de ‘lago’.

O segundo ponto de articulação labial é o labiodental, no qual os lábios inferiores tocam ou se aproximam dos dentes superiores. Em português temos o [f] e o [v]. Entre as consoantes labiodentais não presentes no português, encontramos primeiro uma nasal. Ela ocorre antes de [f], como no alemão *fünf*, no inglês *comfort* e no espanhol *influir*. Seu símbolo é um [m] com um prolongamento e curva para dentro no final: [ɱ]. Existe também o *flepe* ou *flap* labiodental. Por agora, basta a informação de que um flepe é produzido por um órgão que resvala de passagem por outro. Ele é usado, por exemplo, em mono (Congo-Kinshasa). Um símbolo que pode ser usado para transcrever esse fone é um [v] com uma curva para fora no final: [v̥]. Outra forma de transcrevê-lo é por meio do diacrítico *breve*, como faz Olson (2001), segundo o qual há variação livre entre o flepe bilabial e o labiodental na língua mono. A palavra ‘quente’ pode, então, ser pronunciada [w̥égē], com flepe bilabial ou [w̥égē], com flepe labiodental. Detalho a diferença entre *tap* e *flap*, além de propor termos portugueses para os dois na seção 5.8.8.

Por fim, um som bastante comum é a aproximante labiodental transcrita como [ʋ]: ela ocorre no holandês (ou neerlandês, já que a Holanda é só parte dos Países Baixos), como em *wollen* [ˈʋɔlə(n)] ‘querer’. Aparentemente, o que denominamos fricativa labiodental sonora no português com frequência pode ser realizada como aproximante, não sendo acompanhada de ruído.

CORONAIIS

Os pontos de articulação coronais são os mais variados, devido à grande flexibilidade da coroa da língua. Como mencionado, eles são cinco: dental, alveolar, pós-alveolar, retroflexo e palatal.

Das coronais, as únicas inexistentes no português na fala considerada normal (sem desvios fonaudiológicos) são as fricativas laterais. Como elas são produzidas é algo que será discutido na seção 5.4. Por ora, basta dizer que se, na produção de uma lateral, o estreitamento da passagem do ar for muito acentuado, o som produzido acaba sendo uma fricativa. Uma língua que tem tanto a fricativa lateral surda quanto à sonora é o zulu (África do Sul). A rigor, a sonora tem voz frouxa (*slack voice*). O símbolo fonético da fricativa lateral surda é o [ɬ] e o da sonora é o [ɮ]. A partir deste ponto, apresento a transcrição completa das palavras citadas. Como os símbolos vocálicos ainda não foram vistos, eles só serão identificados por completo após a leitura do próximo capítulo. Exemplos em zulu:

(3) *hlala* ‘morar’ [ˈɬaːla] *dlala* [ˈɮaːla] ‘jogar’

Dos pontos de articulação pouco usados ou não usados no português, o primeiro abrange as consoantes retroflexas. Na transcrição, o símbolo usado para representar as consoantes retroflexas sempre termina em um linha descendente que se curva para o lado direito: [ʈ ɖ ɳ ʀ ʁ ʑ ʐ ɭ].

A única consoante retroflexa com ocorrência numa região abrangente do Brasil é a aproximante. Como já vimos neste capítulo, há duas variedades de aproximantes retroflexas no Brasil: uma levemente retroflexa e outra mais acentuada. Podemos usar essa consoante de retroflexão acentuada como base para produzir as outras. Normalmente quando um falante do português brasileiro que usa o retroflexo pronuncia em seguida uma outra consoante coronal, sua língua volta para a posição alveolar ou dental. Todavia, se em vez de fazer isso, ele mantiver a língua na posição retroflexa, todas essas consoantes coronais passarão a ter um ponto de articulação retroflexo. Usando esse artifício, palavras como *perto*, *gordo*, *perna*, *orla*, *força* e *barzinho* teriam na segunda sílaba as seguintes consoantes: [ʈ ɖ ɳ ʀ ʁ ʑ ʐ]. A figura 4 ilustra duas configurações possíveis na articulação de consoantes retroflexas: a região da língua que entra em contato com o céu da boca pode ser o ápice, como em híndi, ou a parte inferior, como em tâmil.

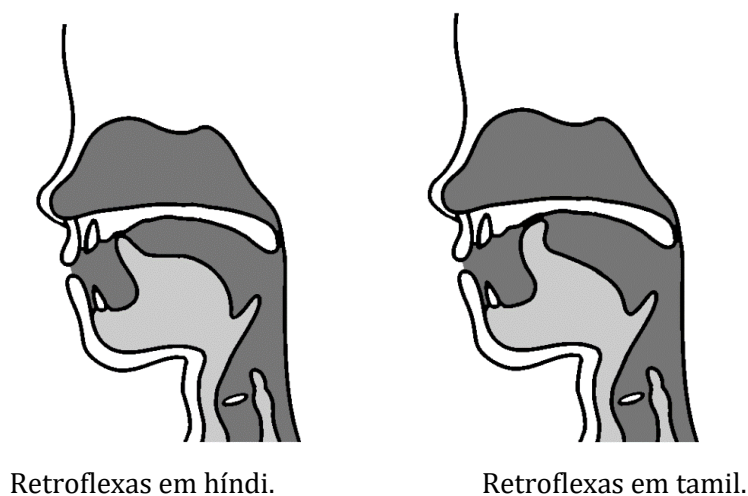


Figura 4.

Várias dessas consoantes são muito comuns em línguas do subcontinente indiano. Na Índia propriamente dita, por exemplo, tanto as línguas indoarianas quanto as dravídicas em geral possuem várias dessas consoantes retroflexas. Na lista que acabamos de ver, temos duas oclusivas, uma nasal, uma lateral e duas fricativas. Existe ainda um tepe retroflexo, transcrito [ɻ]. Seguem exemplos do hîndi, do mandarim e do polonês:

- | | | |
|------------------------|-------------------|---------|
| (4) [ʈa:l] ‘empurrão?’ | [ɖa:l] ‘galho’ | (hîndi) |
| [ʈʰeɳɖa:l] ‘frio’ | [bəɻa:l] ‘grande’ | |

- | | | | |
|-----|----------------------------------|----------------------------|------------|
| (5) | [ʃu] ‘livro’ | [zɛn] ‘pessoa’ | (mandarim) |
| | [tʂ ^h əŋ] ‘completar’ | [tʂuŋ] ‘centro’ | |
| (6) | [ʃus.ti] <i>szósty</i> ‘sexto’ | [ʒɛ.ka] <i>rzeka</i> ‘rio’ | (polonês) |
| | [ʔʂas] <i>czas</i> ‘tempo’ | | |

No sul da Itália também ocorre geminada ou dobrada a oclusiva sonora retroflexa, como na palavra [ka'vaɖɖu] ‘cavalo’, em siciliano.

O segundo ponto de articulação de que vou tratar dentro das coronais é o palatal. Claro que o português tem consoantes palatais. Elas ocorrem nas palavras *filha*, *manha* e *praia*. Mas existem outras importantes de serem mencionadas. Em primeiro lugar as oclusivas. Elas não são tão comuns porque têm uma grande facilidade de passarem a africadas pós-alveolares, mas existem como fonemas, por exemplo, no húngaro, nas palavras *kutya* ['kucɔ] ‘cachorro’ e *magyar* ['mɔɟɔr] ‘húngaro’. Aparentemente elas existem em variação no francês, em palavras como *tiens* ['tjɛ̃] ou ['cɛ̃] e *Dieu* ['djø] ou [ʔø], embora seja difícil encontrar referências na literatura. A configuração do trato vocal na articulação de uma oclusiva palatal aparece na figura 5. O alemão tem a fricativa palatal surda, transcrita com um [ç], em palavras como *ich* ['ɪç] ‘eu’. O sueco tem a fricativa sonora em sílaba tônica inicial: *jaga* ['ja:ga] ‘caçar’. Ela também é uma das variantes da pronúncia de palavras como *yo* em espanhol. O ponto exato da língua que entra em contato com o palato varia.



Figura 5.

Um ponto de articulação distinto, que aparece fora da tabela de consoantes pulmonares, chamado de alveolopalatal, ocorre em várias línguas. São fricativas auditivamente muito próximas de [ʃ] e [ʒ], a tal ponto que em muitos trabalhos são transcritos com esses símbolos,

quer para simplificar, quer por não ter sido identificada a pronúncia exata. Os símbolos são: [ɕ] e [ʐ], surda e sonora, respectivamente. Vejamos alguns exemplos:

- (7) sueco: tjej ['ɕej:] 'menina'
 polonês: zima ['ʐima] 'inverno' musieć ['muɕetɕ] 'ter que'
 japonês: [d̞ʑi'ɕo] 'dicionário' [ɕi'goto] 'trabalho'

Os símbolos [ˈ] e [ˌ] indicam aqui o padrão tonal das palavras em japonês, e serão explicados na seção **Erro! Fonte de referência não encontrada..** As figuras 6 e 7 contrastam os pontos de articulação pós-alveolar e alveolopalatal.



Figura 6. Fricativa pós-alveolar

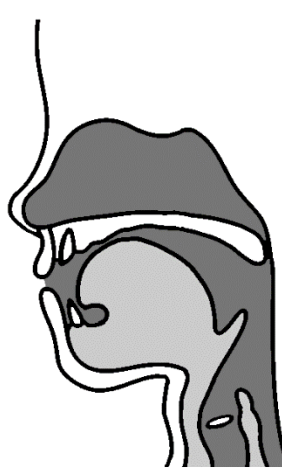


Figura 7. Fricativa alveolopalatal

DORSAIS

Chegando às consoantes dorsais, temos primeiro a nasal velar, que se pode considerar que ocorre no português como segmento ou como parte de um segmento mais complexo, mas sujeita a variação. A nasal velar ocorre de forma mais consistente em línguas românicas, como o espanhol e o italiano, e mais extensamente nas línguas germânicas em geral. Além dos exemplos já vistos do inglês e do alemão, temos o espanhol *cinco*, que é ['θiŋko] na pronúncia da Espanha e ['siŋko] na latinoamericana, e *cinque* ['tʃiŋkwe] no italiano, e ainda no galego *un*, *unha* ['uŋ], ['uŋɐ] 'um, uma' e em piemontês *italian* [itali'aŋ]. Em línguas de outras regiões ela ocorre também em início de palavra, como no cantonês [ŋɐ:ŋ] e no xangaiense [ŋɐ:ŋ] 'dente'. No português, ela pode ocorrer diante de consoantes velares, sendo ainda uma das possibilidades de pronúncia da fase final de vogais nasais, como em *som*, pronunciada ['sõŋ].

Existem ainda duas consoantes velares inexistentes no português: a aproximante velar e a aproximante lateral velar.

A aproximante velar ocorre no grego cipriota, como no exemplo [mauʔa'zi] 'loja', onde no grego há uma fricativa velar. Também ocorre no coreano, apenas diante de [i], como em [uʔise] 'médico'. Esse fone pode ser descrito como semelhante a um [w], mas sem ter arredondamento dos lábios.

A aproximante lateral velar é uma das possibilidades de pronúncia da lateral em coreano antes de consoante velar, como no exemplo [tɛɭɡud̪zi] 'carroça'.

O ponto de articulação uvular é pouco comum nas línguas em geral. Em algumas línguas da Europa existe a vibrante uvular, normalmente em variação com outro tipo de som. É o chamado <r> *grasseyé* do francês, que tem como exemplo prototípico a pronúncia de Edith Piaf, ocorrendo três vezes em "je ne regrette rien". Esse som, transcrito [ʀ], também é usado no alemão. No entanto, é muito comum ele deixar de ser uma vibrante e passar a ser uma fricativa, transcrita [ʁ], se se mantiver sonora.

A oclusiva uvular surda [q] existe em árabe como uma das possíveis pronúncias da consoante *qaf*. Uma língua que faz uso extenso de consoantes uvulares é o quêchua. O tlingit na América do Norte também. No final do capítulo há um link para um vídeo dessa língua. Existem também a oclusiva velar sonora, mais rara ainda, e também a nasal e as fricativas. A figura 8 ilustra uma oclusiva uvular.

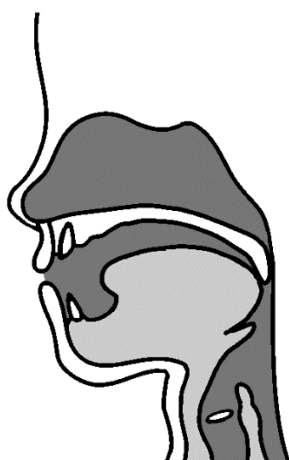


Figura 8.

A oclusiva sonora é transcrita [g], a nasal é transcrita [ŋ]. As fricativas, que também vimos na seção 2.4, são transcritas assim: a surda usa a letra grega *qui* ou *khi* [χ], e a sonora usa um *r* maiúsculo de cabeça para baixo: [ʁ]. As duas últimas ocorrem em variação no português brasileiro e, segundo Rennicke e Martins (2013), são as variantes mais frequentes do arquifonema /R/ no português europeu (trato dos arquifonemas na seção 4.4). A observação de que vários deles são letras em versalete pode facilitar lembrar dos símbolos: [ʀ ʁ ɣ ŋ]. Seguem alguns exemplos:

- (8) [gæh've] 'café' (persa)
 (9) [niho'n] 'Japão' (japonês)
 (10) [pɔχte] 'porta'
 (11) [ma'ʁō] 'marrom'

Exemplos do tlingit:

- (12) [qù:tʃʰ] 'lobo' [q'atʰé:tʃ] 'jarra' [qʰákʰw] 'cesto'
 [ná:qʰw] 'polvo' [té:qʰw] 'red rockfish, red snapper'
 [χá:tʰ] 'peixe, salmão' [χ'a:n] 'fogo'
 [qʷéjnà:] 'toalha' [qʷátʃ] 'panela' [qʰwá:n] 'pessoas de um lugar'
 [ʃanaχʷá:jɪ] 'machado' [χʷá:tʰ] 'truta'
 [tu 'húnχʷ] 'o irmão mais velho dele' [néχʷ] 'yellow cloudberry'

ALÉM DA CAVIDADE ORAL

O último ponto de articulação que veremos é o faringal. As consoantes faringais são muito raras. O árabe é uma das poucas línguas que faz um uso extenso da cavidade faríngea na produção de sons consonantais. Ele tem duas fricativas faringais, a surda [ħ] e a sonora [ʕ]. A sonora tende, nas línguas em que ocorre, a ser uma aproximante, ou seja, não envolver ruído, embora apareça na linha das fricativas na tabela de consoantes.

A fricativa faringal surda ocorre também em variação em galego, o que caracteriza o tipo de pronúncia denominado *gheada*. Exemplos do árabe são:

- (13) [ħam'ma:m] 'banho' [mu'ħamməd] 'Mohammed'
 (14) [ʕu:d] 'madeira, alaúde' [ʕajn] 'olho'

Exemplos do galego, que ocorrem em variação:

- (15) galego [ħa'leħu] magro [maħru]

A figura 9 ilustra o recuo da raiz da língua na produção de uma consoante faringal. O tracejado indica sua posição normal. O traço inteiro, sua posição na produção dessas consoantes.

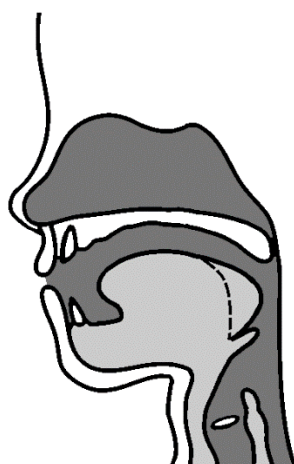


Figura 9. Posição da língua na produção de consoantes faringais.

Trato das consoantes tradicionalmente denominadas enfáticas do árabe, que normalmente envolvem faringalização, na seção 2.9.

2.4 NOSSOS ERRES

Uma dificuldade grande na hora de transcrevermos pronúncias do português é a representação dos sons que na grafia correspondem a um <r> ou a um dígrafo <rr>. Na nossa cultura, em que predominam pessoas alfabetizadas, a observação das letras usadas para grafar as palavras é muito mais difundida do que o exame dos sons utilizados nas línguas. Uma mesma grafia pode corresponder a um grande número de pronúncias diferentes. O caso mais extremo no português é justamente o da letra <r> e do dígrafo <rr>. A palavra *mar* pronunciada isoladamente por falantes nascidos no estado de São Paulo, onde moro, pode ter quatro realizações diferentes. Pode ser pronunciada [r], com uma única batida da ponta da língua na região dos alvéolos, ou seja, como um tepe. Pode ser pronunciada [r], com algumas batidas da ponta da língua na região dos alvéolos, ou seja, como uma vibrante. Pode ser pronunciada [ɹ], com a língua um tanto recuada e curvada, ou também ser pronunciada [ɹ], com a língua bastante recuada e curvada. Uma característica importante que distingue as duas retroflexas do tepe e da vibrante alveolar é que não há contato da língua com a parte superior da cavidade oral.

Essa consoante final pode ainda ser pronunciada como [ʀ], uma vibrante uvular, produzida, portanto, na parte de trás da cavidade oral. Essa variante não é muito comum hoje, mas provavelmente foi através dela que começou o processo de posteriorização do [r]: a vibrante coronal se tornou uma vibrante dorsal, o [ʀ], e a partir dele surgiram as fricativas dorsais, tais como o [x]. Essa fricativização do [ʀ] ocorre também no francês, mas mantendo

o ponto de articulação uvular, seja surda, transcrita [χ], seja sonora, transcrita [ʁ]. As fricativas dorsais que são realizações da consoante escrita <rr>, ou <r>, por exemplo, em início de palavra, podem ter um de dois pontos de articulação vizinhos, e ser surdas ou sonoras. Se forem surdas, a realização mais comum de fricativas em fim de palavra no PB, serão o [x], fricativa velar surda, ou o [χ], fricativa uvular surda. Essa variação ocorre na pronúncia carioca, por exemplo. As variantes sonoras ocorrem mais em posição interna de palavra, antes de consoante sonora, como em *perna*, podendo ser pronunciadas como fricativa velar sonora, transcrita como [ɣ], ou como fricativa uvular surda, transcrita como [ʁ].

Os últimos estágios desse processo de posteriorização do [r] são as fricativas glotais ou laringais. É a pronúncia da palavra *mar* em Belo Horizonte, por exemplo, onde a consoante final é pronunciada como [h] ou [ɦ]. Para os falares do PB que usam essas consoantes glotais, uma característica que a distingue de todas as outras variantes do [r] é o fato de ela não ser articulada na cavidade oral. Se o falante faz algum movimento na cavidade oral ao pronunciar essa consoante, não está pronunciando um [h]. Pronunciando a palavra *mar* com um [h] ao final, a configuração da cavidade oral tem que permanecer a mesma durante a produção do [a] e do [h] que a segue, permanecendo, portanto, bem aberta. Essencialmente o que muda é só o fato de durante o [a] as pregas vocais vibrarem plenamente, enquanto no [h] há o que se denomina aspiração, uma espécie de sopro, uma configuração diferente das pregas vocais, que será vista nas seções 5.7 e **Erro! Fonte de referência não encontrada.** Podemos considerar que nesse exemplo, temos um [a] normal, vozeado, seguido de um [a] desvozeado, aspirado. Isso poderia ser transcrito assim [ma̤], em que o círculo abaixo do segundo [a] indica que ele está ensurdecido. Da mesma forma com vir, [vi̤], e com abajur [aba'ʒṳ]. Podemos observar isso mais claramente produzindo as sequências [aha], [ihi], [uhu], ou quaisquer outras em que tenhamos a mesma vogal duas vezes com um [h] intercalado. Se o som pronunciado entre as vogais for realmente uma glotal, não haverá modificação nenhuma na configuração da cavidade oral. Portanto, se há uma diferença sonora ela deve ser proveniente de outra região e não da cavidade oral. Poderíamos transcrever essas sequências, presentes em palavras como *amarrar*, *irritar* e *surrupiar*, de outra forma, que seria, no entanto, menos intuitiva e mais difícil de visualizar: [a̤a], [i̤i] e [ṳu].

Resumindo as possíveis realizações correspondentes ao que é grafado como <r> em português são [ɾ ɾ̥ ɽ ɽ̥ ʀ x ɣ χ ʁ h ɦ]. Uma observação importante a respeito desse conjunto é que ele não é uma classe natural, no sentido que veremos no capítulo 6. Não há nenhuma característica que esteja presente em todos esses fones, a não ser o fato de serem consonantais, o que é muito pouco, e incluiria várias outros fones. Ainda não vimos os traços

distintivos, o que será feito no capítulo 6, mas podemos tentar agrupar esses fones de acordo com as propriedades articulatórias que já vimos.

O articulador primário não é uniforme para esse conjunto de sons. [r r̥ ɹ ɻ] são coronais, [ʀ x ɣ χ ʁ] são dorsais, e [h ɦ] são glotais. [r r̥ ɹ ɻ ʀ ʁ ɦ] são sonoras, embora a última, a rigor, seja murmurada (v. cap. 5), e [x χ h] são surdas. Mesmo que se argumente que todas são contínuas (v. cap. 6), isso não seria ponto pacífico com relação ao tepe e à vibrante. Em suma, provavelmente seria melhor definir essa classe como tendo semelhança de família, no sentido de que se tomarmos dois fones, por exemplo, o tepe e a vibrante, eles são muito parecidos, mas há fones que são extremamente diferentes um do outro, como o tepe e a aspiração, não tendo nada em comum um com outro: nem ponto de articulação, nem modo, nem o estado da glote. Esse pode ser considerado um argumento a favor de teorias como a linguística cognitiva, que defendem que nossas categorias mentais, e linguísticas, não precisar ser baseadas em características compartilhadas por todos os seus membros, podendo se basear numa semelhança de família, como escreveu Wittgenstein.

2.5 FONES GLOTAIS: CLASSIFICAÇÕES FONÉTICA E FONOLÓGICA DISTINTAS

A classificação fonética, mais concreta, e a fonológica, mais abstrata (v. caps. 4 e 5), de alguns sons pode diferir. É o que se dá com alguns sons glotais. Tanto a parada glotal quanto a aspiração podem ser classificadas de forma diferente na fonologia se comparada com a fonética. Foneticamente, a parada glotal [ʔ] pode ser considerada uma oclusiva, e a aspiração [h] pode ser considerada uma fricativa, pois na primeira há uma interrupção total da passagem do ar, e na segunda ele passa por uma constrição bastante estreita, produzindo ruído. Porém, o fato de esses dois sons serem articulados nas pregas vocais acaba sendo um fator que faz com que eles tenham um comportamento fonológico diferente.

Predominantemente, os sons denominados oclusivas e fricativas têm uma constrição na cavidade oral. Como o [ʔ] e o [h] são produzidos nas pregas vocais, eles são usados de maneira diferente. Ladefoged e Maddieson argumentam que, por isso, eles se assemelham a vogais. Em Chomsky e Halle (1968), eles foram classificados como glides II (*glide* é um termo inglês equivalente a semivogal, talvez aludindo a uma espécie de deslizamento na produção desses sons). Ambos são usados por vezes para evitar hiato, como as semivogais propriamente ditas, [j] e [w]. Vejamos alguns exemplos.

Em holandês, o uso da parada glotal para eliminar hiatos depende das vogais envolvidas e do acento. Ela é usada, por exemplo, entre um [a] e uma vogal tônica, como vemos nos seguintes exemplos. O símbolo ['] indica que a sílaba que vem logo em seguida é a tônica.

(16) chaos [ˈχa:ɔs] ‘caos’

chaotisch [χa'ʔowtis] 'caótico'

Em holandês, um hiato em que a segunda sílaba é átona é tolerado. Mas se a segunda sílaba for tônica, ocorre a inserção de um [ʔ] no início dela.

Também em irlandês, há vários exemplos em que é inserido um [h] em contextos de hiato, como, por exemplo após o artigo definido plural *na* [nə]:

(17)	'o caminho'	an bealach [ə(n) 'bʲalʲəx]
	'os caminhos'	na bealaí [nə 'bʲalʲi:]
	'a idade, o século'	an aois [ənʲ 'i:]
	'idades, séculos'	aoiseanna ['i:]ənʲə]
	'as idades, os séculos'	na haoiseanna [nə 'hi:]ənʲə]

Observe que antes de consoante, o artigo plural *na* é seguido diretamente da consoante inicial do plural pura e simplesmente, mas diante de vogal, há introdução de um [h]. O singular de 'idade, século' é *aois* ['i:], começando por vogal. O plural não precedido do artigo definido *na* é *aoiseanna*, mas *haoiseanna* diante do artigo.

Também em abruzzese, língua regional da província de Abruzzi, geograficamente no centro da Itália, mas linguisticamente pertencente ao grupo itálico meridional, é a fricativa glotal sonora, mais precisamente murmurada, que é usada para evitar hiato, como se vê num exemplo de Rohlfs (1949) da cidadezinha de Bellante: o verbo proveniente do latim *creāre* 'criar', é pronunciado [krə'hæ].

2.6 CONSOANTES NÃO PULMONARES

As consoantes que vimos até agora são todas pronunciadas com uma corrente de ar proveniente dos pulmões. Mas há outros mecanismos de corrente de ar que também são empregados linguisticamente na produção de sons, embora isso não ocorra no português.

2.6.1 CONSOANTES COM CORRENTE GLOTTAL

Se a passagem do ar pela laringe ficar aberta pelo menos intermitentemente, podemos ter uma corrente de ar pulmonar, que é a mais comum. Todas as línguas utilizam a corrente pulmonar e, mesmo entre as línguas que utilizam outros tipos de corrente de ar, a pulmonar é sempre a mais comum. Mas é possível fechar totalmente a laringe e, mesmo assim, produzir sons usados linguisticamente. Como gerar corrente de ar se não vem ar dos pulmões? Há duas possibilidades: a laringe pode ser elevada ou abaixada. Se ela for elevada,

isso empurrará o ar para fora, e será produzida uma consoante ejetiva. Se ela for abaixada, isso puxará o ar para dentro, e será produzida uma consoante implosiva. Vejamos cada um desses tipos separadamente. Retomarei as consoantes com corrente glotal na seção 5.3.

2.6.1.1 EJETIVAS

Há diversas regiões do mundo em que ocorrem consoantes ejetivas. Elas são muito comuns em línguas do Cáucaso, bem como na região da América próxima ao Oceano Pacífico, o que inclui tanto as Montanhas Rochosas na América do Norte quanto os Andes na América do Sul, além da Mesoamérica (sul do México e parte norte da América Central continental), todas elas regiões montanhosas.

As consoantes ejetivas são transcritas com uso do apóstrofo ou de uma parada glotal sobrescrita. As mais comuns são as oclusivas, que podem ser, portanto, transcritas como [p' t' k'] ou como [p^ʔ t^ʔ k^ʔ].

O quêchua e o aimara, falados na Bolívia e países vizinhos, apresentam essas consoantes.

Alguns exemplos de oclusivas ejetivas do aimara:

- | | | |
|------|---------------------------------|----------------------------|
| (18) | k'ata [k'ata] 'veloz' | p'itata [p'itata] 'tecido' |
| | q'ala [q'ala] 'tudo, terminado' | t'ula [t'ula] 'lenha' |

2.6.1.2 IMPLOSIVAS

As consoantes implosivas, produzidas com abaixamento da laringe, frequentemente são uma estratégia alternativa que possibilita que sejam produzidas obstruintes sonoras. Há línguas como o vietnamita, em que todas as oclusivas sonoras são implosivas. O símbolo usado para transcrever essas consoantes apresenta uma curva com final descendente na parte superior do símbolo fonético. Comparem-se [b d g] com [b̥ d̥ g̥].

Entre as línguas que possuem consoantes implosivas se incluem o kwazá (Rondônia), o cambodjano, o vietnamita, o zulu (África do Sul) e a língua maa dos massais (Quênia e Tanzânia). Seguem exemplos do kwazá, em que o [b] varia com [b̥], o qual, embora menos frequente, ocorre mais na fala rápida. O [d] é raro, e não ocorre em sílaba tônica:

- (19) [b̥ɔ̃'r̥ɔ̃ki] ~ [b̥ɔ̃'r̥ɔ̃ki] 'fura/furam/furou/furaram'
[bwa] ~ [bwa] 'termina/terminam/terminou/terminaram'
[dajki] 'pega/pegam/pegou/pegaram'
[dutu're] ~ [dutu're] 'porco'

Alguns exemplos do vietnamita, com indicação dos tons (v. cap. 13):

(20) ba [ɓaːɬ] ‘três’

ɗi [ɗiːɬ] ‘ir, vir’

2.6.2 CONSOANTES COM CORRENTE VELAR: OS CLIQUES

Uma última possibilidade de produzir consoantes é aquela que envolve uma corrente de ar gerada apenas dentro da cavidade oral. Se elevarmos a parte posterior da língua até encostar no palato mole e, ao mesmo tempo, fizermos contato com a parte anterior da língua em outra região da cavidade oral, podemos gerar uma corrente velar. São os chamados cliques. É extremamente raro que eles sejam usados linguisticamente. Em geral, como diz Laver (1994: 175), eles têm um uso paralinguístico, como a parada glotal em português (v. seção 2.2). Podemos até transmitir significado com eles, mas eles não fazem parte de palavras na maior parte das línguas, sendo usados isoladamente, como o clique dental ou alveolar discutido abaixo nesta seção.

Cito alguns exemplos familiares de cliques. O primeiro é próximo de um beijo dado com som, um beijo estalado. Esse é clique bilabial, que é representado pelo símbolo [⊙]. Um exemplo da língua ʼamkoe (Botswana) é [⊙oa] ‘dois’. Como os cliques sempre têm uma oclusão velar, eles podem ser transcritos precedidos de um [k]. Esse mesmo exemplo poderia ser transcrito [k⊙oa].

O clique dental ou alveolar, transcrito [|] equivale ao som usado para expressar a negação em várias línguas europeias: grafado *tsk* ou *tut* em inglês, *tst* ou *tsk* em francês, e expresso pelo verbo *cöcög* [ˈtsøtsøɡ] em húngaro. Observe que sempre há [t] ou [ts], sons muito próximos do clique dental/alveolar, já que têm o mesmo ponto de articulação desse clique. Ele ocorre em exemplos como o zulu *icici* [iːlíːli] ‘brinco (substantivo)’, que também pode ser transcrito [iːklíːkli].

O clique lateral, transcrito [||] equivale ao som utilizado para incitar um cavalo. Um exemplo é o verbo *xoxa* em zulu, pronunciado [||óːlla], podendo também ser transcrito [k||óːklla]. O nome da língua xhosa tem um clique lateral aspirado: *isiXhosa* [isill^hosa] ou [isikll^hosa].

O clique que é descrito como palatal apresenta contato da língua com uma região grande do céu da boca, mas Ladefoged e Traill consideram que o principal contato é na região palatal. Ele é transcrito [†]. Ele está presente no nome da língua ʼhaba (Botswana), que pode ser transcrito [†^habá] ou [k†^habá].

O último clique é o pós-alveolar, transcrito [!]. Esse clique é produzido com a língua formando uma concavidade e sendo puxada para baixo, e não para trás como nos outros cliques. Exemplos: *hoqoqa* [ho!o!a] ‘conversar’ em sotho e *iqanda* ‘ovo’ [i!anda] em xhosa.

Entre as línguas que têm cliques, algumas têm dezenas deles, porque podem combinar esses vários pontos de articulação com vozeamento, nasalidade, aspiração, entre outras

propriedades acústicas. No ʼamkoe, ‘um’ é [õ̃]. Como ele é um clique nasal, uma alternativa de transcrição é ele ser precedido de um [ŋ]: [ŋõ̃].

2.7 CONSOANTES COMPLEXAS

Algumas línguas possuem consoantes com ponto de articulação duplo. Na verdade, o próprio português tem uma semivogal com ponto de articulação duplo: o [w].

Se atentarmos, veremos que um som como o [w] envolve duas regiões distintas da cavidade oral em sua produção. Por um lado, temos a língua erguida em direção ao palato mole, o que faz com que esse seja um som dorsal. Mas, ao mesmo tempo, sua produção envolve o arredondamento dos lábios, o que o torna um som labial. Ou seja, o [w] é um som labiovelar, um som complexo, com dois pontos de articulação. Por isso ele não aparece na tabela de consoantes pulmonares do alfabeto fonético internacional, pois teria que estar em duas colunas ao mesmo tempo.

Diversas línguas da África apresentam outras consoantes que têm exatamente esses dois pontos de articulação. São oclusivas surdas [k̠p], sonoras [g̠b] ou nasais [ŋ̠m]. A ligadura usada em cima dos dois símbolos tem a finalidade de indicar que não se trata de uma sequência de consoantes, mas uma única consoante com dois pontos de articulação.

O iorubá apresenta os dois primeiros tipos, como vemos nos exemplos: *dúpé* [duk̠pɛ] ‘agradecer’ e *gbé* [gbe] ‘morar’. Como o iorubá não tem [p], e só [k̠p], na grafia deste se usa apenas a consoante <p>. Além da África, essas consoantes complexas ocorrem em outras regiões, como no vietnamita, em que se encontram exemplos em final de palavra, como *đũng* [đu̠ŋm̠ʔ] ‘correto’ e *khóc* [xaw̠k̠p̠ʔ] ‘chorar’.

2.8 CONSOANTES COM ARTICULAÇÃO SECUNDÁRIA

Os exemplos acima, de consoantes complexas, envolvem duas oclusões concomitantes, havendo fechamento da passagem do ar em dois pontos da cavidade oral na sua produção. Mas existem consoantes que são produzidas como que com uma vogal concomitante a elas. Se é uma articulação secundária vocálica, não há fechamento da passagem do ar nesses pontos. O mais comum é que se sobreponha o equivalente a um [i] ou o equivalente a um [u] a uma consoante. Voltarei a esse ponto mais adiante, mas essa é uma das pronúncias possíveis de *crepe* e *sapo*. Se a vogal final fica tão breve e sobreposta à consoante precedente que se torna indistinguível temporalmente dela, temos pronúncias como [ˈkreɸʲ] e [ˈsapʷ]. Há línguas que utilizam com frequência esse tipo de articulação, inclusive com valor distintivo.

Exemplos do gaélico escocês, em que a articulação secundária distingue o plural do singular:

- (21) [t̪u:n] 'fortaleza' [t̪u:nʲ] 'fortalezas'

Exemplos do tlingit, com distinção do significado lexical:

- (22) *yaak* [ja:k] 'molusco' *yaakw* [ja:kʷ] 'barco'

Outro tipo de consoante complexa é a série de consoantes coronais tradicionalmente chamadas de enfáticas do árabe, as quais são realizadas com faringalização concomitante à articulação coronal. No árabe padrão são quatro essas consoantes, as quais são transcritas com o diacrítico [ˤ] sobrescrito: [tˤ dˤ sˤ ʔˤ]. Exemplos do árabe:

- (23) [tˤɑ:lɪb] 'estudante' [sˤɑ'ʕi:r] 'jovem'
[dˤɑ:hija] 'subúrbio' [ʔˤanna] 'ele pensou'

A consoante pronunciada [ʔˤ] apresenta muita variação, podendo ainda ser pronunciada como [zˤ] ou como [ðˤ], ou seja, um [z] faringalizado ou um [ð] velarizado.

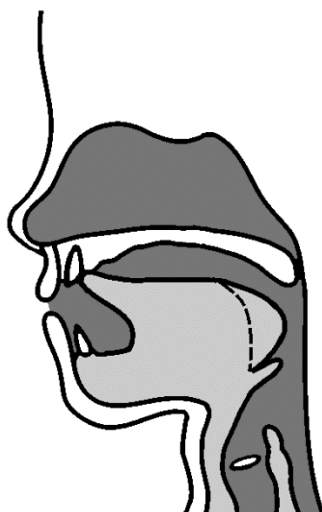


Figura 10.

2.9 EVITANDO CONFUSÕES

É bem possível que os símbolos de duas das consoantes nasais causem confusão. O [ɲ] representa uma nasal palatal. O [ŋ] representa uma nasal velar. Uma dica simples pode ajudar. Talvez tenha sido mero acaso, mas se compararmos a posição das colunas palatal e velar na tabela do alfabeto fonético internacional, que reflete os pontos de articulação como as ilustrações do aparelho articulador que temos visto, veremos que a coluna das palatais fica à esquerda da coluna das velares. Tanto o símbolo [ɲ] quanto o [ŋ] se baseiam no [n]

e têm um prolongamento para baixo que termina em curva para a esquerda. Pois bem, o prolongamento do símbolo [ɲ] fica à esquerda, e do símbolo [ŋ] fica à direita, o que corresponde exatamente à localização das colunas na tabela das consoantes. O outro símbolo que pode ser confundido com esses é o [ɳ], que tem a curva para o lado direito, como todos os símbolos de retroflexos. Se considerarmos que a parte de trás da cavidade oral fica à direita no diagrama e na tabela, esse é um bom mnemônico com relação ao símbolo dos retroflexos.

Outros dois símbolos consonantais que são confundidos com frequência, inclusive em textos de fonologia, são o símbolo [ɬ], em que há uma volta da linha que sai do [l], e o símbolo [ɮ], que pode ser descrito como um [l] com um til que o corta no meio. O primeiro é uma *fricativa lateral surda*, enquanto o segundo é o símbolo de uma *lateral velarizada*, como a que ocorre no português europeu e no inglês, quando ela está após a vogal na sílaba, como em *Portugal* e *ball* ou *film* em inglês. Ela também ocorre no Sul do Brasil, principalmente em falantes de mais idade, por exemplo na pronúncia do próprio nome do país: [bra'ziɮ]. No português europeu há pronúncias que usam a variante velarizada da lateral em qualquer contexto.

Outra fonte possível de confusão é o termo *implosivo*, o qual é usado com dois significados distintos em textos de fonética, fonologia, mas principalmente de linguística histórica. Adoto aqui o padrão atual, que é utilizar o termo no sentido visto na seção 2.6.1.2, ou seja, oclusivas produzidas sem aumento da pressão intraoral. O outro sentido, que evito, não se relaciona a características fonéticas. Há textos que denominam implosiva, por exemplo, a consoante [s] em *posto*. Esse uso se relaciona à posição na sílaba. Em posição na parte final da sílaba uma consoante é chamada, nesse uso, de implosiva. No início da sílaba, como em *sol*, o mesmo [s] é denominado explosivo. A mim parece pouco adequado esse uso, pois consoantes que, como esse [s], não envolvem oclusão não contém plosão, nem explosão, nem implosão.

2.10 ALGUNS DIACRÍTICOS

Nesta seção apresento alguns diacríticos comumente usados com consoantes.

a) Diacríticos que indicam a localização horizontal:

- avançado: na transcrição estrita [k̟], o diacrítico [̟], um sinal de mais (+) subscrito, indica que o [k] não é produzido na região velar, mas sim um pouco mais para a frente, podendo ser, por exemplo, numa região intermediária entre a velar e a palatal.

- recuado (retraído): na transcrição estrita [d̟], o diacrítico [̟], um sinal de menos (-) subscrito, indica que a fase [d] da africada [dʒ] não é produzida na região dental ou alveolar, mas sim na pós-alveolar.
- b) Diacríticos que indicam a localização vertical:
- abaixado: na transcrição estrita [ɔ̞], o diacrítico [̞], indica que o [ɔ] é produzido com menor aproximação dos órgãos articuladores, o que equivale a dizer que se trata de uma aproximante, e não de uma fricativa.
 - alçado: na transcrição estrita [ɪ̟], o diacrítico [̟], indica que o [ɪ] é produzido com maior aproximação dos órgãos articuladores. Essa notação pode ser usada para representar o som grafado <ř> em tcheco, como no nome do compositor *Dvořák* ['dvɔ̟rɑ:k], que é uma fricativa.
- c) Diacríticos que indicam articulação secundária:
- labializado: na transcrição estrita [tʷ], o [ʷ] indica que o [t] é produzido com arredondamento dos lábios.
 - palatalizado: na transcrição estrita [pʲ], o [ʲ] indica que o [p] é produzido concomitantemente com a língua se aproximando do palato. Uma transcrição como [kʲi] pode ser considerada equivalente a [ɣi].
- d) Diacríticos usados só para sons coronais:
- dental: na transcrição estrita [t̪], o diacrítico [̪] indica que o som transcrito é dental, e não alveolar. Sem o diacrítico, a princípio poderia ser tanto um quanto o outro.
 - apical: na transcrição estrita [t̟], o diacrítico [̟] indica que o som transcrito é pronunciado com o ápice da língua, e não com a lâmina. Na prática, pode ser usado para representar um som alveolar.
 - laminar: na transcrição estrita [t̟̞] o diacrítico [̟̞] indica que o som transcrito é pronunciado com a lâmina da língua, e não com o ápice. Uma transcrição como [θ̟̞] pode representar uma fricativa dental surda em que a ponta da língua passa dos dentes, e a região da língua que se aproxima dos dentes superiores é a lâmina.
- e) Um diacrítico normalmente usado com oclusivas, mas que pode ocorrer com nasais:
- soltura não audível (ou som aplosivo): na transcrição estrita [t̚] o diacrítico [̚] indica que o som transcrito é pronunciado sem a fase de explosão audível. Na prática, a corrente de ar é interrompida antes de ser aberta a oclusão.

EXERCÍCIOS

- 1) Transcreva apenas as consoantes das palavras a seguir. Se puder, transcreva mais de uma pronúncia. A que você normalmente usa e outra(s), que pode(m) ser de outra região, mais formais, etc.

chuva	gorro	molho	vinho
próximo	táxi	exato	agitado
casa	soro	pasta	feliz
ativo	joelho	zona	aqui

- 2) Qual a diferença articulatória existente em cada par de consoantes?

[p b]	[n d]	[s ʃ]	[β b]	[l r]
[h h̥]	[c k]	[m ɲ]	[d d̥]	[k x]
[s θ]	[f φ]	[ɹ r]	[ʒ ʒ̃]	[r r̃]
[g k]	[z d]	[n m]	[b m]	[ŋ g]

- 3) Verifique quais dos pares a seguir têm consoantes com o mesmo ponto de articulação.

[ʎ ɲ]	[ŋ ɣ]	[v w]	[θ f]	[g x]
[z ʒ]	[r ʀ]	[l r]	[b v]	[h ʔ]
[l n]	[h h̥]	[ŋ n]	[j ɲ]	[j ʝ]
[t ʈ]	[φ p]	[r s]	[q k]	[d t]

- 4) Verifique quais dos pares a seguir têm consoantes com o mesmo modo de articulação.

[ʀ r]	[d s]	[x h]	[x r]	[z ʒ]
[l d]	[ʈ ʈ̥]	[ŋ n]	[x q]	[f p]
[ʒ ð]	[z ʒ̃]	[d̥ ʒ̃ p̃f]	[m w]	[h r]
[n k]	[q p]	[ʃ ɲ]	[ʎ ʎ̃]	[ð t]

- 5) Verifique quais dos pares a seguir têm consoantes com a mesma situação com relação ao vozeamento: presença ou ausência.

[g ŋ]	[ɣ l]	[ʈ v]	[z β]	[f φ]
-------	-------	-------	-------	-------

[ʁ r]	[z t]	[ʒ θ]	[t t]	[m Δ]
[t ɲ]	[b k]	[ʔ h]	[ɬ ʃ]	[x ç]

- 6) Qual é a consoante intrusa em cada grupo? Por quê? Em alguns casos há mais de uma possibilidade.

[b z ʃ]	[m l ɲ]	[r v s]	[f g x]	[ʌ β p]
---------	---------	---------	---------	---------

- 7) Quais consoantes na transcrição fonética seriam transcritas por símbolos diferentes da letra nas palavras abaixo?

aquela	nome	preço	fino	gelo
sol	rosa	gomo	faixa	doze
senha	foto	vaza	achava	beijo

- 8) No finlandês existe a chamada gradação consonantal, em que uma classe de consoantes tem sua pronúncia alterada na sílaba final da palavra, passando a ter uma pronúncia que pode ser considerada mais fraca. O nome desse processo é lenição. Neste exercício aparecem dados de gradação quantitativa e qualitativa. As denominações devem ser transparentes, mas na gradação qualitativa apenas se altera o tipo de consoantes, enquanto na quantitativa é a quantidade que muda: em vez de uma consoante dobrada ou geminada, grafada dupla, temos uma consoante simples. Qual a classe de consoantes que sofre lenição em finlandês? Quais consoantes sofrem gradação quantitativa e quais sofrem gradação qualitativa?

Nominativo	Genitivo	
tapa	tavan	‘hábito’
kukka	kukan	‘flor’
kauppa	kaupan	‘loja’
kohta	kohdan	‘lugar’
luku	luvun	‘número’
tumma	tumman	‘escuro’
äiti	äidin	‘mãe’
isä	isän	‘pai’
loma	loman	‘férias’
ovi	oven	‘porta’

linna	linnan	‘castelo’
pyörä	pyörän	‘bicicleta’
liha	lihan	‘carne’
pallo	pallon	‘bola’
hissi	hissin	‘elevador’
katto	katon	‘teto, telhado’
seinä	seinän	‘parede’
puoli	puolen	‘lado’

9) Nossa tradição ocidental tem a impressão de que tudo começou na Grécia. Mas em outras regiões há tradições riquíssimas mais antigas ou bem mais ricas, se comparadas ao que havia na Europa na mesma época. A escrita do sânscrito revela uma análise muito requintada das propriedades fônicas. Vejamos a seguir a sequência das consoantes na escrita devanágari do sânscrito, a qual é usada, por exemplo, para ordenar os termos num dicionário, como nossa ordem alfabética. Essa escrita é usada desde os primeiros séculos da nossa era, mas tem princípios já presentes em sistemas de escrita de épocas anteriores. As consoantes que aparecem com duas pronúncias, como [c/tʃ] representam à esquerda uma pronúncia que talvez tenha sido a original e à direita a pronúncia atual nas línguas indoarianas. Uma consoante aparece com três pronúncias (v/v/w) porque tanto no sânscrito quanto nas línguas indoarianas atuais ela parece alternar entre algumas dessas pronúncias. Obs: a rigor, cada letra dessas representa a consoante indicada aqui seguida de um [a] breve, ou *schwa* [ə], mas não incluí essa vogal na transcrição.

As 25 primeiras consoantes formam um grupo à parte, organizado em cinco colunas e cinco linhas. O que caracteriza cada coluna? O que caracteriza cada linha? Que princípio determina a ordem dessas cinco primeiras linhas?

A primeira linha do grupo de consoantes que vem após as primeiras 25 são as chamadas semivogais. A segunda linha são as sibilantes. O que determina a ordem em cada uma dessas linhas? É o mesmo princípio que determina a ordem das cinco primeiras linhas? Um detalhe é que a mesma raiz pode ter < ɾ > ou [ɽ] em certos contextos, mas apenas um retroflexo em outros.

क	ख	ग	घ	ङ
k	k ^h	g	g ^h	ŋ
च	छ	ज	झ	ञ

c/tʃ	c ^h /tʃ ^h	ʃ/dʒ	ʃ ^h /dʒ ^h	ɲ
ट	ठ	ड	ढ	ण
t	t ^h	ɖ	ɖ ^h	ɳ
त	थ	द	ध	न
t	t ^h	d	d ^h	n
प	फ	ब	भ	म
p	p ^h	b	b ^h	m

य	र	ल	व	
j	r	l	v/v/w	
श	ष	स		
ʃ	ʂ	s		
ह				
ɦ				

- 10) Na descrição das consoantes retroflexas, vimos que se pode falar de dois graus de retroflexão no português do Brasil. Se há esses pelo menos dois graus de retroflexão, embora a rigor se deva considerar que há um *continuum*, como seria mais adequado classificar a retroflexão? Como um ponto de articulação, que é a forma como a retroflexão é classificada na tabela das consoantes pulmonares do alfabeto fonético internacional?
- 11) Observe a formação do caso ilativo em finlandês. Esse caso indica movimento para dentro. O finlandês grafia as vogais longas como vogais dobradas, por exemplo, um [u:] é grafado <uu>. Duas vogais juntas na grafia normalmente indicam a existência de um ditongo. Há dois subconjuntos nos dados abaixo. Como é construída a forma do caso ilativo a partir do nominativo nesses dois subconjuntos? O que determina a inclusão em cada subgrupo? Em um dos subconjuntos é introduzido um segmento. Que função ele tem nesse processo? [PCS]

nominativo	ilativo	
talo	taloon	‘casa’
koulu	kouluun	‘escola’
maa	maahan	‘terra’

Helsinki	Helsinkiin	'Helsinque'
tee	teehen	'chá'
tuoli	tuoliin	'cadeira'
suu	suuhun	'boca'

REFERÊNCIAS E LEITURAS COMPLEMENTARES

- Laver, John (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Ladefoged e Maddieson. *The Sounds of the World's Languages*.
- Ladefoged e Maddieson. *Consonants and Vowels*.
- Olson, Kenneth S. (2001). *The Phonology and Morphology of Mono*. University of Chicago: Tese de Doutorado.
- Oushiro, Livia; Mendes, Ronald Beline (2011). A pronúncia de (-r) em coda silábica no português paulistano. *Revista do GEL*, 8(2), 66-95.

LINKS

O quê? An Introduction to Tlingit Sounds with Words:

<https://www.youtube.com/watch?v=gr-x6EL39PY&t=586s>

CAPÍTULO 3

VOGAIS

Neste capítulo vemos o que caracteriza as vogais como classe de sons, o que distingue cada vogal das demais, e também o que caracteriza as semivogais e os ditongos.

3.1 LOCALIZANDO-SE NO ESPAÇO VOCÁLICO

Como vimos no capítulo anterior, a característica que define uma consoante é a presença de um obstáculo significativo à passagem da corrente de ar pelo nosso trato vocal. Diferentemente das consoantes, as vogais são articuladas sem nenhum obstáculo significativo à passagem do ar. A cavidade oral pode estar mais aberta ou mais fechada, chegando ao grau de fechamento presente num [i], por exemplo, mas nunca fechada o suficiente para produzir ruído.

A descrição das consoantes se baseou justamente no local e no tipo de obstáculo à passagem do ar, além de indicar se as pregas vocais estão vibrando ou não em sua produção. Exatamente o fato de não haver um obstáculo, uma constrição acentuada, torna mais difícil descrever as vogais. Um /b/ pode facilmente ser identificado como um som produzido com fechamento total da passagem do ar na altura dos lábios, e com um exame da laringe, a verificação de que as pregas vocais vibram em sua produção. Mas como é produzido o som representado por <é> na palavra *até*, ou o representado por <ô> na palavra *avô*? Não é nada óbvio.

Para a representação gráfica do espaço acústico vocálico, se costuma empregar um trapézio. Os ângulos do trapézio utilizado podem variar. Utilizo aqui um trapézio escaleno (com os quatro ângulos diferentes). Os ângulos mais agudos na parte superior indicam que a língua consegue avançar e recuar mais na região superior da cavidade oral do que na região inferior. O formato do trapézio utilizado aqui não coincide exatamente com o do AFI, no qual a aresta da direita é vertical. A explicação da escolha de outro formato está no final deste capítulo, na seção 3.8.

Por convenção, esse trapézio representa uma pessoa de perfil olhando para o lado esquerdo da ilustração. O trapézio foi inserido na ilustração que mostra a pessoa de perfil, para facilitar o entendimento. O lado esquerdo da ilustração corresponde à região anterior da cavidade oral e seu lado direito à sua região posterior, como se representasse um falante de perfil com o rosto do lado esquerdo e a nuca no lado direito da página, a configuração padrão adotada em fonética, como dito no capítulo anterior. Compare as figuras 11 e 12. Se o símbolo de uma vogal está colocado na parte superior da ilustração, isso representa que, quando ela é pronunciada, a língua fica próxima do céu da boca. Se a vogal estiver na parte inferior do diagrama, isso indica que a língua fica bem afastada do céu da boca na produção da vogal em questão. Um outro parâmetro de distinção vocálica é o papel dos lábios. A principal distinção feita é a que classifica as vogais em dois grupos: as que são pronunciadas com arredondamento dos lábios, e as que não são.



Figura 11. Espaço vocálico no trato vocal.

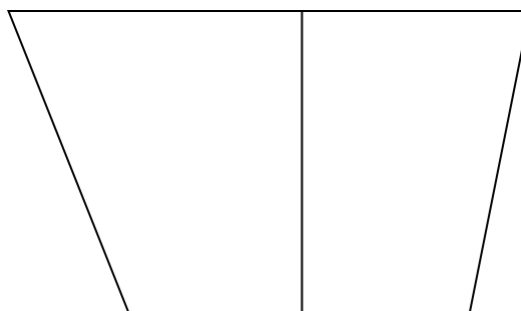


Figura 12. Espaço vocálico.

Voltando à questão de como descrever as vogais em termos articulatórios, um passo importante nesse sentido foi dado pelo foneticista inglês Daniel Jones. Ele propôs que a classificação das vogais partisse das vogais extremas, mais fáceis de descrever, e fosse situando as outras com relação a essas.

As duas primeiras vogais extremas que ele propôs que se tomasse como baliza para situar as outras foram a vogal mais fechada e anterior possível, e a vogal mais aberta e posterior possível.

Essas duas primeiras vogais são:

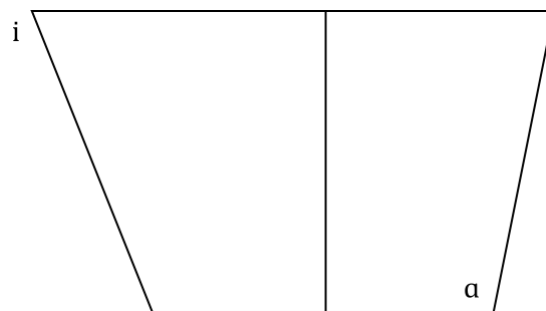


Figura 13. As duas primeiras vogais cardeais de Daniel Jones.

Como a representação no trapézio utilizada aqui é essencialmente a utilizada na tabela do AFI, algumas vogais ficam fora do trapézio. Isso será justificado no próximo parágrafo. Tendo partido dessas duas vogais, Jones toma outras duas vogais extremas como pontos de referência complementares.

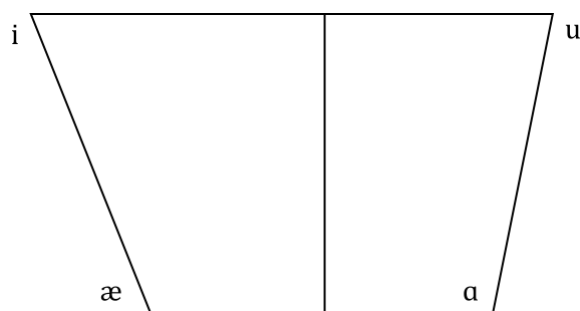


Figura 14. As quatro primeiras vogais cardeais de Daniel Jones (com uma retificação).

As duas vogais da parte superior desse diagrama correspondem essencialmente ao nosso [i] ou <i> em *ví*, e ao nosso [u] ou <u> em *tu*. Observe que das quatro vogais, a única que está à direita de uma das linhas não horizontais é a única pronunciada com arredondamento dos lábios. Esse é o motivo de ela estar à direita. As vogais que aparecem à esquerda dessas linhas não horizontais não envolvem arredondamento dos lábios. As que estão à direita são arredondadas.

As duas vogais da parte inferior do diagrama correspondem a vogais que não existem no português padrão, nem no do Brasil nem no de Portugal. A vogal [ɑ] é encontrada na pronúncia norteamericana mais difundida de *hot*, uma vogal bem aberta como o nosso [a] mas posterior como o nosso <ó>, e não central. A vogal [æ] é encontrada na pronúncia de *hat*, uma vogal bem aberta como o nosso [a] mas anterior como o nosso <é>, e não central. Neste ponto, preciso fazer uma observação importante. Na quase totalidade dos casos, sigo aqui as convenções de transcrição da Associação Fonética Internacional, mas com relação às vogais, discordo enfaticamente, assim como Barry e Trouvain (2008), da utilização do símbolo [a], oficialmente utilizado para transcrever uma vogal anterior aberta. Por isso, o

[æ] aparece na tabela deste livro ligeiramente abaixo de onde aparece na tabela oficial do AFI. Discuto isso mais extensamente na seção seguinte.

As quatro primeiras vogais ficam sendo então duas diagonais: [i ɑ] e [æ u]. É com relação a essas quatro vogais, pontos extremos no espaço vocálico, que vamos situar as demais. Segundo a proposta de Jones, em cada trajetória vertical, a mais anterior e a mais posterior, são acrescentados dois pontos aproximadamente equidistantes acusticamente, respectivamente, entre [i] e [æ], e entre [u] e [ɑ]. No primeiro caso, temos as vogais [e] e [ɛ], que ocorrem, respectivamente, na pronúncia de *dedo* e *perto*, no português, e no segundo caso, as vogais [o] e [ɔ], que ocorrem, respectivamente, na pronúncia de *hoje* e *fora*. Ficamos, então, com oito vogais.

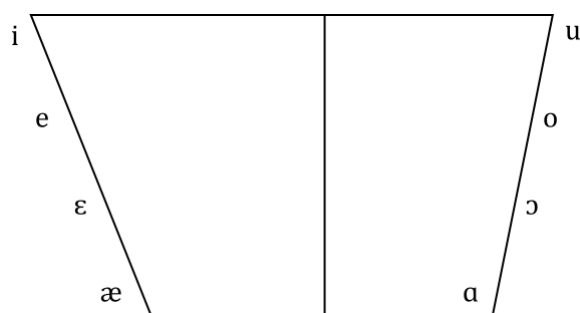


Figura 15. As oito vogais cardeais primárias de Daniel Jones (com uma retificação).

Resumindo a pronúncia de cada uma dessas oito vogais, temos o seguinte:

- a vogal [i] é encontrada no português *vi* ['vi].
- a vogal [e] é encontrada no português *vê* ['ve].
- a vogal [ɛ] é encontrada no português *pé* ['pe].
- a vogal [æ] tem o mesmo grau de abertura do nosso [a] central, mas a língua avança, sendo, portanto, uma vogal anterior. É encontrada no inglês em *hat* ['hæt] 'chapéu', *cat* ['kʰæt] 'gato'.
- a vogal [ɑ] tem o mesmo grau de abertura do nosso [a] central, mas a língua recua, sendo, portanto, uma vogal posterior. É uma pronúncia comum na América do Norte das palavras *not* ['nat] 'não', *hot* ['hat] 'quente', *shop* ['ʃap] 'loja'.
- a vogal [ɔ] é encontrada no português *só* ['sɔ].
- a vogal [o] é encontrada no português *avô* [a'vo].
- a vogal [u] é encontrada no português *tu* ['tu].

Essas são as oito vogais cardeais primárias. O termo *cardeais* remete aos pontos cardeais (leste, oeste, norte e sul), com base nos quais nos localizamos geograficamente. Essas vogais cardeais incluem várias das vogais mais difundidas nas línguas do mundo. Três vogais, na

verdade, ocorrem em cerca de 90% das línguas ou mais. Isso faz com que às vezes as vogais sejam representadas não num trapézio mas num triângulo, como o seguinte:

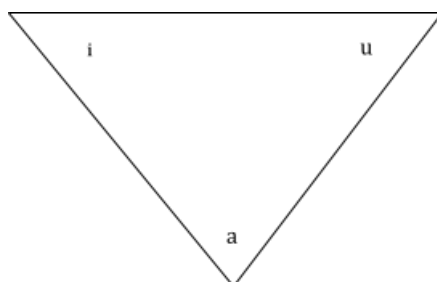


Figura 16. As três vogais extremas.

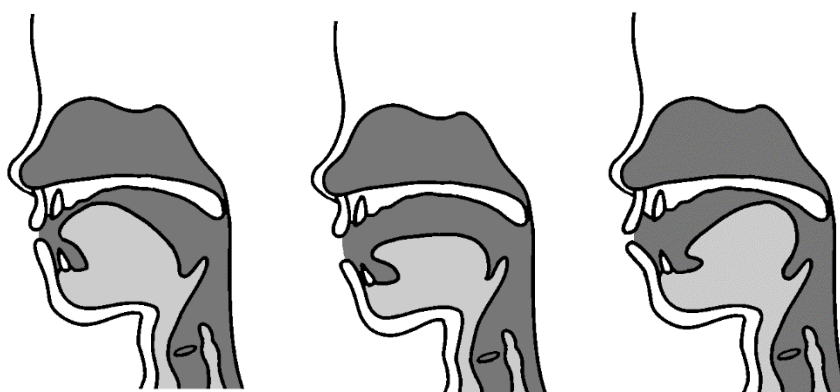


Figura 17. Articulação das três vogais extremas (na ordem, [i], [a] e [u]).

As vogais cardeais primárias [i e ε æ α ɔ o u] recebem inclusive uma numeração, sendo o [i] chamado de vogal cardinal primária 1; o [e] de vogal cardinal primária 2; e assim por diante, até o número 8, que é a vogal [u].

Detalhando parcialmente essas vogais, temos que [i] e [u] são vogais fechadas ou altas. São chamadas vogais fechadas se o critério adotado é o grau de abertura da cavidade oral. São chamadas de vogais altas se nos baseamos na posição da língua, que fica o mais alto possível para uma vogal na produção dessas duas vogais.

No outro extremo temos as vogais abertas ou baixas [æ] e [ɑ], além do [a], que não figura entre as vogais cardeais. As outras quatro vogais são denominadas médias, e se subdividem em médias fechadas ou médias altas, as vogais [e] e [o], e médias abertas ou médias baixas, as vogais [ɛ] e [ɔ].

O ideal da transcrição fonética é que cada símbolo corresponda essencialmente a um único fone e vice-versa, ou a uma região do espaço acústico com pequena variação. Há, no entanto, uma certa ambiguidade na transcrição das vogais médias que não aparece na notação. A rigor as vogais [e] e [o] são médias fechadas ou médias altas, como no português. Mas com frequência elas são usadas para transcrever vogais médias médias sem o uso de nenhum

diacrítico. Numa língua com sete vogais, como o português, [e] e [o] são de fato médias fechadas, e [ɛ] e [ɔ] médias abertas. Contudo, em línguas que não distinguem dois graus de abertura entre as vogais médias, como é o caso do espanhol, é comum que se tenha vogais intermediárias entre essas, o que poderíamos redundantemente ou por motivos de clareza denominar médias médias. A rigor, essas vogais deveriam ser transcritas como [ɐ̝] e [ɔ̝], usando o mesmo diacrítico que indica abaixamento visto no capítulo anterior, quando falei das aproximantes do espanhol, mas na prática em geral mesmo nessas línguas são usados esses dois símbolos vocálicos sem diacríticos. Flemming (2004) é um dos poucos trabalhos que distingue vogais médias altas e médias médias, mas, surpreendentemente, ele transcreve o que seriam as médias altas com o diacrítico *elevado* ou *alçado* [ɛ̥] e [ɔ̥] e deixa as médias médias sem diacrítico nenhum: [e] e [o].

Como já vimos, além do que poderíamos denominar posição horizontal (anteriorizada, neutra ou central, ou posteriorizada) e posição vertical da língua (baixa, média baixa, média alta, alta), há ainda o papel dos lábios na produção das vogais. Sua configuração mais característica é a que denominamos de arredondamento dos lábios.

É com base na presença ou ausência desse arredondamento que Jones propôs que, além dessas oito vogais cardeais primárias, se considerasse outras oito vogais cardeais, as secundárias, na mesma posição de cada uma das primárias, mas com o arredondamento dos lábios invertido. Se a vogal cardinal primária não tem arredondamento, a secundária tem, e vice-versa.

Assim se obtém as oito vogais acrescentadas abaixo.

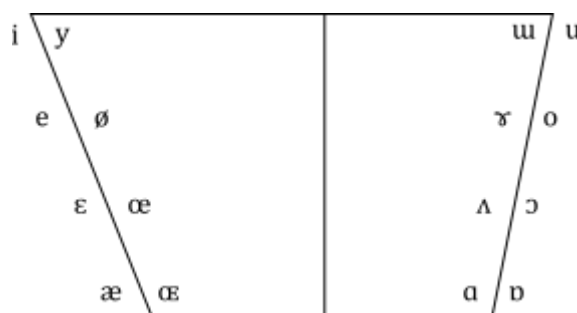


Figura 18. As 16 vogais cardeais de Daniel Jones (com uma retificação).

Detalhando a pronúncia de cada uma dessas oito vogais secundárias, temos o seguinte:

- a vogal [y] corresponde a um [i] pronunciado **com** arredondamento dos lábios. É encontrada no francês *tu* ['ty] 'tu, você', e no alemão *müde* ['my:də] 'cansado'.
- a vogal [ø] corresponde a um [e] pronunciado **com** arredondamento dos lábios. É encontrada no francês *feu* ['fø] 'fogo' e *Peugeot* [pø'ʒo] e no alemão *Flöte* ['flø:tə] 'flauta'.

- a vogal [œ] corresponde a um [ɛ] pronunciado **com** arredondamento dos lábios. É encontrada no francês *fleur* ['flœʁ] 'flor' e no alemão *Löffel* ['lœfl] '(a) colher'. Como o [œ] pode ser longo diante de [ʁ], às vezes ele aparece transcrito com [ɔ̃], que indica essa duração maior, ficando assim a transcrição: ['flœ:ʁ].
- a vogal [œ] corresponde a um [æ] pronunciado **com** arredondamento dos lábios. É extremamente rara, sendo encontrada, embora um pouco menos aberta no sueco *förr* 'antes', que, portanto, poderia ser transcrito como ['fœr:] ou ['fœr:].
- a vogal [u] corresponde a um [u] pronunciado **sem** arredondamento dos lábios. É encontrada no japonês *sushi*, cuja primeira sílaba é [su] e no vietnamita *Đức* [duk] 'Alemanha'. A rigor, a vogal do japonês tem labialização, embora não arredondamento, mas sim uma certa compressão dos lábios. Esse detalhe da pronúncia pode ser representado numa transcrição estrita por um [β] sobrescrito: [su^β].
- a vogal [ɤ] corresponde a um [o] pronunciado **sem** arredondamento dos lábios. É encontrada no vietnamita *ở* [ɤ:ɺ] '(estar) em' e no estoniano *sõna* ['sɤna] 'palavra'.
- a vogal [ʌ] corresponde a um [ɔ] pronunciado **sem** arredondamento dos lábios. É encontrada, por exemplo, no dinamarquês *komme* ['kʰʌmə] 'vir' e no coreano em [sʌl], a pronúncia do nome da capital Seul. A grafia *Seoul*, embora dê a impressão de ter um <e> seguido de um <ou>, tem um <eo>, que representa a vogal [ʌ], seguido de um <u>. A vogal do coreano pode ter um leve arredondamento, o que seria transcrito com o diacrítico [ɤ̠], ficando assim: [ʌ̠].
- a vogal [ɒ] corresponde a um [ɑ] pronunciado **com** arredondamento dos lábios. É encontrada na pronúncia britânica de *not* ['nɒt], 'não'.

Recapitulando, as vogais cardiais primárias são [i e ɛ æ ɑ ɔ o u] e as vogais cardiais secundárias são [y ø œ ɐ ʊ ɤ ʌ ɯ]. Há outras listas de vogais cardiais, que acrescentam vogais centrais, além dessas 16, chegando a um total de 18 ou 22.

Encerro esta seção mencionando um termo emprestado da música que é empregado para se referir à qualidade vocálica: timbre. Roederer (1998: 21) define assim o termo: "A qualidade, ou timbre, é o que nos permite distinguir os sons de diferentes instrumentos, mesmo que eles tenham a mesma altura e a mesma intensidade." Na música, podemos ouvir a mesma nota tocada num piano, num cravo, numa guitarra, num violino, e perceber que foram produzidos por um desses instrumentos, e não por outro. O timbre vocálico é a qualidade sonora que nos permite distinguir as vogais umas das outras.

3.2 A VOGAL CENTRAL ABERTA

Como disse na seção anterior, um ponto bastante incômodo, para mim, mas também para diversos outros pesquisadores, é a vogal transcrita como [a]. Na transcrição oficial do AFI, a vogal [a] é usada para representar uma vogal anterior aberta. Estranhamente, não há um símbolo oficial para representar sem o uso de diacríticos a vogal central aberta, o [a] (!), presente no português, uma das vogais mais comuns entre as línguas do mundo. A meu ver, considerando as vogais das línguas em geral, transcrever a vogal aberta anterior como [a] foi uma escolha ruim. Neste livro, como muitos outros fonólogos, utilizo o [a] para transcrever a vogal aberta central.

Para justificar essa escolha, preciso deixar claro que a transcrição de uma vogal se relaciona com todo o esquema de vogais cardeais. Ao que tudo indica, na determinação das vogais cardeais não foram levadas em consideração apenas fatores como localização em pontos extremos do espaço acústico ou da cavidade oral. É evidente que as vogais cardeais primárias são mais comuns que as secundárias, e justamente por isso são primárias: por exemplo, [i e ε] são muito mais comuns que [y ø œ].

É por esse motivo que uma das duas primeiras vogais cardeais é o [a], vogal baixa não arredondada, e não a correspondente arredondada [ɒ], já que esta é menos comum. Repare também que ela é a única vogal cardeal primária posterior não arredondada. Isso pode ser justificado por dois motivos. Por um lado, dos quatro graus de abertura das vogais posteriores, o único que ocorre com mais frequência sem arredondamento é o das vogais abertas, e o fato de ser mais comum nas línguas em geral é um fator na determinação de quais vogais são cardeais primárias ou secundárias. Por outro lado, o grau de arredondamento é gradativamente menor à medida que aumenta o grau de abertura da cavidade oral, por simples dificuldade articulatória.

É bem provável que a dificuldade do arredondamento de vogais abertas ou o fato de elas serem menos comuns tenha feito Daniel Jones considerar que a número 5 (cardeal primária) seria a vogal não arredondada enquanto que a vogal 13 (cardeal secundária) seria arredondada. Os dois motivos, na realidade, estão correlacionados.

Outra questão problemática é que, como já comentei, da forma como o alfabeto fonético internacional foi elaborado, a tendência é que os sons mais comuns sejam representados por símbolos sem diacríticos. Nenhuma vogal cardeal é representada com diacríticos. Ora, a vogal aberta central é uma das mais comuns. Se é assim, por que ela haveria de ser representada com diacríticos? Isso faz com que vogais pouco comuns, como a vogal alta anterior arredondada [y] ou mesmo a vogal anterior aberta arredondada [œ], quase inexistente nas línguas em geral, sejam representadas com símbolos sem diacríticos, e uma vogal como a vogal aberta central da palavra *já* seja representada como [ä] ou [a], o que não faz sentido.

Se mantivermos o símbolo [a] para representar a vogal anterior aberta, essas duas possibilidades de transcrição para a vogal central aberta, [ä] e [a], têm diacríticos que representam, respectivamente, centralização e posteriorização. Além do inconveniente de representar uma das vogais mais comuns com diacrítico, esses diacríticos podem indicar um certo grau de centralização ou posteriorização, e não necessariamente que ela chega a ser central. A vogal representada poderia ser algo entre anterior e central com esses diacríticos.

Um possível fator que determinou a escolha de [a] para representar uma vogal anterior no sistema de Daniel Jones é que as vogais do inglês, pelo menos em sua versão denominada *Received Pronunciation* não incluem a vogal aberta central como monotongo, mas só fazendo parte de ditongos como [aj] e [aw], encontrados em palavras como *like* e *house*.

Um texto detalhado sobre o tema é Barry e Trouvain (2008), que retoma a discussão sobre a ausência de um símbolo que represente a vogal aberta central. Eles afirmam que o tema já tinha sido “oficialmente” discutido no encontro da Associação Fonética Internacional em Kiel em 1989 mas não apresentado para discussão pela comunidade fonética de forma mais ampla. Ao final do artigo eles listam três possíveis soluções, sem optar explicitamente por uma. Refiro os interessados ao artigo e adoto a segunda solução proposta (p. 354): redefinir a vogal [æ] como vogal anterior aberta, em vez de vogal anterior quase aberta, passando a ser a vogal cardeal número 4. Eles afirmam que essa opção tem mais de um argumento em seu favor. Em primeiro lugar, embora não oficial, [a] é o símbolo de fato usado em sistemas vocálicos triangulares, os que têm apenas uma vogal aberta. Em segundo lugar, no inglês britânico padrão, a vogal de *cat*, a provável justificativa anglocêntrica para a inclusão do símbolo [æ] para uma vogal quase aberta, se tornou aberta. O próprio *Handbook of the IPA* faz essa afirmação com relação ao inglês britânico meridional. Como benefício fortuito, eles aludem que o símbolo [æ] combinaria perfeitamente com o símbolo [œ] da vogal anterior aberta arredondada.

Mais cautelosamente, Roca e Johnson (1999: 128) lamentam o fato de o símbolo [a] ser encontrado com dois valores distintos nos textos, representando uma vogal anterior ou uma vogal central, e dizem que vão dar o passo “ousado” de usar a notação que também adoto aqui: usam [æ] para representar a vogal aberta anterior e [a] para representar a vogal aberta central. Para concluir, na prática, um grande número de autores usa de fato [a] para representar a vogal aberta central, que é muito mais comum que a vogal aberta anterior.

Por fim, falando como professor, vejo que para estudantes falantes do português, a classificação e a tabela têm o enorme inconveniente de sugerir que a vogal [a] do português é anterior, por mais que já tenha sido dito repetidas vezes em aula que nosso [a] é uma vogal central.

Foi uma argumentação um tanto prolongada, mas acredito ter assim justificado o fato de neste livro usar o símbolo [a] para representar a vogal aberta central, e deslocar o símbolo [æ] um pouco para baixo no espaço acústico, usando-o para representar a vogal cardeal número 4.

3.3 AS VOGAIS ORAIS DO PORTUGUÊS

Tendo visto as vogais cardeais primárias e secundárias, e examinado a questão da vogal central aberta, veremos nesta seção as vogais do português. As vogais orais tônicas do português são as seguintes: [i e ε a ɔ o u]. As vogais orais átonas na última sílaba no português do Brasil em geral não são exatamente iguais a nenhuma dessas.

Por isso, embora tenhamos aqui já um número grande de símbolos vocálicos, eles não são suficientes para transcrever com detalhe suficiente a pronúncia de todas as vogais do português. Examinemos algumas palavras que aparentemente têm mais de uma vez a mesma vogal. Se compararmos a palavra *vatapá* e a palavra *mapa*, é possível perceber que em *vatapá* temos três vezes a vogal [a] pronunciada sem grande diferença de uma sílaba para a outra. Mas na palavra *mapa* a vogal da segunda sílaba não é exatamente igual. Podemos até pronunciar a palavra *mapa* diante do espelho. Talvez isso nos ajude a perceber que o segundo <a> não representa exatamente a mesma vogal que o primeiro. Enquanto o primeiro <a> representa uma vogal bem aberta, o segundo <a> não é tão aberto. Como a transcrição fonética pretende representar as diferenças de pronúncia, existem símbolos diferentes para representar essas duas vogais. A vogal central com grau máximo de abertura é representada como [a]. Já a vogal um pouco menos aberta é representada pelo símbolo [ɐ]. Essa vogal menos aberta é chamada de *reduzida*.

Tomando um par de exemplos mais semelhantes entre si, se compararmos a vogal <a> da sílaba final de *fala* com a de *falar*, veremos que elas não são iguais. A vogal da sílaba final de *falar* é tônica, podendo até ser o último som da palavra, caso o som grafado <r> não seja pronunciado. É essa a pronúncia, sem o <r>, que será transcrita aqui. Além da diferença da tonicidade, isto é, o fato de que a vogal final de *fala* é átona e a de *falar* é tônica, podemos identificar a diferença já mencionada do grau de abertura. A vogal central da sílaba final de *falar* é pronunciada bem aberta, ao passo que a vogal central de *fala* não é tão aberta. A transcrição dessas duas formas ficaria assim:

(1) falar: [fa'la] fala: ['fale]

Repare que o acento é indicado pelo sinal ['] antes da sílaba tônica. Alguns dicionários fazem o contrário, colocando o ['] depois da sílaba tônica, enquanto outros dicionários colocam o

['] antes da vogal da sílaba tônica, transcrevendo [fale], por exemplo, embora nenhum dos dois seja o padrão da Associação Fonética Internacional. É importante verificar qual convenção um dicionário adota.

Da mesma forma que com a vogal átona final grafada <a>, as vogais grafadas como <o> e <e> em sílaba átona final têm, na maior parte das regiões em que se fala português, uma pronúncia que não corresponde à de nenhuma vogal tônica. Esse <o> não representa nem um [o] nem um [u], exatamente. O <e> também não corresponde exatamente a um [e] nem a um [i].

Se compararmos as palavras *urubu* e *suco*, ou as palavras *vovô* e *soco*, é possível que consigamos perceber que em *urubu* temos três vezes a vogal [u] pronunciada de forma equivalente de uma sílaba para a outra, mas em *suco*, a vogal grafada <o>, embora pareça um [u], é um pouco diferente. Comparando as duas vogais de *vovô* e as duas de *soco*, podemos perceber algo análogo. As duas vogais de *vovô* são idênticas mas as duas de *soco* não são.

Da mesma forma, comparando novamente formas do verbo *falar*, as formas *falo* e *falou* têm diferenças parecidas com as que vimos com relação à vogal central. A pronúncia mais comum no Brasil da forma *falou* é sem ditongo na sílaba final. Comparando a pronúncia das formas *falo* e *falou*, vemos que, além do fato de a sílaba final de *falo* ser átona e a sílaba final de *falou* ser tônica, a qualidade do som das vogais é diferente.

Então, a vogal final de *suco*, *soco* e *falo*, para a maioria dos falantes do português não é nem [u] nem [o]. É uma vogal intermediária, chamada de *u reduzido*. Ela é bem semelhante à vogal encontrada em palavras inglesas como *look* ou *book*. O desenho do símbolo fonético é um ômega maiúsculo pequeno de cabeça para baixo: [ʊ]. As formas *falou* e *falo* são transcritas assim:

(2) *falou*: [fa'lo] *falo*: ['falʊ]

Embora ainda não tenhamos visto os ditongos, repare que, na transcrição de *falou*, a segunda sílaba tem apenas a vogal [o], e não um ditongo. Faço isso para representar a pronúncia mais difundida no Brasil e em Portugal, em que não há mais ditongo, mas sim uma vogal simples, ou monotongo.

Encontramos o mesmo tipo de diferença quando comparamos as vogais das palavras *tititi* e *tive*, ou as das palavras *bebê* e *teve*. É possível que consigamos perceber que em *tititi* temos três vezes a vogal [i] pronunciada de forma equivalente de uma sílaba para a outra, mas em *tive*, a vogal grafada <e>, embora pareça um [i], é um pouco diferente. Comparando as duas vogais de *bebê* e as duas de *teve*, podemos perceber algo análogo. As duas vogais de *bebê* são idênticas, mas as duas de *teve* não são. Então, a vogal final de *tive* e de *teve* não é nem [i] nem

[e]. É uma vogal intermediária, transcrita com o símbolo [ɪ], que é o <I> em versalete, e chamada de *i reduzido*. Ela é bem semelhante à vogal encontrada em palavras inglesas como *sit* ou *miss*.

As formas *tevé* e *teve* são transcritas assim:

- (3) tevé: [te've] teve ['tevi]

Em partes do Sul do Brasil, não ocorrem nem o [ʊ] nem o [ɪ] em posição final, mas sim [o] e [e]. É possível que o contato com o espanhol seja um fator nessa diferença. É possível ainda que as duas vogais de *vovô* e *bebê* não sejam iguais para alguns falantes do português do Brasil.

As três vogais [ɐ ʊ ɪ] são chamadas de vogais reduzidas no português. O termo é adequado em dois sentidos. Primeiro, em termos de duração, essas são vogais muito breves, as mais breves do português, juntamente com algumas outras ocorrências de vogais átonas em meio de palavra. Além de temporalmente terem uma duração reduzida, os movimentos feitos para produzi-las são menos extremos comparados aos das vogais tônicas correspondentes [a u i]. Isso pode ser verificado também no diagrama das vogais do Alfabeto Fonético Internacional. As três vogais tônicas [a u i] ocupam posições bem periféricas, estando nos extremos da localização possível de vogais no espaço acústico-articulatório. As três vogais reduzidas estão ligeiramente mais para dentro nesse espaço acústico-articulatório.

Embora só ocorram em sílabas átonas no português, essas vogais podem ocorrer em sílaba tônica em outras línguas, como já mencionado, não sendo, então, reduzidas nessas línguas. Duas delas ocorrem em inglês, as vogais [ʊ] e [ɪ], como nos exemplos a seguir:

- (4) lick ['lɪk], licking ['lɪkɪŋ] look ['lʊk], looking ['lʊkɪŋ]

Uma outra vogal que é importante mencionar ao falarmos do português, embora ela seja inexistente no português do Brasil, é uma vogal muito frequente no português europeu. Ela é central, assim como o [a] e o [ɐ], mas não é aberta como a primeira, nem quase aberta como a segunda. Trata-se da vogal central fechada, transcrita com um <i> cortado: [ɨ]. Em sílaba átona, normalmente essa é a pronúncia de uma vogal grafada como <e>, como nas palavras *perigo* e *pessoa*, que seriam transcritas [pɨ'riɡu] e [pɨ'soɐ]. Ocorre que, inclusive por só ocorrer em contexto átono, essa vogal frequentemente acaba não sendo pronunciada, dando origem a pronúncias como: ['prigu] e ['psoɐ]. Se se quiser representar numa transcrição só as duas pronúncias, pode-se usar parênteses para indicar o som que é opcional na pronúncia: [p(ɨ)'riɡu] e [p(ɨ)'soɐ]. Cruz-Ferreira (1995), no entanto, descreve essa vogal como quase posterior quase fechada não arredondada (*near-close near-back unrounded vowel*), o que poderia ser transcrito com símbolos *ad hoc* tais como [ũ]

(medializada e centralizada, *mid-centralized*), [ɥ] (anteriorizada) and [ʊ] (menos arredondada, ou não arredondada). Mas no espaço acústico-articulatório a diferença entre essas duas vogais é mínima, o que justifica transcrever a vogal do português europeu como [i].

Com isso já podemos transcrever quase todas as vogais do português com algum detalhe fonético, sem considerar se essa diferença afeta o significado ou não. Faltam as vogais nasais, que veremos na próxima seção. Para resumir o que já foi visto, ou seja, as vogais cardiais primárias e secundárias, as três vogais reduzidas do português e a vogal alta central do português europeu, apresento aqui o trapézio contendo todas essas vogais.

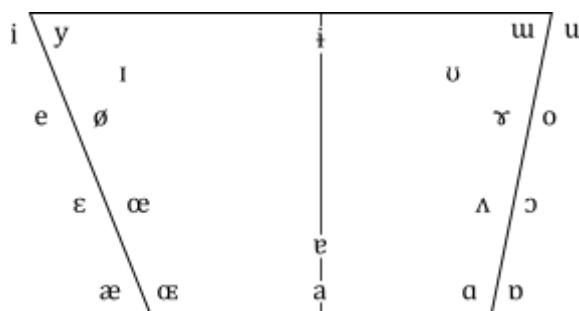


Figura 19. As 16 vogais cardiais e outras do português.

3.4 AS VOGAIS NASAIS DO PORTUGUÊS

Uma possibilidade razoavelmente comum, ainda não vista neste capítulo, é a existência de vogais nasais ou nasalizadas. Não farei aqui distinção entre esses dois termos, e usarei apenas o termo *vogais nasais*. As vogais orais são pronunciadas com a passagem entre a faringe e a cavidade nasal, a abertura nasofaríngea, fechada, de modo que a corrente de ar não penetra na cavidade nasal, passando da faringe apenas para a cavidade oral. Porém, como ocorre no português, existe a possibilidade de a abertura nasofaríngea estar aberta quando se pronuncia uma vogal, resultando na produção de uma vogal nasal. Todas as línguas possuem vogais pronunciadas com a corrente de ar passando apenas pela cavidade orofaríngea, ou seja, pela faringe e pela boca. Parte das línguas tem também vogais em que a corrente de ar passa ao mesmo tempo pelas duas cavidades, a oral e a nasal, as vogais nasais.

As vogais nasais são transcritas colocando o *til* acima delas, como fazemos na ortografia do português, em palavras como *põe* ou *irmã*. Em princípio, qualquer vogal oral pode ter uma correspondente nasal, com o mesmo timbre, isto é, com as mesmas características de grau de abertura, posição da língua etc. É importante mencionar um detalhe, no entanto. De um modo geral na pronúncia do português, em vez das sete vogais orais, temos apenas cinco

vogais nasais. Não há [ɛ] nem [ɔ] nasais. As pronúncias mais comuns das palavras *tempo*, *vim*, *som* e *atum* seriam transcritas como se segue: ['tẽpu], ['vĩ], ['sõ] e [a'tũ]. Com essa transcrição queremos dizer que essas vogais nasais são produzidas com uma configuração da cavidade oral equivalente à das vogais orais correspondentes. Um [ĩ] é basicamente um [i] produzido com passagem do ar não só pela cavidade oral, mas também pela cavidade nasal. Em suma, em português, temos menos vogais nasais do que orais. As orais tônicas são em número de sete: [i e a ɔ o u]. As nasais apenas cinco: [ĩ ã õ ã ã]. É muito comum o inventário de vogais nasais ser menor do que o de vogais orais, como vemos nos exemplos do kaingang e do dâw na seção **Erro! Fonte de referência não encontrada..** Comento a vogal central alguns parágrafos mais abaixo.

A nasalidade vocálica é uma área bastante complexa no português por vários motivos. Em primeiro lugar porque as análises diferem. Além disso, existe variação na pronúncia. Existem pronúncias que mantêm a nasalidade de todas as vogais, mas existem também aquelas em que uma ou todas as vogais são parcial ou totalmente desnasalizadas. Neste capítulo tratarei inicialmente da realização tida como mais prototípica, ou pelo menos a realização mais tradicional da nasalidade vocálica no PB.

Um outro ponto importante de assinalar é que uma análise acústica nos revela que as vogais nasais não têm uma pronúncia uniforme em toda a sua duração, permitindo, por exemplo, que vejamos que a pronúncia da primeira e da segunda porções da vogal não são exatamente iguais, apresentando diferenças sutis, como será visto na seção **Erro! Fonte de referência não encontrada..** Abstraio essas diferentes fases das vogais nasais, tratando delas como se fossem uniformes em toda a sua duração.

Em princípio encontramos cinco vogais nasais no português. Podemos relacioná-las com as vogais orais /i e a o u/. Nos pares a seguir existem pronúncias no Brasil em que basicamente a única diferença é a nasalidade ou não da vogal tônica.

(5)	vi ['vi]	vim ['vĩ]
	cedo ['sedu]	sendo ['sẽdu]
	lobo ['lobu]	lombo ['lõbu]
	mudo ['mudu]	mundo ['mũdu]

O caso da vogal central nasal no português, no entanto, é diferente. Se pronunciamos as palavras *lá* e *lã*, provavelmente vamos perceber que o grau de abertura na cavidade oral é muito diferente. A vogal da primeira palavra é bem aberta, ao passo que a vogal da segunda é bem menos aberta. Quando se nasaliza, a vogal central fica menos aberta de uma forma muito nítida. Assim, o timbre da vogal nasal é muito diferente do da vogal oral. Varia bastante a transcrição dessa vogal nasal. A mais simples seria utilizar o [ɐ] já visto com um

til, ou seja, [ẽ], já que ele representa uma vogal menos aberta. Aparentemente, no entanto, essa vogal central nasal é mais fechada ainda. Duas alternativas para representar esse maior fechamento da cavidade oral são os símbolos [ɜ] e o símbolo [ə]. O primeiro é um pouco mais fechado que o [e], e representa o som encontrado em palavras inglesas como *bird*. O segundo é um pouco mais fechado ainda, e se encontra por exemplo na primeira sílaba da palavra inglesa *about*.

A palavra *fama* é transcrita, então, de pelo menos três formas diferentes: ['fẽme], ['f̃me] ou ['f̃me]. É possível que isso reflita uma certa variabilidade no grau de fechamento ou alçamento que os falantes utilizam ao pronunciar essa vogal. Além dessas três transcrições, embora ocorra em textos da área uma transcrição como ['fãme], ela é a menos adequada de todas, porque pode dar a entender que a vogal central nasal é tão aberta quanto a oral correspondente. Claro que num texto que só fale do português pode-se até deixar isso subentendido, mas como este livro trata da fonologia das línguas em geral, é mais correto evitar esse tipo de transcrição. Além do mais, se estamos fazendo uma transcrição *fonética*, devemos tentar representar diferenças de pronúncia. E é fácil um falante do português perceber que a vogal central nasal não é aberta como a vogal central oral.

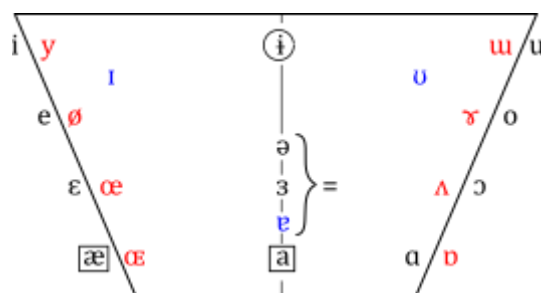
O que há em comum entre as três transcrições apresentadas acima da palavra *fama*: ['fẽme], ['f̃me] ou ['f̃me]? São todas vogais centrais, assim como o [a], mas com grau de abertura decrescente, indo de uma aberta, mas não no grau máximo, o [e], passando por uma vogal média aberta, o [ɜ], e podendo chegar a uma média propriamente dita, o [ə], chamado *schwa* pronunciado *xuá*, como aliás escreveu Mattoso Câmara Jr. (1981: 246). O [ə] é uma vogal central e média, até por isso também denominada *vogal neutra*.

Pelo menos em São Paulo e no Rio Grande do Sul existem trabalhos que apontam uma tendência de desnasalização da vogal central nasal. Talvez isso seja justamente consequência da mudança grande no timbre da vogal. Se a vogal oral e a nasal correspondente já têm grau de abertura tão diferente, essa simples diferença no grau de abertura já é suficiente para distingui-las, fazendo com que, mesmo que a vogal média não seja quase nasalizada ou se desnasalize totalmente, ela seja percebida como se fosse nasal. Nesse caso podemos ter pronúncias como ['f̃me].

Um incidente que me chamou atenção por volta do ano 2000 e deixou claro como já era presente em São Paulo essa pronúncia desnasalizada aconteceu quando desapareceu o cachorro de uma família que morava perto de mim. Para tentar localizar o cachorro, a família colocou avisos nos postes com o nome do cachorro, que era *Lucky*. Para representar a pronúncia dessa palavra, que é inglesa, eles a escreveram numa grafia aportuguesada, dessa forma: *Lãqui*. Foi o til que chamou a minha atenção. Ele não era uma tentativa de representar uma vogal nasal, algo como *Lanque* ['lõki], que em inglês seria *lunky*, uma

Outro *causo* que mostra o quanto nossa vogal central nasal tem um timbre muito diferente da vogal oral correspondente me foi contado por um aluno argentino. Ele e a namorada brasileira estavam no Paraguai e iam comprar uma pulseira, se não me engano, e queriam gravar o nome dela, Ana, na pulseira. A pessoa que ia gravar o nome perguntou o nome da moça e ela disse: [ˈɛnɐ]. O falante de espanhol respondeu: Hã??? A moça repetiu o nome algumas vezes. Só depois de várias tentativas, sabe-se lá como, caiu a ficha e o outro disse: Ah, [ˈana]?! O que dificultou a comunicação aí foi o timbre muito diferente da vogal.

Tendo visto as vogais cardeais primárias e secundárias e as vogais do português, vejamos agora um diagrama que contém todas as vogais cardeais, bem como todas as vogais do português que discutimos.



Nesse diagrama, as duas séries de vogais próximas às linhas diagonais são as vogais cardeais. Dessas, as que estão em preto são as cardeais primárias e as que estão em vermelho são as cardeais secundárias. As que estão em azul são as vogais orais reduzidas do PB. A chave, }, com o sinal de igual à direita de três vogais no centro do diagrama indica que qualquer um desses símbolos pode ser usado com o til para representar a vogal central nasalizada do português. As duas vogais que aparecem enquadradas estão com a posição alterada com relação ao diagrama oficial da Associação Fonética Internacional. O [a] coloco no centro, e o [æ] na mesma linha horizontal do [a], ocupando a posição que o [a] ocupa no diagrama oficial. A vogal dentro do círculo no meio da parte superior não ocorre no português do Brasil, mas é muito frequente no de Portugal.

55

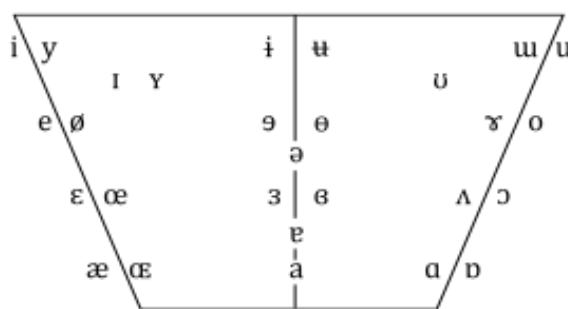


Figura 21. As vogais do alfabeto fonético internacional.

3.5 DITONGOS E SEMIVOGAIS

Para nos referirmos aos sons produzidos sem obstáculo à passagem do ar, em vez de usar o termo vogais, podemos usar o termo vocóides, mais abrangente, que desconsidera a posição desses fones na sílaba. Tanto vogais quanto semivogais podem ser incluídas na classes dos monotongos, embora o uso do termo não seja muito comum. Os chamados encontros vocálicos se dividem em dois tipos: aqueles em que cada vocóide está numa sílaba separada, e aqueles em que todos os vocóides estão na mesma sílaba.

Restringindo a discussão à situação em que são dois os vocóides que estão em sequência numa palavra, temos os hiatos e os ditongos. São exemplos dos dois tipos de encontros vocálicos as formas *saí* e *sai*. Em *saí* temos duas sílabas e em *sai* apenas uma. Sem entrarmos ainda na questão da transcrição, podemos marcar essa diferença colocando um ponto para separar as sílabas: *sa.í* e *sai*. Vemos que em *saí* cada vocóide está numa sílaba separada. Isso é o que define um hiato. O substantivo *hiato* provém do particípio do verbo latino *hiō*, *hiāre*, que significava ‘estar de boca aberta; estar boquiaberto; bocejar’. O hiato contém então uma separação, um “intervalo”, embora concretamente, mesmo num hiato, não haja pausa nenhuma. Já na forma *sai*, como temos apenas uma sílaba, os dois vocóides se encontram na mesma sílaba. Logo, não há hiato. Há ditongo, termo de origem grega que significa literalmente ‘(com) dois sons’, mas que adquiriu o significado de dois sons na mesma sílaba. Uma definição comum diz que um ditongo tem duas vogais na mesma sílaba. Essa definição tem o inconveniente, para a fonética e a fonologia, de estar mais voltada para a escrita. No português, independente da posição do vocóide anterior fechado na sílaba, ele é escrito predominantemente como <i>, e o vocóide posterior fechado como <u>.

Para entendermos melhor a estrutura de um ditongo, examinemos dois tipos de exemplos do português do Brasil. Existem várias palavras ou formas de palavras que apresentam variação entre uma pronúncia com ditongo e outra sem ditongo. São formas como *doutor*,

outro, peixe e queijo. A sílaba inicial de todas essas palavras pode ser pronunciada com ditongo ou apenas com monotongo (uma só vogal). Isso faz com que Bisol (1994) os analise como ditongos falsos. Ainda sem fazermos a transcrição fonética, se formos representar essa variação com base na ortografia, poderíamos escrever essas palavras assim: *do(u)tor, o(u)tro, pe(i)xe e que(i)jo*. Os sons que estão entre parêntesis podem ser pronunciados ou não.

Só veremos a sílaba e sua estrutura no capítulo 9, mas podemos considerar que em todos esses exemplos, o vocoide principal, o mais proeminente, é sempre pronunciado, enquanto que o vocoide secundário em termos de estrutura silábica pode ser elidido. Essa importância maior do vocoide mantido na sílaba está relacionada ao fato de ele receber a denominação de *núcleo*. Assim como na sintaxe, se temos em *casas novas* um sintagma nominal, pois seu núcleo é *casas*, se não houver a palavra *casas*, mas apenas *novas*, não temos mais um sintagma nominal. De forma semelhante, o que caracteriza uma sílaba como sílaba é ela ter um núcleo. O núcleo, principalmente de uma sílaba tônica, normalmente não é elidido, mas um segmento que está numa margem da sílaba, pode ser. Caracterizaremos, então, assim o contraste entre vogais e semivogais: as vogais ocorrem no núcleo (ou no núcleo do núcleo) e as semivogais numa margem, antes ou depois do núcleo (ou numa margem do núcleo).

As semivogais são também denominadas semiconsoantes por alguns autores. Como as duas denominações devem sugerir, elas têm características intermediárias entre as das vogais e as das consoantes. Se por um lado, não há obstáculo acentuado à passagem do ar na sua produção, elas costumam ocorrer vizinhas a vogais, que ocupam o centro, ou núcleo ou pico da sílaba. Sua análise na fonologia é controversa, mas uma análise comum distingue esses dois tipos de vocoides pelo fato de as semivogais estarem nas margens da sílaba, e as vogais no núcleo. Uma forma simples de formular isso seria dizer que uma semivogal fica vizinha a uma vogal (antes ou depois), ao passo que uma vogal não precisa estar vizinha a uma semivogal.

No português temos duas semivogais. Uma delas aparece na linha de aproximantes da tabela de consoantes pulmonares do alfabeto fonético internacional: o [j]. Essa é a semivogal que em português é representada na escrita na maior parte das vezes como <i>, embora às vezes seja grafada como <e>. As duas grafias alternam em nomes próprios como *Morais* ou *Moraes*. Temos ainda outra semivogal, o [w], que representa a semivogal mais frequentemente grafada com <u> no português, embora às vezes ela ocorra grafada com <o>. As duas grafias aparecem em termos homófonos como *Mao (Tsé-Tung)* e *mau*. O <o> é também a forma de grafar essa semivogal no ditongo <ão>, como em *chão*.

Como vimos na seção 2.7, surpreendentemente, o [w] não aparece na mesma linha do [j] na tabela do alfabeto fonético internacional. Ele está fora dessa tabela, no meio da página à esquerda, no conjunto denominado *outros símbolos*. A Figura 21 contém apenas a coluna da esquerda desses outros símbolos. O [w] é o segundo. Sua descrição contém a explicação de por que ele não está na tabela de consoantes pulmonares. Ele é descrito como aproximante labiovelar vozeada. Se ele é labiovelar, tem dois pontos de articulação, sendo ao mesmo tempo labial e velar. Sendo assim, ele não teria como figurar na tabela das consoantes, pois elas apresentam as consoantes como classificadas pelo seu *único* ponto de articulação.

OUTROS SÍMBOLOS

ɱ	Fricativa labiovelar não vozeada
ʋ	Aproximante labiovelar vozeada
ɥ	Aproximante labiopalatal vozeada
ħ	Fricativa epiglotal não vozeada
ʕ	Fricativa epiglotal vozeada
ʡ	Plosiva epiglotal

Figura 21. Quadro intitulado *Outros símbolos* na tabela do AFI.

Nessa lista aparece também a semivogal equivalente na pronúncia ao [y], representada pelo símbolo [ɥ], um [h] de cabeça para baixo, mas que apresenta semelhança gráfica com o [y], por ter a haste direita comprida, descendo abaixo da linha. Ele ocorre no francês *suís* ['sɥi] 'sou; sigo' e *muet* ['mɥe] 'mudo', isto é, antes de vogais anteriores.

Há outro tipo comum de transcrição dos ditongos, que discutirei na próxima seção.

Mas já cabe nessa seção a menção a um equívoco comum na transcrição de ditongos: usar sobrescritos. Os sobrescritos são articulações secundárias. No *Handbook of the International Phonetic Association*, o [w] e o [j] sobrescritos, ou seja, [ʷ] e [ʲ], são usados para indicar, respectivamente, labialização e palatalização do segmento precedente, como vimos no capítulo anterior. Transcrever *mau* e *sai*, respectivamente, como ['maʷ] e ['saʲ], é um equívoco, porque não se quer dizer que a vogal [a] que antecede as semivogais na transcrição fica labializada ou palatalizada. Além disso, como os sobrescritos não representam segmentos independentes, essa notação implica que [aʷ] e [aʲ], se existirem, não são ditongos.

Transcrever ditongos ascendentes com sobrescritos já é outra questão. Pode-se considerar que há análises diferentes. Transcrevendo *quadro* e *água* como ['kwadru] e ['agwɐ] estamos considerando que há um ditongo crescente. Transcrever essas palavras como ['kʷadru] e

[^wag^e] implica que se está analisando as sequências <qua> e <gua> como tendo apenas dois segmentos: uma consoante velar labializada, portanto, uma consoante complexa, seguida da vogal [a]. De acordo com essa análise, como há apenas um vocoide, não há ditongo.

3.6 ALGUNS DIACRÍTICOS

Como já vimos no capítulo anterior, a transcrição fonética pode ser mais lata, se contiver menos detalhes, ou mais estrita, se contiver mais detalhes. Com sinais diacríticos podemos acrescentar mais detalhes, como numa espécie de *zoom*. A tabela de diacríticos do Alfabeto Fonético Internacional aparece na parte inferior esquerda da página.

Um diacrítico muito familiar para falantes do português, como já vimos, é o til, <~>, usado com o mesmo valor no Alfabeto Fonético Internacional. Colocado em cima de um símbolo fonético (não necessariamente uma vogal), ele indica que o fone correspondente é nasal.

Como já vimos neste capítulo, em termos de vogais, ele pode ser usado acima de qualquer uma delas. As palavras *um* e *sim* são transcritas, portanto, respectivamente: [ˈũ] e [ˈsĩ].

Um outro diacrítico, o qual tem um uso variável, é o que indica que um segmento é assilábico: [̥]. Ele aparece abaixo de um símbolo vocálico. Como normalmente, por si sós, as vogais formam uma sílaba, se uma vogal ocorre junto a outra, sem formar sílaba separada, ou seja, compondo um ditongo, é comum representar os ditongos não com [j] ou [w], mas com vogais que apresentam esse diacrítico. Pode-se considerar que transcrever um ditongo com uma semivogal ou com o diacrítico de assilabidade são variantes notacionais. A palavra *mau* aparece transcrita de pelo menos três maneiras distintas em textos de fonologia: [ˈmaw], [ˈmaʊ̥] ou [ˈmaʊ̥].

Dependendo da língua que está sendo transcrita, nem se usa o diacrítico assilábico. Em inglês, por exemplo, é comum transcrever *now* como [ˈnaʊ]. Se o objetivo é transcrever apenas uma língua, não há necessariamente problema nisso. Num livro como este, no entanto, que não trata de uma única língua, considero menos adequado representar um ditongo dessa forma. A rigor, uma transcrição como [ˈnaʊ] é ambígua entre representar um ditongo ou um hiato. O uso do diacrítico elimina essa ambiguidade. Transcrever *mau* como [ˈmaʊ̥] ou [ˈmaʊ̥] ou *now* como [ˈnaʊ̥] deixa claro que nas três transcrições temos ditongos, já que o vocoide final é assilábico. Portanto, há apenas uma sílaba.

Não vejo vantagem, no entanto, na obrigação de usar esse diacrítico nesses casos. Uma transcrição como [ˈmaw] já é transparente o suficiente, sem necessitar de diacrítico. Pode-se até argumentar que esse vocoide final na verdade é uma vogal em algum nível mais abstrato, mas em se tratando de uma transcrição fonética, não se justifica, e é menos direto. Por isso, em todo o livro, sempre que houver ditongos em que a semivogal for um vocoide

alto equivalente a [i] ou [u], utilizo os símbolos [j] e [w], independente da tradição de transcrição da língua em questão. Em suma, defendo que transcrever *mau* como ['maw] é melhor do que ['maɯ] ou ['maʊ], e transcrever *pai* como ['paj] é melhor do que ['pai] ou ['paɪ]. Os vocoides que funcionam como semivogais, como já foi dito, têm um caráter intermediário entre vogais e consoantes. Se lembrarmos da definição segundo a qual as vogais não apresentam obstáculo à passagem do ar e as consoantes apresentam, uma semivogal vai normalmente estar no limite entre apresentar um obstáculo ou não. Dentre os vocoides, os que mais se aproximam de apresentar um obstáculo são os mais fechados: os vocoides altos. Por isso, as semivogais por excelência, [j ɥ w], são equivalentes a vogais altas, respectivamente [i y u]. Há línguas, no entanto, que utilizam outros vocoides como semivogais. É o caso do romeno, em que tanto [e] quanto [o] podem ser semivogais. Temos palavras como *veac* 'século' e *poartă* 'porta'. As duas contêm ditongos crescentes em que o vocoide que está em posição na margem da sílaba é um vocoide médio. Como essa é uma situação incomum, não há símbolos sem diacríticos para representar isso. Nesse caso, o ideal é que sejam transcritos com o diacrítico [̘] subscrito: ['vɛ̘ak] e ['pɔ̘artə]. No experimento feito em Chitoran (2002), os ditongos [ja] e [ɛ̘a] no romeno apresentaram uma diferença significativa estatisticamente, ao passo que [wa] e [ɔ̘a] não apresentaram, ou seja, [ɛ̘a] é pronunciado distintamente de [ja], mas [ɔ̘a] pode alternar com [wa] na mesma palavra.

No português temos, além desses ditongos orais, ditongos nasais. Como fica implícito no termo usado, o ditongo inteiro fica nasal, não só a vogal. A semivogal, também. Nesse caso, fazendo uma transcrição fonética, devemos usar o til tanto sobre a vogal quanto sobre a semivogal. No teclado normal não podemos fazer isso, mas nos editores de texto, no menu inserir símbolo, há os subconjuntos IPA, caracteres de modificação de espaços e combinando sinais diacríticos. Esses últimos são usados quando precisamos colocar, por exemplo, um <~> em cima de um <j> ou um <w>. Vejamos como fica a transcrição dos ditongos nasais do português com alguns exemplos, lembrando que estou apresentando três possibilidades de transcrição da vogal central nasal.

- (6) mãe: ['mẽj], ['mɜ̃j] ou ['mẽj̃]
 não: ['nẽw̃], ['nɜ̃w̃] ou ['nẽw̃j̃]
 põe: ['põj̃]
 muito: ['mũjt̃u]

Observe agora essas palavras:

- (7) quando: ['kwẽdu]
 pinguim: [pĩ'gwĩ]

Repare na diferença entre os ditongos crescentes (semivogal seguida de vogal) e os decrescentes (vogal seguida de semivogal). Nos decrescentes, o ditongo inteiro se nasaliza. Nos crescentes, não. A semivogal permanece oral. Portanto, apesar de podermos considerar que há ditongo nessas palavras, não há ditongo nasal.

No capítulo anterior, ao tratar das consoantes nasais, mencionei a nasal palatal do português: [ɲ]. Apesar de normalmente pensarmos no som grafado <nh> em português como uma consoante nasal, nem sempre ele tem uma pronúncia estritamente consonantal. O que define as consoantes nasais é o fato de haver o fechamento total da cavidade oral, enquanto o ar passa pela cavidade nasal. Ora, na fala de muitos falantes do português do Brasil não chega a haver contato da língua com o palato na pronúncia de palavras como *unha*. Nesse caso, não temos uma consoante nasal, mas sim uma semivogal nasal.

(8) unha: ['ũɲɐ] ou ['ũjɐ] (como se fosse *uia* nasalizado)

O fato de haver essa variação entre a consoante nasal e a semivogal nasal faz ficar perfeita a rima raríssima, genial, de Caetano Veloso em *Qualquer Coisa*, nos versos: “*Quero que você ganhe / Que você me apanhe / Sou o seu bezerro gritando mamãe!*” A forma *apanhe* pode ser pronunciada com três sílabas. Considerando a variação na produção do som que vem logo após a vogal tônica, temos as pronúncias [a'pẽɲɪ] e [a'pẽjɪ]. A segunda pronúncia apresenta uma semivogal que equivale a um [i] nasal, seguida de um [i] reduzido. A distintividade dos dois sons na sequência é mínima. Com isso, se tem a pronúncia [a'pẽj], que rima perfeitamente com ['mẽj]. O mesmo se aplica à palavra *ganhe* na mesma obra.

Um último diacrítico que representa uma sutileza na pronúncia à qual nem sempre atentamos é um pequeno círculo subscrito. Ele indica que o fone em questão, embora normalmente seja sonoro ou vozeado, está sendo pronunciado sem vibração das pregas vocais, sendo, portanto, surdo ou desvozeado.

Como as vogais átonas finais no português são bastante reduzidas, tanto em duração quanto na energia dispendida em sua pronúncia, pode ocorrer que a corrente de ar fique tão fraca que não faça vibrar as pregas vocais. Querendo transcrever esse detalhe da pronúncia, usamos então o círculo subscrito. A palavra *gato* aparece a seguir transcrita com duas pronúncias: na primeira a vogal final tem vibração das pregas vocais e a segunda não. A palavra *gato*, por exemplo, pode ser pronunciada com a vibração plena das pregas vocais, ou seja, [ˈɡatʊ], mas pode apresentar desvozeamento. Nesse caso, a transcrição mais adequada seria [ˈɡatʊ̹].

3.7 OUTROS DIACRÍTICOS

Assim como no capítulo precedente, apresento aqui mais alguns diacríticos usados na transcrição de vogais.

a) Relacionados ao estado da glote (ou das pregas vocais):

- soproso ou murmurado (*breathy*): as pregas vocais vibram com uma pequena abertura entre elas, produzindo um som de qualidade diferente. O diacrítico [̤] é colocado abaixo da vogal, como em [a̤]. Está presente em línguas como o guzerate (Índia), em que o advérbio ‘muito’ pode ser pronunciado [bəɦu] ou [bə̤u]. A segunda pronúncia tem vogal murmurada.
- crepitoso, rangido, tremulado, laringalizado ou glotalizado: graficamente tem o mesmo formato do til, mas é colocado embaixo da vogal, como em [a̰]. Esse diacrítico indica que a vogal é pronunciada no extremo mais grave da nossa extensão vocal. A voz fica tão grave que quase se ouvem rangidos. Esse tipo de estado da glote está presente no munduruku (Amazonas e Pará), como em [wḭ̃d̰a̰], ‘onça’.

b) Posição horizontal:

- avançado: como vimos no capítulo anterior, o diacrítico [̟] indica que o som transcrito é pronunciado num ponto mais anterior. A transcrição estrita [ɔ̟] representa uma vogal que fica entre posterior e central. Um bom exemplo é a vogal de *good* ['gʊ̟d] ‘bom’ em inglês, a qual é bastante anteriorizada, se comparada à de *look* ['lʊk] ‘olhar’.
- recuado (retraído): como vimos no capítulo anterior, o diacrítico [̠] indica que o som transcrito é pronunciado num ponto mais posterior. A transcrição estrita [ɪ̠] representa uma vogal que fica entre anterior e central. **exemplo?**
- centralizado: o trema pode ser utilizado como diacrítico de centralização. As duas vogais transcritas nos dois parágrafos anteriores poderiam ser transcritas assim: [ö̥] e [i̥]. **exemplo?**

c) Posição vertical:

- alçado (ou elevado): na transcrição estrita [ɥ], o diacrítico [̥] indica que a vogal transcrita não tem abertura máxima, tendo uma abertura intermediária entre [ɥ] e [ɔ]. A língua fica, portanto, numa posição ligeiramente elevada, ou alçada. Uma transcrição alternativa seria [ɥ̥], usando o diacrítico seguinte. **exemplo?**
- abaixado: na transcrição estrita [ɛ̣], o diacrítico [̣] indica que a vogal é pronunciada com a língua numa posição ligeiramente mais baixa do que sem o diacrítico. Nesse caso,

[e] é uma vogal média fechada, ou média alta. A vogal [ɐ] é média média. Uma transcrição alternativa seria [ɛ], usando o diacrítico precedente. **exemplo?**

3.8 A REPRESENTAÇÃO DO ESPAÇO VOCÁLICO NUM TRAPÉZIO

Como disse no início do capítulo, embora eu siga a Associação Fonética Internacional em utilizar um trapézio para representar o espaço vocálico, o trapézio oficial de vogais tem o lado direito com ângulos retos. Isso implicaria que não há distinção em termos de recuo da língua entre as vogais [u], [o] e [ɔ], ou talvez, melhor ainda, acusticamente o F2 seria das três vogais seria equivalente (ver seção 5.6). Escudero et al. (2009), no entanto, trazem dados que indicam que há, sim, aumento de F2 médio de [u] para [o] e de [o] para [ɔ] tanto no português do Brasil quanto no de Portugal. Esses dados talvez só sejam compreendidos depois da seção 5.6, mas cito aqui o F2 das vogais tônicas em vozes femininas (F) e masculinas (M) de falantes do português do Brasil (PB) e do português europeu (PE) citados pelos autores. Assinalo aqui que o F2 é fortemente correlacionado ao avanço ou recuo da língua das vogais anteriores e posteriores.

	/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
F (PB)	2676	2468	2271	1627	1054	893	812
M (PB)	2198	2028	1831	1329	927	804	761
F (PE)	2760	2508	2360	1662	1118	921	862
M (PE)	2161	1987	1836	1365	934	843	814

Tabela 1. Escudero et al. (2009).

A tabela apresenta as médias geométricas do segundo formante (F2), não incluído o desvio padrão. Como podemos ver, a diferença do F2 é de fato menor do que a encontrada entre as vogais anteriores. Comparando a diferença entre as vogais mais distantes entre anteriores e posteriores, vemos que a diferença entre o F2 de /i/ e de /ɛ/ é sempre bem maior do que a que há entre /ɔ/ e /u/.

	/i/ - /ɛ/	/ɔ/ - /u/
F (PB)	405	242
M (PB)	367	166
F (PE)	400	256
M (PE)	325	120

Tabela 2. Baseada em dados de Escudero et al. (2009).

O $\Delta F2$ das vogais posteriores com relação ao das vogais anteriores fica por volta de 60% e nas vozes femininas e por volta de 40% nas vozes masculinas. Ou seja, ele oscila por volta

de 50%. Trocando em miúdos, a diferença no avanço da língua entre /i/ e /ε/ é aproximadamente o dobro da diferença no recuo da língua entre /ɔ/ e /u/.

Podemos pensar em três alternativas para o trapézio:

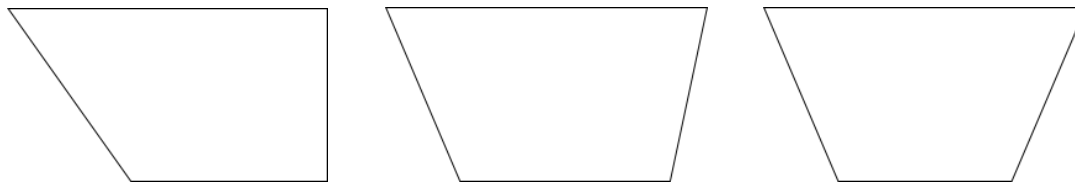


Figura 22. Três representações alternativas do espaço vocálico.

Na primeira representação, a oficial do AFI, o lado direito é vertical, formando ângulos retos com as linhas horizontais. Talvez não seja muito fácil ver a diferença entre o segundo e o terceiro. A segunda representação é um trapézio escaleno, com todos os lados e ângulos diferentes, tendo um ângulo mais fechado na parte superior à esquerda do que na parte superior à direita. Essa representação seria adequada para indicar a menor diferença entre os F2 das vogais posteriores do que entre as anteriores, o que sugere um avanço maior nas posteriores do que o recuo nas posteriores. A terceira representação é um trapézio isósceles, com dois ângulos iguais entre si em cima e dois iguais entre si embaixo. É uma representação que abstrai a diferença entre o avanço e o recuo máximos da língua, representando apenas que há avanço nas anteriores e recuo nas posteriores.

Em vista dos dados apresentados nas duas tabelas acima, utilizei aqui a segunda representação. Cumpre observar que a situação do português é recorrente nas línguas em geral. Principalmente se considerarmos o contraste entre vogais abertas e as demais, é muito comum haver um recuo bem maior da língua nas vogais posteriores não abertas.

Um último ponto a assinalar é que há ainda a questão de se esse diagrama representa dados articatórios ou acústicos, ou seja, a posição da língua ou os sons que ouvimos, o que não necessariamente seria consensual entre todos os pesquisadores.

3.9 DUAS VOGAIS DO INGLÊS

Dada a difusão do inglês e a conseqüente repercussão que os equívocos no que se diz quanto à fonologia do inglês podem trazer, considere importante desfazer dois deles. A primeira vogal presente no inglês de que quero tratar é o *schwa*. É comum lermos frases como *I'd like to be a schwa, because it's never stressed* (Eu gostaria de ser um schwa, porque ele nunca é/está acentuado/estressado), que brincam com a polissemia de *stressed*, termo que pode

significar tanto ‘acentuado’ quanto ‘estressado’. Pode ser divertido, ou até *cute*, mas a rigor isso não é verdade de forma mais abrangente.

À primeiro vista, poderíamos dizer que isso se aplica à pronúncia de referência do inglês britânico, a *Received Pronunciation* (RP), pois nela o [ə] como vogal simples nunca é acentuado. Mas mesmo na RP, ocorre *schwa* tônico em ditongo, como em *so* [ˈsəʊ]. Como o núcleo da sílaba é a vogal, e não a semivogal, se formos dizer que algum segmento é acentuado, é justamente o [ə], e não o [w]. Portanto, nem em RP, essa afirmação é correta. No inglês da América do Norte, ela não é verdadeira nem com relação às vogais puras, pois algumas pronúncias de *love* lá encontradas podem muito bem ser transcrita [ˈləʊv].

Não ficando com nosso olhar restrito ao inglês, deve ser ressaltado também que há ainda diversas línguas em que o [ə] claramente pode ser tônico. Ilustro com algumas línguas, mencionando um exemplo de cada. Em galês a palavra *ysbyty* ‘hospital’ é pronunciada [əsˈpəti]. Temos também: em romeno, palavras como *măr* ‘maçã’ [ˈmər], e em luxemburguês, palavras como *Dës*ch [dɛʃ] ‘mesa’; no catalão das Ilhas Baleares temos também palavras como *set* ‘sede’ [ˈsət]. Em suma, como brincadeira e sem examinar em detalhe a pronúncia do inglês, pode-se afirmar que o *schwa* nunca é acentuado, mas como afirmação fonológica, ela não se sustenta.

Outra questão relacionada é o uso do [ʌ] para transcrever a pronúncia britânica de palavras como *cut* e *must*. Se consultarmos o trapézio vocálico do alfabeto fonético internacional, vemos que esse símbolo representa uma vogal posterior, e não central. Como apontam Roca e Johnson (1999: 135), a vogal dessas palavras tinha uma pronúncia realmente posterior na *Received Pronunciation* antes da II Guerra Mundial. Depois disso, essa vogal se centralizou. O inconveniente é que, apesar disso, a transcrição não foi alterada. Dada a difusão do inglês, essa imprecisão da transcrição tem o grande inconveniente de não fazer a distinção entre uma vogal central e outra posterior, de maneira análoga ao que discuti com relação ao símbolo fonético [a]. Uma transcrição provavelmente mais adequada para essa vogal do inglês britânico seria [ɐ]. Essa transcrição, aliás, é usada no próprio *Handbook of the International Phonetic Association*. As palavras *worry* e *hurry*, normalmente transcritas [ˈwʌɹi] e [ˈhʌɹi] são transcritas como [ˈwɛɹi] e [ˈhɛɹi] no *Handbook* (p. 5 da edição de 1999), o que é mais adequado.

3.10 EVITANDO CONFUSÕES

É preciso cuidado para não confundir os símbolos [ɣ] e [ʁ]. É muito comum eles serem trocados, inclusive em bons textos de fonética e fonologia, seja por engano dos autores, seja por confusão na preparação do texto. O primeiro é a letra gama minúscula do grego e

O símbolo [ɣ] representa uma **vogal** média alta posterior não arredondada, presente, por exemplo, no estoniano *sõna* ['sɔna] ‘palavra’. Observe que a parte inferior da letra não fica abaixo da linha. Até 1928, esse som era representado por um A invertido, [ʏ]. De 1928 até 1989, o símbolo foi o chamado *baby gamma*, um gama minúsculo com a volta na parte inferior reduzida e inteiramente escrita acima da linha. Justamente para diferenciá-lo melhor do [ɣ], a partir de 1989 ele passou a ter a parte superior curva: [ɣ̥]. A figura 3 contrasta o *baby gamma* e o símbolo atual. Como o *baby gamma* deixou de fazer parte do sistema unicode, é difícil representá-lo de forma correta no texto, pois ele é automaticamente substituído pela versão com curvas na parte superior. Por isso, apresento as duas na figura 24.



Com relação às vogais baixas posteriores, não é demais lembrar que o símbolo [a] representa uma vogal não arredondada, e o símbolo [ɒ], uma vogal arredondada.

EXERCÍCIOS

- 1) Transcreva as palavras seguintes. Se puder, transcreva mais de uma pronúncia. A que você normalmente usa e outra(s), que podem ser de outra região, mais formais, etc. Em alguns casos já indico entre parênteses um número de pronúncias possíveis a ser transcrito.

nata	leve	ovo	sujo	nossa
chego	livro	quero	graxa	cola
hélice	cinto	Brasil	dezoito	
barraca	certo	ponte (3)	circuito (2)	folha (2)
vinho	junta	ponto	sempre	janta
põe	mãe	vem	ruim (2)	chão

- 2) Neste exercício, transcreva as palavras a seguir, sempre representando variação. Os segmentos ou sequências cuja variação na pronúncia deve ser representada nas transcrições estão sublinhados. O número entre parêntesis sugere um número de pronúncias diferentes a serem transcritas, mas pode haver mais possibilidades.

poeta (3)	hom <u>e</u> m (4)	mul <u>h</u> er (3)	
grat <u>u</u> ito (2)	fol <u>h</u> a (2)		
f <u>e</u> ira (2)	pou <u>co</u> (2)	nós	três
sust <u>o</u> (2)	ves <u>g</u> o (2)	mes <u>m</u> o (3)	
sol <u>t</u> o (3)	ant <u>i</u> go (2)	<u>d</u> ireto (2)	

- 3) Verifique nos pares de vogais a seguir quais têm o mesmo grau de abertura.

[y u]	[ɔ ʌ]	[ε ɔ]	[ɑ ɐ]	[y ʉ]
[æ ɑ]	[i u]	[ɪ ʊ]	[ʌ o]	[e ʉ]
[ø ɐ]	[ɒ a]	[y ʏ]	[ʌ ε]	[e i]
[ʉ u]	[ε ø]	[ʌ e]	[ɪ u]	[œ ɔ]

- 4) Verifique nos pares de vogais a seguir quais coincidem com relação ao arredondamento dos lábios: presença ou ausência.

[y u]	[i y]	[ε ɔ]	[ɑ ɔ]	[ɒ u]
[i ɪ]	[ʉ ʊ]	[ə e]	[ʌ u]	[ø ʉ]
[ʌ a]	[æ ɑ]	[e ə]	[ʌ ɔ]	[ɪ o]

- 5) Verifique nos pares de vogais a seguir quais têm a mesma posição horizontal da língua, ou seja, são ambas anteriores, ambas centrais ou ambas posteriores.

[ʌ a]	[ɔ ɑ]	[e ə]	[ʌ ʏ]	[ɪ ɨ]
[ʊ ɔ]	[e ə]	[ɛ ɜ]	[ø œ]	[æ ɛ]
[ɨ ə]	[a ɨ]	[ʊ ʌ]	[ɤ o]	[y u]
[ɑ æ]	[e ɜ]	[ɪ ʏ]	[ʊ u]	[o ø]

- 6) Qual é a vogal intrusa em cada grupo? Por quê? Em geral há mais de uma possibilidade, dependendo do critério adotado.

[ʌ ɔ a]	[ɔ u ɑ]	[e e ə]	[y ʏ u]	[ɪ ɪ ɨ]
[ɜ u ʌ]	[ʊ ɤ ɑ]	[ɛ ɤ o]	[o ø œ]	[a æ ɛ]
[ɨ ə u]	[y ø ɐ]	[ʊ u o]	[ɛ ɜ ɔ]	[a ɐ u]

- 7) O que distingue articulatoriamente as vogais de cada par? Pode haver mais de uma distinção.

[y u]	[ɔ ɤ]	[ə e]	[ɛ ɔ]	[ʌ o]
[œ ɔ]	[y ʊ]	[ɪ ɨ]	[ɔ ɑ]	[æ e]
[ɑ a]	[ʊ u]	[ø e]	[e ə]	[ɐ œ]

- 8) Observe os dados a seguir do islandês. Algumas palavras apresentam variação na pronúncia, podendo ter uma vogal longa (seguida do diacrítico [:] na transcrição) ou um ditongo. Qual classe de vogais apresenta em variação uma pronunciada ditongada?

['li:ta] 'olhar'	['li:tyr] 'cor'
['lɛ:ti] / ['lɛɛti] 'preguiça'	['tʰa:la] 'fala'
['sy:maɹ] 'verão'	['fœ:ɣyɹ] / ['fœœɣyɹ] 'bonita'
['kʰɔ:ma] / ['kʰoo:ma] 'vir'	['kʰu:la] 'bola'

- 9) Observe os dados a seguir, que representam os ditongos em frísio ocidental (Holanda). Desenhe um diagrama vocálico ou consulte a tabela do alfabeto fonético internacional, e classifique os ditongos em dois grupos, com base no segundo vocóide do ditongo. Como poderíamos descrever a trajetória de cada ditongo no espaço vocálico? Obs: com

vimos na seção 3.5, o símbolo fonético [ɥ] representa a semivogal correspondente à vogal [y].

stien	['stiɐ̃n]	'pedra'	beam	['biɐ̃m]	'árvore'
flues	['flyɐ̃s]	'membrana'	hout	['hawt]/ ['hɔwt]	'madeira'
boat	['boɐ̃t]	'barco'	stuit	['stœɥt]	'momento'
bloed	['bluɐ̃t]	'sangue'	kaai	['ka:j]	'cais'
hy	['hɛj]	'ele'	hoed	['huɐ̃t]	'chapéu'
maits	['majts]	'larva'	freon	['frøɐ̃n]	'amigo'

10) Embora eu tenha afirmado que as vogais nasais do português são cinco, **tóme, nóme**

LINK

A página <https://ipa.typeit.org/full/> tem quase todos os símbolos fonéticos. É possível digitar transcrições *online* e copiá-las para o editor de texto. Também é possível instalar o programa para PC.

REFERÊNCIAS E LEITURAS COMPLEMENTARES

- Bossaglia, Giulia (2019). *Linguística Comparada e Tipologia*. São Paulo: Parábola.
- Chitoran, Ioana (2002). *A perception-production study of Romanian diphthongs and glide-vowel sequences*.
- Lindblom e...
- Roederer, Juan G. (1998). *Introdução à Física e Psicofísica da Música*. (Trad.: Alberto Luís da Cunha.) São Paulo: EDUSP.
- Trubetzkoy, Nikolai (1939). *Grundzüge der Phonologie*. Praga: Travaux du cercle linguistique de Prague. (há traduções em inglês e francês)
- Walker, Rachel. **Nesse capítulo?**

CAPÍTULO 4

DISTRIBUIÇÃO, CONTRASTE, NEUTRALIZAÇÃO, SISTEMAS

Neste capítulo nos familiarizaremos com os conceitos que foram propostos na fonologia estruturalista, no começo do século XX. São eles a distribuição, que trata de onde os fones podem ocorrer dentro de uma palavra; o contraste, que trata da possibilidade de distinguirmos significado usando sons diferentes; a neutralização, processo que faz com que dois ou mais fones que podem distinguir significado em certos contextos não possam fazê-lo em outros; e os sistemas, que contêm o conjunto de sons, e suas relações e oposições numa língua ou variedade linguística específica.

4.1. PRIMEIROS CONCEITOS DA FONOLOGIA

Depois de termos vistos os sons de uma perspectiva essencialmente de fonética articulatória, ou seja, examinando principalmente como cada consoante e cada vogal é produzida, iniciamos neste capítulo o exame dos sons da língua da perspectiva específica da fonologia.

Uma coisa é pensar o som, ou o fone, apenas em seus aspectos sonoros: como ele é produzido, transmitido, ou mesmo percebido. Outra é considerar como ele funciona linguisticamente. Para isso, incorporamos ao estudo do som a perspectiva gramatical. Um ponto de partida conveniente é a dicotomia saussureana que considera as duas faces do signo linguístico. Para Saussure, o signo linguístico, do qual um exemplo simples seria uma palavra monomorfêmica como *sol*, é formado de duas faces. Uma delas é o significado, o conteúdo que se quer transmitir. No caso da palavra *sol*, é o do astro que é o centro do sistema solar, que brilha no céu durante o dia e fica oculto à noite. Podemos considerar que o significado é algo presente na mente dos falantes, algo mental. Para comunicar o conteúdo de nossos pensamentos a outras pessoas, não há como transmitir o significado diretamente a essas pessoas na forma de significado. Como o significado é mental, e não temos o dom da

telepatia, ele precisa ser transmitido através de algo captável pelos sentidos, o significante. O significante pode ser composto de gestos feitos principalmente com as mãos, como no caso das línguas de sinais; pode ser construído com sons produzidos pelo aparelho fonador; ou pode ainda ser constituído de sinais gráficos. Como na maioria das sociedades os falantes de línguas orais são em maior número, o nome dado à disciplina que estuda essa forma de manifestação do significado é fonologia, o estudo do som. Apesar de o significante nas línguas de sinais não ser sonoro, também se usa o termo fonologia ao falar delas.

Embora muitos tratem a fonologia e a fonética como disciplinas absolutamente distintas, considero que é mais adequado tratá-las como aspectos diferentes de uma mesma área, questão que retomarei no capítulo 14. O que as distingue é que o peso dado por cada uma das duas disciplinas a cada aspecto dos sons é diferente. A fonética em grande parte se baseia em aspectos concretos, observáveis, mensuráveis. Já a fonologia, embora leve em conta esses aspectos, pende muito mais para o lado gramatical e cognitivo. Em termos saussureanos, a fonética estuda os sons essencialmente como significantes, enquanto que uma característica importante da fonologia é que ela examina com atenção também a relação entre esses significantes e o significado.

Já desde antes do estruturalismo, no século XIX, se sabe que uma propriedade importante dos sons linguísticos é a capacidade de eles transmitirem significado. Sua relação com o significado está na base de uma série de conceitos do estruturalismo que veremos a seguir. Iniciemos a análise fonológica por etapas. Se tivermos duas sequências sonoras como [i'fo] e [i'vo] numa língua qualquer, ou seja, que duas sequências que apresentem diferença em pronúncia em apenas uma posição da cadeia sonora, em um único som ou fone, que podemos denominar *segmento*, temos um par suspeito. O termo *par suspeito* se refere à etapa da análise em que ainda não começamos a relacionar o significante com o significado. Por isso me referi a uma língua qualquer, para não vincular o significante a um possível significado. Na maioria dos casos, o mais relevante é compararmos fones foneticamente semelhantes. É altamente improvável que seja relevante compararmos [k] e [a], por exemplo.

Passando para o português, consideremos dois fones vistos no capítulo das consoantes: a oclusiva coronal sonora e o tepe, que também é coronal e sonoro. Os dois ocorrem no português. Temos sequências pronunciadas ['fado] e ['faru] ou ['mɔdɐ] e ['mɔrɐ]. Os membros de cada par diferem apenas em um segmento, sendo, por conseguinte, pares suspeitos.

Se consideramos também o significado, vemos que a troca de uma consoante pela outra nesses exemplos alterou completamente o significado: no par *fado* e *faro*, a palavra *fado* é um estilo de música típico de Portugal, e *faro* é o sentido do olfato bem desenvolvido.

Quando temos duas formas assim, que se distinguem por um único fone, e isso afeta o significado, esse par de formas é denominado *par mínimo*. Como foi dito, na maioria dos casos, os fones mais relevantes para serem comparados são os que apresentam uma grande semelhança fonética. No caso do [d] e [r], as duas consoantes são pronunciadas com a coroa da língua entrando em contato com o céu da boca. Basicamente, a única diferença entre os dois fones é que na produção do [d] a coroa da língua encosta no céu da boca e é mantida ativamente em contato com ele, permanecendo assim alguns centésimos de segundo, enquanto que no caso do [r], a coroa da língua não fica em repouso lá: ela atinge o céu da boca e logo se afasta. Embora possa variar o sentido atribuído ao termo *céu da boca*, utilizo-o aqui para me referir a toda a parte superior da cavidade oral que vai dos alvéolos à região velar.

É importante deixar claro que não são só palavras distintas que contam como par mínimo. Se são sequências sonoras com significado diferente funcionam como par mínimo. Um exemplo é o da palavra *man* ['mæn] 'homem' em inglês e seu plural *men* ['mæn]. São duas formas da mesma palavra, mas têm significados distintos: a primeira é singular e a segunda é plural. Exemplos semelhantes em português seriam *foi* ['foi] e *fui* ['fui]: a primeira é uma forma de terceira pessoa, e a segunda é uma forma de primeira pessoa. Logo, é um par mínimo também.

Mas para entendermos mais claramente os conceitos de par suspeito e par mínimo e examinar o papel dos sons no significado por etapas, nos detendo primeiramente apenas no significante, é mais proveitoso olhar uma língua que não falamos. Isso é difícil na nossa própria língua materna, porque olhamos o significante e imediatamente já pensamos no significado. Observemos agora, então, os seguintes dados do shanenawa (Acre). Nessa língua encontramos os seguintes significantes: [ru'no?] e [du'no?]. Em princípio, temos a mesma situação encontrada no português. A única distinção entre os dois é que no primeiro temos um tepe no início e no segundo temos uma oclusiva coronal sonora. Temos, então, um par suspeito. Foi feita a mesma substituição de fones que fizemos no português. Se observarmos o significado, no entanto, veremos que o efeito é diferente. No português, a troca do [d] pelo [r] produziu uma alteração no significado. No shanenawa, no entanto, isso não aconteceria, já que as duas são pronúncias diferentes da palavra que significa 'cobra'. Essa é uma distinção fundamental. Às vezes uma pronúncia diferente faz com que tenhamos uma nova palavra, mas às vezes não. No estruturalismo, exemplos como os do português fazem com que consideremos que os dois fones considerados, [d] e [r], são *fonemas* diferentes, são fones que têm a possibilidade de distinguir significado no português. Para representar esse estatuto diferente, e indicar que estamos falando de fonemas, e não mais de fones, representamos os fonemas entre barras: /d/ e /r/.

Uma lição importante do estruturalismo é que o que funciona como fonemas distintos numa língua não necessariamente funciona da mesma forma em outra língua. Se eles não forem fonemas diferentes, são realizações diferentes do mesmo fonema ou, usando o termo técnico, são *alofones*, em que *alo-* significa ‘outro’. Isso é equivalente ao que ocorre na escrita: <a> e <ɑ> são realizações diferentes da mesma letra ou grafema. Se eu escrever meu nome como <Paulo> ou como <Paulo>, isso não faz diferença. É o mesmo nome que está sendo escrito. Da mesma forma, em shanenawa, se alguém disser [ru'noʔ] e [du'noʔ], estará pronunciando a mesma palavra de duas maneiras diferentes. Portanto, [d] e [r] são alofones em shanenawa.

Essas duas formas do shanenawa, [ru'noʔ] e [du'noʔ], são um par suspeito, pois diferem pela presença de um [r] numa delas e de um [d] na mesma posição na outra, mas como essa diferença sonora não afeta o significado, esse par não chega a constituir um par mínimo.

Para exercitar a distinção entre esses dois tipos de conceitos, vejamos mais alguns dados. Se forem fones encontrados no português ou em alguma língua com que temos mais familiaridade, é sempre interessante comparar a situação encontrada nessas línguas.

Examinemos inicialmente alguns dados do armênio oriental (o diacrítico sob a vogal indica que ela é um pouco posteriorizada):

- | | | |
|-----|-------|-------|
| (1) | [gin] | [kin] |
| | [bɑh] | [pɑh] |
| | [dɑt] | [tɑt] |

Olhando inicialmente apenas a pronúncia dos pares de dados de cada linha acima, vemos que em cada um deles, a única diferença é que a forma na primeira coluna tem uma oclusiva sonora, ao passo que a forma na segunda coluna apresenta uma oclusiva surda. São pares suspeitos, isto é, pode ser que sejam pares mínimos, mas só podemos ter certeza quando soubermos o significado de cada forma.

O significado de cada uma é o seguinte:

- | | | |
|-----|--------------------|-----------------|
| (2) | [gin] ‘lugar’ | [kin] ‘mulher’ |
| | [bɑh] ‘pá’ | [pɑh] ‘momento’ |
| | [dɑt] ‘julgamento’ | [tɑt] ‘avó’ |

Vemos, portanto, que em cada linha temos de fato um par mínimo. Considerando os três exemplos conjuntamente, constatamos que a diferença entre surda e sonora é *contrastiva* ou *distintiva*: não apenas são fones diferentes, mas a troca de um por outro produz diferença ou contraste no significado. A situação é a mesma do português com essas consoantes, já que temos pares mínimos como *gola* e *cola*, *bar* e *par*, ou *dom* e *tom*.

Vejamos agora outros pares do armênio oriental, os quais envolvem fones inexistentes no português.

- (3) [t^hur] [tur]
 [k^hujr] [kujr]

Aqui os fones iniciais de cada forma são fones inexistentes no português, oclusivas surdas aspiradas, transcritas com um [h] sobrescrito após a consoante, isto é, [t^h]. O par de cada linha envolve uma oclusiva aspirada e outra não aspirada. A questão, portanto, é saber se o fato de a oclusiva ser aspirada ou não faz diferença em termos de significado. Observemos o significado dessas formas:

- (4) [t^hur] ‘sabre’ [tur] ‘dê!’
 [k^hujr] ‘irmã’ [kujr] ‘cego’

Novamente, como visto nos exemplos do grupo anterior, a distinção entre oclusiva aspirada e não aspirada é contrastiva: cada par contém fonemas distintos.

Examinemos agora um último par de fones do armênio oriental: o tepe e a vibrante.

- (5) [sar] ‘montanha’ [sar] ‘frio’
 [lur] ‘notícia’ [lur] ‘silêncio’

Como vemos nos dados, temos também um par de fonemas. Esses dados podem ser surpreendentes para falantes de português, língua em que só é possível contraste entre esses dois fones em posição intervocálica.

Vejamos agora as duas vibrantes encontradas no holandês: [r] e [ʀ]. Elas variam com diversas outras possibilidades de realização, as quais não são vibrantes, mas me limito aqui a apresentar essas duas.

- (6) [rɛjk] ou [ʀɛjk] ‘rico’
 [ˈpra:t] ou [ˈpʀa:t] ‘fala’
 [ˈdo:r] ou [ˈdo:ʀ] ‘através’
 [ˈstra:t] ou [ˈstʀa:t] ‘rua’

Como se pode ver, no holandês (ou neerlandês), os fones [r] e [ʀ] são alofones. A situação é de fato, bastante próxima à do português do Brasil, pois esse <r> corresponde a inúmeras realizações diferentes, embora haja diferenças.

Por fim, vejamos o estatuto fonológica da relação [x]-[h] em hebraico.

- (7) [jɛʃ la'xɛm] ‘vocês (m.) têm’ [jɛʃ la'hɛm] ‘eles têm’
 [ʃɛla'xɛm] ‘de vocês (m.)’ [ʃɛla'hɛm] ‘deles’

Embora no português do Brasil [x] e [h] sejam alofones, pois palavras como *rua* e *par* poder ser pronunciadas, respectivamente, como ['xue] ou ['hue], e como ['pax] ou ['pah], vemos que no hebraico eles distinguem significado, sendo, portanto, fonemas diferentes.

Esses exemplos ilustram claramente se a relação entre dois fones se trata de contraste fonológico, também denominado contraste fonêmico, ou não. Entretanto, às vezes, por exemplo, quando se trata de fones com baixa frequência de ocorrência, não conseguimos nem encontrar pares suspeitos. Nesses casos, pode-se recorrer aos chamados pares análogos, pares que apresentam mais de um segmento diferente, caso se suponha que um não interfere no outro.

Um exemplo frequentemente citado é do inglês. Nele até existe par mínimo para ilustrar o contraste entre [ʃ] e [ʒ]: o par *Aleutian* [ə'lu:ʃən] 'habitante ou nativo das ilhas Aleutas' e *allusion* [ə'lu:ʒən] 'alusão', nenhuma das quais é exatamente uma palavra comum. É um dos pouquíssimos pares do tipo em inglês. Caso não se esteja conseguindo localizar um par desses, com diferença na pronúncia apenas de um segmento, pode-se usar pares com duas diferenças, contanto que seja razoável considerar que a diferença em uma posição não interfere na diferença da outra.

No inglês, pode-se usar também o par *vision* e *fission*, por exemplo. Essas duas palavras são pronunciadas ['fɪʃən] e ['vɪʒən], e cognatas de suas traduções em português, *visão* e *fissão*, têm também duas diferenças na pronúncia. A consoante inicial é [f] numa e [v] na outra. A consoante do início da segunda sílaba é [ʃ] e [ʒ] na outra. Como normalmente a sonoridade de uma consoante não afeta a sonoridade de outra consoante que não esteja vizinha a ela, podemos desconsiderar a presença de um segundo contraste na palavra e, apesar de não ser o tipo de evidência perfeito, concluir que /f/ e /v/ são fonemas distintos em inglês. Os pares análogos não são o recurso ideal para a análise fonêmica, mas em último caso, podem ser usados. Deve ser mencionado, contudo, que nem todos vêm com bons olhos esse recurso.

Na verdade, a abordagem da distintividade não necessariamente se restringe a pares suspeitos (e mínimos). É possível chegar a um quadro dos fonemas (ou segmentos distintivos) de uma língua, mesmo sem nada que se assemelhe a pares suspeitos, partindo apenas dos contextos em que os segmentos ocorrem. Isso será visto na seção 4.3.

4.2. TIPOS DE ALOFONES

Quando detectamos a impossibilidade de dois fones distinguirem significado numa língua, chegamos à conclusão de que eles são alofones. Ocorre, porém, que nem todos os alofones

são do mesmo tipo. Há uma classificação que distingue dois tipos de alofones, os quais não podem distinguir significado por motivos completamente diferentes.

4.2.1 ALOFONES EM VARIAÇÃO LIVRE

É bastante comum que um determinado fonema admita mais de uma pronúncia no mesmo contexto. É como vimos nos exemplos [ru'noʔ] e [du'noʔ] do shanenawa, que significam, ambos, 'cobra'. Um outro exemplo, que já foi dado na seção 3.6, envolve o segmento cuja ortografia em português é <nh>. Como vimos na seção 3.6, há duas pronúncias possíveis no português do Brasil para esse segmento. Uma em que o segmento tem uma pronúncia consonantal, transcrita com o símbolo fonético [ɲ], e outra semivocálica em que a língua se aproxima mas não encosta no palato. Esse som é transcrito como [ɲ̃].

O termo variação livre, do estruturalismo, que se desenvolveu na primeira metade do século XX, é anterior à sociolinguística, iniciada na década de 1960 por Labov, a qual já demonstrou suficientemente que a rigor nenhuma variação é livre. Dizer que a variação é livre se limita a implicar que uma alteração no significante não afeta o significado. Essa análise desconsidera todos os fatores extralinguísticos como idade e região de nascimento de quem fala, além de formalidade da situação, entre outros fatores, atendo-se apenas ao sistema linguístico.

4.2.2 ALOFONES EM DISTRIBUIÇÃO COMPLEMENTAR

Além dessa situação em que tanto faz usar um alofone ou outro, os alofones podem ter outro tipo de relação entre si. É comum que, havendo dois alofones, um só possa ser usado em alguns contextos e não em outros, e o segundo só possa ser usado onde o primeiro não pode. Um exemplo bom disso envolve consoantes laterais. Poderíamos até usar a situação mais difundida no português do Brasil, mas como ela envolve um fone que pode ter outro valor fonológico, o [w], utilizarei aqui o inglês que apresenta uma situação equivalente, mas mais simples que a do português do Brasil, pois ambos os fones só são realizações do mesmo fonema.

Os dois fones em questão são a lateral dental ou alveolar simples, transcrita como [l], e um outro tipo de consoante lateral, que apresenta uma articulação dupla, como visto no capítulo 2. Como vemos na figura 25, na lateral não velarizada, há apenas o contato da língua na região dos alvéolos, enquanto que na velarizada há dois pontos de articulação: um ponto de contato na região dos alvéolos e um ponto de aproximação na região velar.

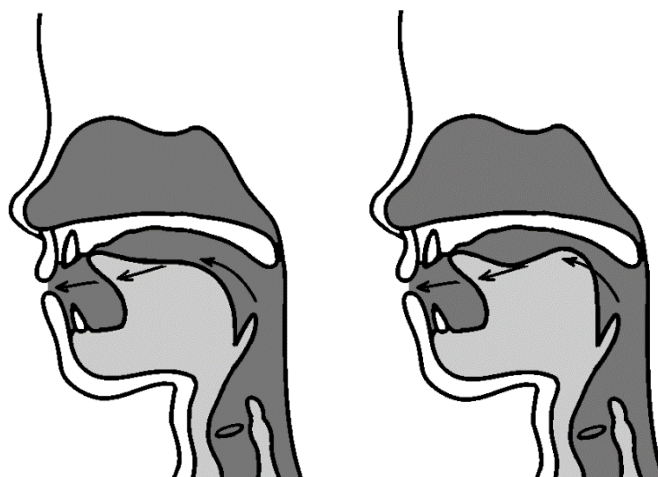


Figura 25. lateral não velarizada e lateral velarizada.

- | | | |
|-----|--------------------------------|------------------------------|
| (8) | sell ['sɛɫ] 'vender' | seller ['sɛlə] 'vendedor' |
| | film ['fɪlm] 'película, filme' | shelter ['ʃɛltə] 'abrigo' |
| | lip ['lɪp] 'lábio' | aloud [ə'laʊd] 'em voz alta' |

Esses dados podem ser analisados com relação à estrutura da sílaba, mas como a sílaba só será vista em detalhes no capítulo 9, pensemos apenas em termos de vogais e consoantes vizinhas e da posição na palavra. Três dos exemplos apresentam a lateral não velarizada: *seller*, *lip* e *aloud*. Os demais apresentam a lateral velarizada: *sell*, *film* e *shelter*. É suficiente pensarmos em termos de vogais e consoantes, além da posição na palavra. A lateral não velarizada só ocorre diante de vogal, e a velarizada só diante de consoante ou no final da palavra. Fica claro que onde uma das laterais pode ser usada a outra não pode. É isso que caracteriza a distribuição complementar. A distribuição de um fone são os contextos em que ele pode ser usado. Como cada um desses dois fones pode ser usado onde o outro não pode, a distribuição de um complementa a do outro.

Como analisar a situação das africadas [tʃ] e [dʒ] no português do Brasil? Depende do conjunto de dados ou de falantes que estivermos considerando. Se tivermos dados apenas da pronúncia comum dos falantes que nasceram nas cidades do Rio de Janeiro e São Paulo, por exemplo, o quadro é o que veremos a seguir. Primeiro há exemplos em que só ocorre a pronúncia [tʃ]:

- (9) taco teto telha touro toca tudo
 traço atleta

Nenhum falante pronunciaria ['tʃetu] nem ['tʃoru], a não ser por brincadeira. Mas por outro lado, temos exemplos em que só há a pronúncia [tʃ] para falantes dessas regiões:

(10) tia tinta parte

Nesse tipo de situação, faz sentido considerar que temos alofones em distribuição complementar.

Se, no entanto, estivermos levando em consideração o Brasil como um todo, o quadro muda, já que no segundo conjunto de dados seria possível também a pronúncia [t], além de [tʃ]. Temos, portanto, variação livre nesse subconjunto de dados, mas só nele. É variação livre restrita a um determinado contexto: diante de vogais anteriores altas, sejam elas orais [i], nasais [ĩ], ou reduzidas [ɪ].

Há ainda variedades do português do Brasil, em regiões da Bahia, por exemplo, em que a presença de uma semivogal palatal antes da oclusiva provoca sua palatalização, o que verificamos em pronúncias como as seguintes:

(11) muito ['mũjtʃu] oito ['ojtʃu]

É possível ainda uma outra análise. Cristófaró Silva e Cantoni (2011) defendem que pares como os seguintes nos indicam que a situação já mudou:

(12) pato pátio tal tchau

Como a pronúncia dessas palavras é ['patu], ['patʃu], ['taw] e ['tʃaw], há razões para considerar que eles já funcionam como fonemas no português do Brasil. O intuito aqui não é fechar a questão, mas ilustrar como é complexo esse tipo de análise, e que é necessário argumentar para defender uma determinada análise.

4.3. ESTUDO DE CASOS

Como foi dito na seção 4.2, a análise da distintividade não necessariamente se restringe a pares suspeitos (e mínimos). Principalmente em casos de alofonia, é possível fazer a análise fonológica de uma língua, mesmo sem nada parecido com pares suspeitos, partindo apenas dos contextos em que os segmentos ocorrem. Examinaremos duas línguas para ver como isso é feito.

Um procedimento útil e que facilita muito uma análise desse tipo é fazer um levantamento dos sons que ocorrem antes e depois de cada ocorrência de determinado fone no conjunto de dados, dispondo esses dados em tabelas. Procedendo assim podemos fazer uma análise fonêmica mesmo sem ter pares suspeitos à disposição.

O primeiro caso a estudar é o do russo, que possui duas consoantes laterais. A rigor, nenhuma delas é equivalente a nossa lateral alveolar/dental. Para distinguir as duas, o ideal

é usarmos diacríticos. A primeira é muito próxima à nossa lateral palatal, e é transcrita com o *j* sobrescrito, [l^j], que indica palatalização. A segunda é próxima da lateral velarizada do inglês e do português europeu mas apresenta um grau de velarização mais acentuado, podendo ser transcrita com um [ɣ] sobrescrito, [l^ɣ], ou com usando o símbolo [ɭ] visto no capítulo 2. Para diminuir a quantidade de diacríticos na transcrição, opto aqui pela segunda opção. Vejamos os dados:

- (13)
- | | |
|---|---|
| [l ^j ʊ'bof ^j] 'amor' | [t'ol ^j kə] 'apenas' |
| [be'l ^j el ^j ɕ:ɪk] 'torcedor' | [ʔizɪ] 'esqui' |
| [l ^j ʌɡʊ] 'vou me deitar' | [gəʔe'va] 'cabeça' |
| [l ^j ist] 'folha' | [dʊɭ] 'soprar (perfeito m. sg.)' |
| [b ^j i'l ^j et] 'bilhete, entrada' | [gʔe'goɭ] 'verbo' |
| [pɭute'vat ^j] 'trapacear' | [et'l ^j it ^ɣ enij] 'excelente' |
| [p ^j l ^j aʃ] 'praia' | [zəl ^j] 'pena' |
| [zəɭ] 'apertar (perfeito m. sg.)' | [m ^j el ^j] 'giz' |
| [toɭk] 'empurrão' | [b ^j iɭ] 'bater (perfeito m. sg.)' |
| [m ^j et ^j] 'ter (perfeito m. sg.)' | [nʊl ^j] 'zero' |
| [ʔejbɭ] 'etiqueta' | [ʔokət ^j] 'cotovelo' |
| [miɭij] 'encantador' | [b ^j il ^j] 'história verdadeira' |

Pode ser que diante de alguns conjuntos de dados, já à primeira vista consigamos detectar a relação entre os fones, mas nem sempre temos essa sorte. Por isso, é bom termos um procedimento para facilitar a visualização dos contextos em que cada fone ocorre. Vamos, então, fazer algumas tabelas que conterão em cada linha o segmento precedente e o seguinte a cada ocorrência do segmento em questão. Começemos por listar os contextos em que ocorre o segmento [l^j]. Ele ocorre nas sequências sonoras [l^jʊ'bof^j], [t'ol^jkə] e assim por diante. Em [l^jʊ'bof^j], o [l^j] aparece no início da palavra e é seguido da vogal [ʊ]. Já em [t'ol^jkə], ele é precedido da vogal [o] e seguido da consoante [k]. São essas as informações que colocaremos na tabela. Quando o segmento ocorrer no início da palavra, colocamos o símbolo de fronteira de palavra (#) antes do segmento para representar isso. Se ele estiver no final da palavra, o símbolo # virá na coluna que indica o contexto subsequente. As ocorrências de [l^j] e de [ɭ] aparecem nos seguintes contextos:

(14)

[l ^j]	
#	ʊ
o	k
e	e

[ɭ]	
#	i
ə	e
u	#

e	ɛ
#	a
#	i
i	e
t	i
p	a
a	#
#	e
e	#
u	#
i	#

e	o
o	#
p	u
a	#
o	k
i	#
e	#
#	e
#	o
i	i

Embora não seja um conjunto muito numeroso de dados, podemos fazer algumas observações. De início, é importante mencionar que o acento não interfere na distribuição dessas duas laterais. Como se pode ver também, ambas ocorrem antes de consoante, em final de palavra, após qualquer vogal, e praticamente antes de qualquer vogal. A conclusão que se impõe é que elas são fonemas diferentes. Os únicos contextos em que não encontramos as duas são: antes da vogal [i], onde só se encontra [lʲ]; e antes da vogal [ɨ], onde só se encontra [ɬ]. Se essas duas vogais são fonemas distintos ou alofones é uma questão espinhosa da fonologia russa, o que complica a análise de qualquer fenômeno que envolva as duas.

Passemos agora a analisar dados da língua gen-gbê (Togo), para verificar se é possível elaborar uma regra que explique a distribuição dos fones [r] e [l] ou se eles são fonemas distintos.

- | | | | |
|------|----------------------------------|-------------------------|---------------------|
| (15) | aḡble ‘fazenda’ | agonglo ‘lagarto’ | aṅoli ‘fantasma’ |
| | ak̂plo ‘lança’ | sabule ‘cebola’ | sra ‘esforço’ |
| | alo ‘mão’ | avlo ‘isca’ | blafog̃be ‘abacaxi’ |
| | dre ‘estender os braços’ | edro ‘sonho’ | exlo ‘amigo’ |
| | exle ‘pulga’ | hle ‘ler’ | ṅlo ‘escrever’ |
| | t̂r̂r̂õ ‘exterminar’ | ṅrã ‘ser feio’ | klb ‘lavar’ |
| | tre ‘cola’ | vlu ‘esticar uma corda’ | lo ‘gostar’ |
| | m̂la ‘bater tambor’ | pleplelu ‘pombo que ri’ | wla ‘esconder’ |
| | zro ‘mosca’ | esro ‘cônjuge’ | etro ‘balança’ |
| | epr̂õ ‘cobra que cospe’ | ḡzro ‘dica’ | |
| | atitrwε ‘pombo de bico vermelho’ | | |

Seguindo o mesmo procedimento que utilizamos com o russo, listemos o segmento precedente e o seguinte de cada ocorrência de [r] e de [l].

(16)

l		r	
gb	e	s	a
g	o	d	ε
ɔ	i	d	ɔ
kp	ɔ	tʃ	õ
u	ε	ɲ	ã
a	ɔ	t	e
v	ɔ	z	o
b	a	s	ɔ
x	ɔ	t	o
x	e	ɲ	õ
h	e	dʒ	o
ŋ	ɔ	t	w
k	ɔ		
v	u		
#	ɔ		
m	a		
p	e		
e	u		
w	a		

Uma observação: como os dois primeiros [l] da palavra *pleplelu* estão exatamente no mesmo contexto, não repeti essa informação na tabela.

Deve ser perceptível que temos aqui uma situação bem diferente da do russo. Apesar de não haver diferenças muito grandes no contexto seguinte a esses dois fones, se olharmos o contexto precedente, vemos que antes de [r] só ocorrem consoantes. Mais do que isso: não é qualquer tipo de consoante que ocorre antes de [r]. São somente consoantes coronais: [s d tʃ ɲ t z dʒ]. Tendo essa informação, se voltamos a examinar a tabela dos contextos do [l], vemos que nenhuma consoante coronal ocorre antes dele. Portanto, temos aí claramente um caso de distribuição complementar. O [r] só ocorre após consoante coronal, ao passo que o [l] ocorre em todos os demais contextos, e nunca após consoante coronal.

O uso dessas tabelas facilita muito a visualização do conjunto de dados, sendo, então, muito útil nesse tipo de análise. Uma observação importante é que esses dados servem de argumento para se considerar que os fones pós-alveolares e palatais são coronais, já que eles funcionam como as coronais anteriores [s d t z] e não como as dorsais e labiais que podem preceder o [l].

4.4. NEUTRALIZAÇÃO

Examinaremos agora dados um pouco mais complexos. Por isso, é bom voltar para o português do Brasil, com que temos mais familiaridade. Examinaremos as consoantes sibilantes, que são quatro: [s z ʃ ʒ].

Examinemos essas consoantes duas a duas. Elas podem ser colocadas num quadrado.

(17)	s	ʃ
	z	ʒ

Da forma como elas estão dispostas, há somente uma diferença entre as duas consoantes que estão na mesma linha ou na mesma coluna. Ou seja, os pares, [s] e [z], [s] e [ʃ], [z] e [ʒ], e [ʃ] e [ʒ], apresentam uma única diferença entre cada membro do par. As mais diferentes são as consoantes [s] e [ʒ], ou as consoantes [ʃ] e [z]. A consoante [s] é uma fricativa alveolar surda, enquanto a consoante [ʒ] é uma fricativa pós-alveolar sonora: elas apresentam duas diferenças. A consoante [ʃ] é uma fricativa pós-alveolar surda, enquanto a consoante [z] é uma fricativa alveolar sonora: elas também apresentam duas diferenças.

Comparemos os fones mais próximos.

Comparando os fones [s] e [ʃ] no português, vemos que eles podem distinguir significado, sendo portanto fonemas distintos. Temos pares mínimos como *lance* e *lanche*, pronunciados respectivamente ['lɛsi] e ['lɛʃi], e os substantivos *soro* e *choro*, que são pronunciados respectivamente ['sɔɾu] e ['ʃɔɾu]. Temos, então, os fonemas /s/ e /ʃ/.

Mas não é isso que acontece em todos os contextos no português do Brasil. Há exemplos em que a troca de [s] por [ʃ] não altera o significado. Palavras como *pasta* podem ser pronunciadas tanto ['paste] quanto ['paʃte], e palavras como *giz* podem ser pronunciadas ['ʒis] ou ['ʒiʃ].

Mas afinal, [s] e [ʃ] são fonemas distintos ou não? Em certos contextos eles distinguem significado mas em outros não distinguem? Na verdade, o que devemos verificar é se os dois fones *podem* distinguir significado *em algum contexto*. Se isso for *possível* na língua em questão, eles são fonemas diferentes nessa língua. Nos contextos em que eles perdem a

capacidade de distinguir significado, dizemos que o contraste entre eles sofreu *neutralização*. O que foi neutralizado aqui nesses exemplos foi o ponto de articulação.

A mesma neutralização de ponto de articulação que encontramos entre as sibilantes surdas existe também entre as sibilantes sonoras [z] e [ʒ]. Elas funcionam como fonemas no português, o que pode ser confirmado através de pares mínimos como *gelo* e *zelo* pronunciados, respectivamente, ['ʒelu] e ['zelu], se forem substantivos, ou ['ʒɛlu] e ['zɛlu] se forem formas verbais. Também há distinção de significado entre *(h)aja* e *asa*, pronunciados ['aʒɐ] e ['aze].

Novamente constatamos que o contraste só se verifica diante de vogais em português. Diante de consoantes encontramos a mesma situação de variação livre vista com o [s] e o [ʃ]. Palavras como *vesgo* e *asma* são pronunciadas indiferentemente ['vezgu] ou ['veʒgu], e ['azme] ou ['aʒme].

Portanto, o contraste de ponto de articulação nas sibilantes do português brasileiro apresenta neutralização tendo como resultado variação livre.

Há ainda outro tipo de neutralização entre pares de sibilantes. Em pares como *face* ['fasi] e *fase* ['fazi], ou *peço* ['pɛsu] e *peso* ['pɛzu], vemos o contraste fonológico ou fonêmico entre /s/ e /z/, a segunda vozeada e a primeira, não. Como nos dois casos anteriores, esse contraste só se verifica diante de vogal. Diante de consoante, como em *casca* e *vesga*, pronunciadas ['kaskɐ] e ['vezgɐ], nas variedades sem posteriorização da sibilante, vemos que a sibilante e a consoante seguinte precisam ter o mesmo vozeamento: ou são ambas surdas, como em ['kaskɐ], ou são ambas sonoras, como em ['vezgɐ].

Da mesma forma, em *chá* ['ʃa] e *já* ['ʒa], as consoantes [ʃ] e [ʒ] distinguem significado e podemos, então falar do contraste fonêmico entre /ʃ/ e /ʒ/. Os mesmos exemplos *casca* e *vesga*, pronunciadas ['kaʃkɐ] e ['veʒgɐ], nas variedades com posteriorização da sibilante, ilustram agora a neutralização entre esses dois fonemas, mas diferentemente do que ocorre com o contraste de ponto de articulação, novamente com concordância das duas consoantes quanto ao vozeamento. Logo, vemos que o resultado da neutralização do vozeamento das sibilantes é uma distribuição complementar.

Retomando, então, vemos que o resultado da neutralização pode ser variação livre, como no caso de [s] ou [ʃ] diante de consoante não vozeada, assim como no caso de [z] e [ʒ] diante de consoante vozeada, mas também pode ser distribuição complementar, como no caso de [s] e [z] ou [ʃ] e [ʒ] diante de consoante. .

Com relação às vogais, um dos principais contextos em que encontramos neutralização é nas sílabas átonas. O primeiro exemplo a dar dessa situação é a do português, que tem sete vogais orais distintas nas sílabas tônicas, /i e e a ɔ o u/, como vimos no capítulo 3. Mas fora da sílaba tônica não há essas sete possibilidades distintas. Em palavras como *recado*, a vogal

grafada <e> varia entre as pronúncias [e] e [ɛ], e em palavras como *coração*, a vogal grafada <o> entre as pronúncias [o] e [ɔ], o que em grande parte é uma diferença regional.

Examinemos agora os dados do catalão central (Espanha), falado numa região que inclui Barcelona. Primeiramente, vejamos exemplos de vogais em monossílabos.

- | | | | |
|------|---------------------|---------------------|-------------------|
| (18) | deu ['dew] 'deus' | deu ['dɛw] 'deve' | dau ['daw] 'dado' |
| | diu ['diw] 'diz' | duu ['duw] 'conduz' | |
| | molt ['moɫ] 'muito' | mòlt ['mɔɫ] 'moído' | mul ['muɫ] 'mulo' |

Os cinco primeiros exemplos demonstram que /e/, /ɛ/, /a/, /i/ e /u/ distinguem significado, sendo, portanto, fonemas distintos. Os três últimos exemplos complementam o quadro, demonstrando que /o/, /ɔ/ e /u/ também são fonemas distintos. O quadro completo de vogais nesses monossílabos, todos tônicos, é o mesmo quadro das vogais orais do português: /i e ɛ a ɔ o u/.

Observemos agora o que ocorre em sílabas átonas. Examinaremos o que acontece quando um sílaba que é tônica numa forma deixa de ser tônica em outra forma relacionada. Vejamos a primeira para exemplificar.

- | | | |
|------|---------------------|-----------------------------|
| (19) | baix ['baɪ] 'baixo' | baixesa [bə'fɛzə] 'baixeza' |
|------|---------------------|-----------------------------|

A vogal [a] de ['baɪ] não se mantém como [a] quando se torna átona. Podemos ver isso examinando a palavra *baixesa*, pronunciada [bə'fɛzə]. Na sílaba inicial, em vez de um [a], temos agora um *schwa* [ə]. Podemos dizer que o [a] se transforma em [ə] em sílaba átona. Examinemos agora as demais vogais. Como no primeiro par de exemplos, a forma da esquerda ilustra as vogais em sílaba acentuada e as da direita como essas vogais são pronunciadas quando em sílabas não acentuadas. Examine atentamente os exemplos antes de começar a ler os parágrafos seguintes.

- | | | |
|------|--------------------------|-----------------------------------|
| (20) | lletra ['ɫɛtrə] 'letra' | lletrat [ɫə'trat] 'letrado' |
| | peix ['peɪ] 'peixe' | peixet [pə'fɛt] 'peixinho' |
| | cola ['kɔlə] 'cola' | colar [ku'la] 'colar' (verbo) |
| | passió [pəsi'o] 'paixão' | passional [pəsiu'naɫ] 'passional' |
| | dur ['du] 'duro' | duresa [du'rezə] 'dureza' |
| | mig ['mitʃ] 'meio' | mitjà [mi'dʒa] 'médio, mediano' |

Pelos exemplos, podemos verificar que não só o [a], mas também o [e] e o [ɛ] passam a ser pronunciados como [ə] quando em sílaba átona. A possibilidade de contraste entre essas três vogais, que existia na sílaba tônica, desaparece em sílaba átona. Trata-se, portanto, de mais um caso de neutralização.

Também com relação às vogais posteriores [u], [o] e [ɔ] encontramos algo semelhante. Em sílaba tônica é possível o contraste entre as três, como vimos com *molt* [moɫ] ‘muito’, *mòlt* [moɫ̃] ‘moído’ e *mul* [muɫ] ‘mulo’. Em sílaba átona, novamente, esse contraste é impossível. Ele é perdido porque todas essas vogais convergem na pronúncia como [u].

A única vogal em catalão que não sofre neutralização é o [i]. Ele se mantém como [i] em sílaba átona, e nenhuma outra vogal passa a [i] nessa posição.

O sistema de neutralizações do catalão pode ser mais bem visualizado com um diagrama que indica as mudanças sofridas pelas vogais em sílabas átonas. Represento esse processo com setas no trapézio das vogais.

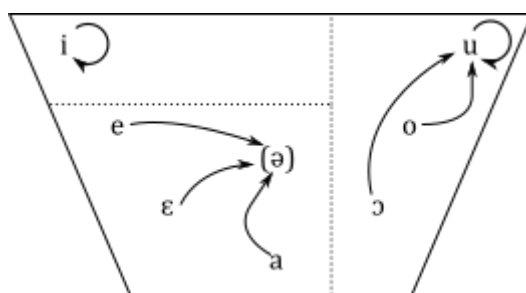


Figura 26. Neutralizações vocálicas no catalão central.

Nessa representação gráfica que proponho, uso as setas para indicar a direção da transformação das vogais em sílabas átonas. O [i] se mantém como era, então tem uma seta que parte dele e volta para ele mesmo. O mesmo ocorre com o [u]. No entanto, a única seta que aponta para o [i] parte dele mesmo, portanto, não há neutralização envolvendo essa vogal. Por outro lado, três setas apontam para o [u]: a que parte do próprio [u], a que parte do [o] e a que parte do [ɔ]. Essa representação gráfica indica que essas três vogais se neutralizam, e que o resultado é o [u], uma das três. Há também três setas que apontam para o [ə]: a que parte do [a], a que parte do [e] e a que parte do [ɛ]. A situação é semelhante à do [u], no qual convergem três setas, mas uma das setas parte do próprio [u], e nenhuma das três setas que apontam para o [ə] parte dele. Isso porque ele não ocorre em sílaba tônica. Os parênteses em volta dele foram usados para ressaltar esse fato: não há um fonema /ə/ no catalão central; ele é apenas o resultado da neutralização desses 3 fonemas vocálicos em sílaba átona.

Contrastando a neutralização de sibilantes do português especificamente com a neutralização das vogais /a e ε/ em catalão, vemos que no primeiro caso o resultado da neutralização é sempre um dos fonemas que se neutralizam, ao passo que na neutralização dessas três vogais em catalão o resultado é um fone totalmente diferente.

Vejamos agora outro tipo de exemplos. Isso será visto com mais detalhes no capítulo relacionado à estrutura silábica, o capítulo 9, mas existem línguas que distinguem vogais breves de vogais longas, como fazia o latim. Examinemos agora o padrão de neutralização das vogais do palauano (Palau). O acento agudo está indicando a sílaba tônica e o hífen separa a raiz do morfema possessivo.

(21)	forma sem possuidor	com possuidor de 1ª p.sg.	
	kar	kər-úk	‘remédio’
	sérs	sərs-ék	‘jardim’
	bsíbs	bsəbs-ék	‘furadeira’
	ŋór	ŋər-ék	‘boca’
	ʔúr	ʔər-ík	‘risada’

Como podemos observar em sílaba átona todas as vogais se neutralizam, passando a ser pronunciadas como *schwa*. Na primeira coluna vemos as vogais /a ε o i u/ na sílaba tônica. Quando a sílaba em que essas vogais se encontram deixa de ser tônica, todas elas passam a ser pronunciadas como [ə]. Esse é um caso extremo de neutralização, em que todas as vogais passam a ter a mesma pronúncia, o *schwa*, que se localiza bem no centro do espaço acústico vocálico.

Ao observarmos o que ocorre com as vogais longas ou ditongos percebemos que o padrão é o outro. Os ditongos contam como vogais longas. Todos eles têm uma das vogais, menos o /a/, seguido de uma semivogal, que é /j/ com as vogais anteriores /i/ e /ε/, e /w/ com as vogais posteriores /u/ e /o/. Nenhuma dessas quatro vogais tem sua qualidade alterada. Apenas a semivogal deixa de ser pronunciada, ou seja, de duas posições (semi)vocálicas passamos a ter somente uma.

(22)	forma simples	forma sufixada	
	rék ^h	rək-él	‘farfalhar (substantivo)’
	ðəkówl	ðəkol-él	‘cigarro’
	ʔíjs	ʔis-él	‘fuga’
	búwʔə	buʔ-él	‘noz de areca’

Um conceito utilizado no estruturalismo para se referir ao resultado da neutralização é o arquifonema. Se os fonemas /z/ e /s/ se neutralizam em determinados contextos, o

resultado é um arquifonema, e pode ser representado com uma letra maiúscula, como /S/. A tendência é utilizar o membro menos marcado da oposição em maiúscula (sobre marcação, ver a seção **Erro! Fonte de referência não encontrada.**). Para representar a pronúncia indiferentemente de vogais médias altas ou médias baixas se pode usar vogais maiúsculas. Um /E/ representa [ɛ e] e um /O/ representa [ɔ o]. O arquifonema apresenta paralelos com a subespecificação, que será vista na seção **Erro! Fonte de referência não encontrada.**

4.4.1 NEUTRALIZAÇÃO ATIVA E PASSIVA

Podemos fazer uma subdivisão entre tipos de neutralização. A primeira distingue a neutralização ativa da passiva. A neutralização ativa, que faz parte do lado dinâmico da fonologia, é o tipo de exemplo que temos explorado mais, com algum som se alterando em algum contexto.

Como exemplo de neutralização ativa em final de palavra podemos citar o finlandês. Nele há as nasais /m/ e /n/ em começo e meio de palavra, mas como só podem ocorrer consoantes coronais ou a parada glotal em final de palavra, onde há uma nasal labial em outras posições, ela se transforma em coronal no final da palavra. Vejamos um exemplo:

- (23) puhelin 'telefone' (nominativo)
 puhelimen 'telefone' (genitivo)

No meio da palavra pode ocorrer o [m], como no genitivo *puhelimen*, mas no final da palavra a nasal labial se transforma em coronal, ocorrendo neutralização entre as duas.

A neutralização passiva pode ser considerada uma característica da fonologia estática das línguas, não sendo o resultado de processos sincrônicos que alteram fones em determinados contextos, como as vogais em sílabas átonas vistas neste capítulo ou as nasais do finlandês.

Por exemplo, em português, temos o contraste entre o tepe e a vibrante em posição intervocálica. Esse contraste, no entanto, só existe nessa posição. Em posição inicial de palavra, os falantes nativos não descendentes de imigrantes utilizam apenas a vibrante e suas variantes. Já em ataques complexos, só se utiliza o tepe. Não há nesses dois exemplos transformações de vibrante em tepe em ataques complexos, nem de tepe em vibrante no início da palavra. Trata-se apenas de inexistência nessas posições do contraste que existe em outra(s).

4.5. SISTEMAS SONOROS

Uma mudança importante ocorrida com o surgimento do estruturalismo foi o desenvolvimento de uma perspectiva que passou a examinar claramente sistemas sonoros, deixando o exame de sons isolados como um aspecto não central na fonologia. O foco passou a ser a análise das relações entre os elementos do sistema ou estrutura, termos que podem ser considerados sinônimos.

Embora Saussure tenha utilizado o termo sistema, o termo que prevaleceu depois, com o mesmo significado, foi estrutura. Daí se origina o nome da corrente teórica *estruturalismo*. Como afirmou Saussure, o valor dos elementos lingüísticos depende dos demais elementos no sistema (*un système où tout se tient*). Um conceito importante proposto por Saussure foi o de valor, segundo o qual o um elemento de um sistema se define pelas relações que tem com os demais elementos do sistema.

Podemos ilustrar isso com um som que é pronúncia de maneira equivalente em cinco línguas: o som [p]. Para caracterizá-lo de uma forma que englobe todas as línguas a seguir, partamos da classificação que o denomina uma oclusiva bilabial. Isso ele é em qualquer língua em que ocorra. Mas faltam detalhes que vão variar, dependendo do sistema da língua em análise.

No português, as oclusivas bilabiais são duas: o [p] e o [b], presentes em *roupa* e *rouba*. Para conseguir delimitar um [p], precisamos dizer que ele é uma oclusiva bilabial surda (desvozeada) e não sonora (vozeada), como seria o [b].

Se examinarmos os dados do mandarim, por outro lado, veremos que há o [p], mas não existe [b]. Por outro lado, existe [p^h]. Descrever o [p] como oclusiva bilabial surda não acrescenta nada, já que as duas oclusivas bilabiais são surdas. O que é relevante é que, além de ser uma oclusiva bilabial, o [p] não é aspirado, ao passo que o [p^h] é. Assim, o valor do [p] em mandarim fica sendo uma oclusiva bilabial não aspirada.

O tailandês, por sua vez, apresenta tanto [p], quanto [b], e também [p^h]. Sendo assim, para delimitarmos o [p] entre as oclusivas bilabiais, precisamos dizer que ele é surdo e não aspirado. As duas características são imprescindíveis.

Uma outra língua que tem [p] e [b], e na qual a caracterização do [p] seria a mesma do tailandês é o híndi, que tem, na verdade: [p], [p^h], [b] e [b^h]. Se a caracterização do [p] nas duas línguas é a mesma, a caracterização do [b] já não é. Em tailandês basta dizer que o [b] é oclusiva bilabial sonora. Em híndi, entretanto, como existem [b] e [b^h], é necessário descrever o [b] como oclusiva bilabial sonora não aspirada (a rigor, não murmurada).

Por fim, numa língua como várias da família ianomâmi, em que o [p] é a única oclusiva bilabial, basta isso para descrevê-lo.

Portanto, para caracterizarmos adequadamente os fones de uma língua e o papel que eles têm na língua é necessário ter um quadro da rede de oposições dessa língua. Deixa-se de lado uma visão atomística para adotar uma visão sistêmica da língua e de sua fonologia.

4.6. SIMETRIA

Esse olhar que examina as características da língua como um sistema é bem ilustrado pelo trabalho de Trubetzkoy, o primeiro autor a examinar um número considerável de sistemas vocálicos para chegar a um levantamento de suas propriedades e classificar esses sistemas. Uma propriedade muito frequente dos sistemas sonoros é a sua simetria. Ela não é algo absoluto, mas uma tendência relativamente forte. Sistemas vocálicos com 3 vogais como o do árabe, o do quêchua e o do inuktitut, com /i a u/, de 5 vogais como os do hebraico, o do havaiano e o do espanhol, com /i e a o u/, e de 7 vogais, como os do português, do catalão e de diversas línguas bantas, com o sistema /i e ε a ɔ o u/, têm o mesmo número de vogais anteriores e posteriores, além de uma vogal central. Há outros tipos de sistemas vocálicos simétricos com configurações distintas dessas. A língua chamorro (Ilhas Marianas) apresenta seis vogais: /i e æ u o α/. Vemos que três delas são posteriores, as vogais /u o α/, e três são anteriores, as vogais /i e æ/. Não há vogais centrais. O amárico (Etiópia) apresenta sete vogais, mas com uma configuração diferente da existente no português. São elas: /i e i̯ ə a u o/. Vemos que são duas vogais anteriores, /i e/, três centrais /i̯ ə a/ e duas posteriores /u o/.

Embora seja comum, a simetria não é um universal absoluto. Um bom contraexemplo é o francês, que tem entre as vogais orais três posteriores, /ɔ o u/, mas seis vogais orais anteriores: /i e ε y ø œ/, além da vogal central /a/.

Embora seja muito comum analisar a simetria de sistemas vocálicos, podemos analisar a simetria também de sistemas consonantais. O latim apresentava contraste de vozeamento nos três pontos de articulação das oclusivas, sendo assim, o sistema de oclusivas era simétrico. Seguem exemplos de alguns pares mínimos baseados apenas no contraste de vozeamento:

(24)	bellō ‘guerrear’	pellō ‘impulsionar’
	densus ‘espesso’	tensus ‘estendido’
	gustōs ‘sabores’ (acus. pl.)	custōs ‘guarda’ (nom. sing.)

No entanto, o sistema de obstruintes latinas (oclusivas e fricativas) não era simétrico, no sentido de que esse contraste de vozeamento só existia entre as oclusivas. As fricativas do latim eram /f s h/, se se quiser incluir o /h/ entre as fricativas. É possível considerar que o

surgimento das fricativas sonoras nas línguas românicas tenha sido reflexo da tendência de simetria também nos sistemas consonantais.

EXERCÍCIOS

- 1) Qual o estatuto fonológico dos fones [p b t d k g r s fʃ] na língua sanõma da família ianomâmi (Roraima, Venezuela)? Obs: estatuto fonológico se refere ao fato de eles serem fonemas ou alofones, e se forem alofones, de qual tipo.

[ˈboa] ‘faca’	[koaˈraɡa] ‘fogo’
[koaˈtaɡa] ‘fogo’	[mãˈdu] ‘água, rio’
[ãñõkoiˈmẽ] ‘lagarta’	[ˈfʃuiði] ‘mulher’
[koaˈdaɡa] ‘fogo’	[ˈpoa] ‘faca’
[mãˈtu] ‘água, rio’	[ãñõgoiˈmẽ] ‘lagarta’
[mãˈru] ‘água, rio’	[ˈsuiði] ‘mulher’

- 2) Em várias línguas originárias do Brasil ocorrem consoantes orais pré-nasalizadas (por ex., [mb, nd]) e nasais pré ou pós-oralizadas (bm, mb). Em wayoro êmêto (Rondônia), ocorrem tanto consoantes nasais simples quanto pós-oralizadas. Examine qual o estatuto fonológico dos pares [m, mb], [n, nd], [ŋ, ŋg] e [ŋʷ, ŋgʷ]. Nos exercícios daqui para a frente, considere a ausência de dados de um determinado tipo como evidência de que eles não ocorrem na língua em questão.

[mãˈtʰi] ‘papagaio’	[ˈnõ] ‘outro’
[mbiˈkʰip] ‘fonte de água’	[ŋʷẽˈndap] ‘tamanduá’
[õˈmõ] ‘beiju’	[aˈnĩŋ] ‘verme’
[ˈtĩm] ‘neto’	[ndoːˈja] ‘calango’
[ˈŋgʷɛj] ‘paxiúba’	[ˈŋgɛ] ‘roça’
[ˈŋẽm] ‘seio’	[õˈmẽn] ‘meu marido’
[ˈŋʷãm] ‘duro’	[ãˈŋã] ‘largo’
[õŋʷãˈẽ] ‘minha panela’	[ˈjĩːŋ] ‘fumaça’
[õˈŋgi] ‘meu cuspe’	[ẽˈŋgʷit] ‘mel’
[aĩˈmbɛ] ‘pó da árvore de angico’	

- 3) Qual o estatuto fonológico dos fones [k c q] e [ʃ s] em shanenawa (Acre)? Obs: [tʃ] não se comporta do mesmo modo que [ʃ], e não faz parte do que deve ser analisado no exercício.

[pahĩ'ci] 'orelha'	[iʃ'cĩ] 'peixe'
[qus'qoʔ] 'urubu'	[qo'saʔ] 'cedro'
[ka'kã] 'abacaxi'	[faki'huʔ] 'menino'
[ʃi'pi] 'banana'	[quʃ'qoʔ] 'sapo'
[pi'siʔ] 'casa'	[ʃa'kaʔ] 'casca'
[ri'ci] 'nariz'	[di'ci] 'nariz'
[ju'ʃiʔ] 'bicho'	[ʃi'naʔ] 'aranha'
[tʃa'soʔ] 'veado'	[wa'siʔ] 'capim'
[fu'suʔ] 'face'	[ʃi'ciʔ] 'milho'
[i'si] 'doença'	[ma'sũʔ] 'veado'
[fus'taʔ] 'nuvem'	[sawʔ] 'osso'
[fu'seʔ] 'lontra'	

- 4) Qual o estatuto fonológico das vogais médias em wayana (Suriname/Guiana Francesa/Pará)? A distribuição delas sofre alguma restrição? Para determinar isso, pense se todas as combinações esperadas são possíveis.

[pom] 'deitar'	[pɔm] 'deitar'
[kopə] 'chuva'	[kɔpə] 'chuva'
[muɾej] fruta (esp.)	[ʃijew] roedor (esp.)
[muɾej] fruta (esp.)	[ʃijew] roedor (esp.)
[pepta] 'grande'	[pɛpta] 'grande'
[ahnep] 'amendoim'	[ahnɛp] 'amendoim'
[koko] 'noite'	[kɔkɔ] 'noite'
[kinoɾo] 'arara'	[kinɔɾɔ] 'arara'
[wewe] 'madeira'	[wɛwɛ] 'madeira'
[ɛɾɛkit] 'ferida'	[ɛɾɛkit] 'ferida'

- 5) Qual o estatuto fonológico de [β f], [i i u o], [n ŋ] e [s ʃ s] em yawanawá (Acre)? Obs: a africada [tʃ] não faz parte do que deve ser analisado no exercício.

[ʃuriʃ'pi] 'esterno'	[wa'si] 'capim'
----------------------	-----------------

[fɨ'tu] 'reboto'	[tɨf'ki] 'torto'
[pus'tu] 'barriga'	[fɨ'to] 'reboto'
[nis'kan] 'suar'	[βu'ta] 'nuvem'
[tɨf'ki] 'trouxa'	[pos'to] 'barriga'
[suj] 'soprar'	[ma'ki] 'piranha'
[ʃaʃin] 'mucura da água'	[fi'pi] 'macaco soim'
[nawi] 'fumo'	[ki'ja] 'rachado'
[rantuŋ'ku] 'joelho'	[fu] 'assar'
[ki'ʃa] 'lábios'	[pahiŋ'ki] 'orelha'
[aʃ'fa] 'boca'	[aʃ'βa] 'boca'
[ma'ku] 'careca'	[niwi] 'vento'
[si'ni] 'nambu pedrês'	[ʒumu'ʒi] 'agulha'
[tʃa'nas] 'cotiara'	[huʃ'ka] 'dor de cabeça'
[ramaʃta'ma] 'antigamente'	[is'ku] 'japó'
[tis'ki] 'jandaia'	[fa'βa] 'dia'
[fara] 'bom'	[ku'fu] 'cujubim'
[iftin] 'estrela'	[sa'tã] 'caracol'
[kiʃ'ni] 'barba'	[pas'pi] 'ponta de lança'

- 6) Analise os dados a seguir da língua mosetén (Bolívia) e determine o estatuto fonológico das consoantes [h] e [x].

[aha] 'gripe'	[rəx] 'todos'	[həmʔ] 'bom'
[hame] 'flor'	[ax] 'já'	[hoʔ] 'água'
[nõxnõx] 'amanhã'	[mix] 'stone'	[naihokiʔ] 'amanhecer'
[hikox] 'sal'	[maʔhoi] 'feliz'	[ixme] 'flecha'

- 7) Há análises que consideram que o mandarim possui apenas uma vogal média fonêmica, a qual tem, no entanto, quatro alofones. Verifique quais os contextos de ocorrência de cada uma delas: [ə ɤ e o]. Os símbolos à direita, dentro dos colchetes, indicam os tons, que não fazem diferença neste exercício.

[ɤ ˥] 'ganso'	[lje ˥] 'rachar'	[lwo ˥] 'cair'
[ɕe ˥] 'bota'	[ɕjow ˥] 'descansar'	[ən ˥] 'favor'
[ow ˥] 'Europa'	[xwən ˥] 'casamento'	[xej ˥] 'preto'
[ləŋ ˥] 'distráido'	[ɕjen ˥] 'fresco'	[ɕqen ˥] 'anunciar'

[lɤ ʋ] ‘feliz’

[ljow ʋ] ‘seis’

[ʃweɪ ʌ] ‘água’

EXERCÍCIOS DE PALATALIZAÇÃO

A região palatal está envolvida em muitos fenômenos fonológicos. Como na maioria dos aspectos da fonologia, os fenômenos que a envolvem variam de língua para língua. A palatalidade pode se apresentar como a articulação primária (ou única) de uma consoante, mas também como articulação secundária, acrescentada à articulação principal de uma consoante. Os exercícios a seguir englobam dados dos dois tipos.

- 8) Observe os dados a seguir e analise a relação entre fricativas velares e palatais no grego moderno, para determinar o estatuto fonológico delas. Caso haja alofonia, verifique em que contextos cada alofone ocorre.

[ma'çeri]	‘faca’	[ixɔs]	‘som’
[xa'ra]	‘alegria’	[çeri]	‘mão’
[xanɔ]	‘perder’	[xris'tɔs]	‘Cristo’
[xliar'ɔs]	‘morno’	[ɔçi]	‘não’
[xɔ'revɔ]	‘dançar’	[xno'tizɔ]	‘feder’
[xrɔma]	‘cor’	[tsa'ruçi]	‘sapato com pompom na ponta’
[xulja'ra]	‘colher de sopa’	[psi'çi]	‘alma’
[rixɔ]	‘arremessar’	[çerɔmɛ]	‘se alegrar’
[zaxari]	‘açúcar’	[ta'çitita]	‘velocidade’
[paxɔs]	‘gordura’	[maxɔmɛ]	‘lutar’
[paçis]	‘gordo’	[ðixɔ]	‘mostrar’

- 9) Observe os dados a seguir e analise a relação entre fricativas velares e palatais no alemão, para determinar o estatuto fonológico delas. Caso haja alofonia, verifique em que contextos cada alofone ocorre.

[gə'ɪçt]	‘boato’	[hø:çstəns]	‘no máximo’
[zu:xə]	‘procura’	[vax]	‘acordado’
[kʏçə]	‘cozinha’	[niçt]	‘não’
[vɔxə]	‘semana’	[ho:x]	‘alto’
[lœçə]	‘buracos’	[fluxt]	‘fuga’
[ɛçt]	‘verdadeiro’		

Observe agora as duas palavras a seguir. A pronúncia padrão acaba sendo substituída por outras pronúncias, dependendo da região. Abaixo vemos a distribuição aproximada das pronúncias regionais. Por que a pronúncia padrão é substituída por outras? Que características são preservadas em cada pronúncia regional, se comparada à pronúncia padrão?

	‘química’	‘China’
padrão	[çe'mi:]	['çi:na]
centro e norte da Alemanha	[je'mi:]	['fi:na]
sul da Alemanha, Áustria	[ke'mi:]	['ki:na]
Suíça	[xe'mi:]	['xi:na]

- 10) Observe os dados a seguir e analise a relação entre consoantes palatalizadas e não palatalizadas no irlandês, para determinar o estatuto fonológico delas. Caso haja alofonia, verifique em que contextos cada alofone ocorre. A rigor, as consoantes não palatalizadas são velarizadas, e poderiam ser transcritas com o diacrítico [ʷ], como [bʷ], que representa uma oclusiva labial sonora velarizada, mas para deixar a transcrição mais fácil de visualizar, omito esse diacrítico.

['bɑ:d]	‘barco’	['mʲɑ:l]	‘bola’
['fʲo:lʲ]	‘carne’	['kɑ:s]	‘caso’
['gʲal]	‘claro’	['ni:]	‘criança’
['fuər]	‘frio’	['mɑ:l]	‘lento’
['fo:lʲ]	‘ligeiro’	['kad]	‘o quê?’
['fuərʲ]	‘obteve’	['kʲad]	‘permissão’
['nʲi:]	‘filha’	['kɑ:sʲ]	‘casos’
['bo:]	‘vaca’	['bɑ:dʲ]	‘barcos’
['gal]	‘vapor’	['bʲo:]	‘vida’

- 11) Observe os dados a seguir e analise a relação entre consoantes palatalizadas e não palatalizadas no tchuvache (Rússia), para determinar o estatuto fonológico delas. Caso haja alofonia, verifique em que contextos cada alofone ocorre.

[pʲirʲ]	‘lona’	[ʂuv]	‘água’
[tʲirʲ]	‘pele, couro’	[pʷl]	‘mel’
[pʲilʲ]	‘bênção’	[vʲitʲ]	‘cobrir’

[tʲirʲ]	‘há pouco’	[numaj]	‘muito(s)’
[kʲilʲ]	‘vir’	[pʲitʲ]	‘rosto’
[pʲilʲ]	‘saber’	[xəvat]	‘força’
[tut]	‘segurar’	[man]	‘meu’
[kʲimʲegʲe]	‘livro’	[tʲynʲ]	‘cair’
[ɕun]	‘pessoa’	[xəvɐrt]	‘rápido’
[sən]	‘rosto’	[pʲyrʲe]	‘rim’
[tavax]	‘obrigado’	[tus]	‘amigo’
[pʲyrʲnʲe]	‘dedo’	[sʲimʲ]	‘escuridão’
[sʲimʲy]	‘sentença, oração’		

EXERCÍCIOS DE NEUTRALIZAÇÃO

12) Faça um levantamento das vogais em búlgaro em sílaba tônica e em sílaba átona e analise sua neutralização. O padrão dos dados é encontrado nos dialetos orientais do búlgaro. Qual o inventário vocálico em sílaba tônica? Qual o inventário vocálico em sílaba átona? Quais conjuntos de vogais se neutralizam? Qual tipo de contraste é perdido?

- | | | | |
|------------|------------|-------------|---------------|
| a. [rɔguf] | ‘córneo’ | [ru'gat] | ‘chifrudo’ |
| b. ['mig] | ‘instante’ | [mignu'ven] | ‘instantâneo’ |
| c. ['selu] | ‘vila’ | [si'la] | ‘vilas’ |
| d. ['dux] | ‘espírito’ | [du'xɔvin] | ‘espiritual’ |
| e. ['grat] | ‘cidade’ | [grə'dɛts] | ‘cidadezinha’ |
| f. ['mɐf] | ‘homem’ | [mɐ'ʒɛ] | ‘homens’ |

13) Faça um levantamento das vogais em belarusso e analise sua neutralização. Os dados completos envolveriam mais fatores. Esses dados aqui apresentados têm um funcionamento mais claro. Qual o inventário vocálico em sílaba tônica? Qual o inventário vocálico em sílaba átona? Quais conjuntos de vogais se neutralizam? Qual tipo de contraste é perdido? Obs: atenção para não confundir as vogais [i] e [ɨ].

Primeiro, alguns monossílabos para ilustrar as vogais existentes em sílaba tônica.

- | | | | |
|----------|-----------|----------|------------|
| ['lʲist] | ‘folha’ | ['ɣɛtʲɨ] | ‘este’ |
| ['mɨ] | ‘rato’ | ['kat] | ‘carrasco’ |
| ['fʲum] | ‘barulho’ | ['kɔt] | ‘gato’ |

Agora algumas alternâncias na posição do acento para verificarmos se há mudanças correlacionadas.

[ˈstɔɫ] ‘mesa’	[staˈɦ] ‘mesas’
[ˈrɛki] ‘rios’	[raˈka] ‘rio’
[ˈstarasʲtsʲ] ‘envelhecer’	[staˈrʲ] ‘velho’
[ʃuˈmjɛtsʲ] ‘barulho’	[ˈʃum] ‘barulho’
[ˈlʲist] ‘folha’	[lʲistaˈpat] ‘novembro’
[ˈbʲɪ] ‘ele era’	[ˈbʲɪˈla] ‘ela era’

- 14) Faça um levantamento das vogais em russo e analise sua neutralização. Qual o inventário vocálico em sílaba tônica? Qual o inventário vocálico em sílaba átona? Quais conjuntos de vogais se neutralizam? Qual tipo de contraste é perdido? Não incluí a vogal [ɪ] no exercício porque seu estatuto é controverso.

Primeiro, alguns monossílabos para ilustrar as vogais existentes em sílaba tônica.

[ˈdom] ‘casa’	[ˈsam] ‘próprio (masc.)’
[ˈmʲɛx] ‘pele’	[ˈdʲisk] ‘disco’
[ˈʃum] ‘barulho’	

Agora algumas alternâncias na posição do acento e mudanças correlacionadas. Aqui verifique quais as neutralizações.

[ˈdom] ‘casa’	[dɛˈma] ‘casas’
[ˈsam] ‘próprio (masc.)’	[sɛˈmo] ‘próprio (neutro)’
[ˈmʲɛx] ‘pele’	[mʲɪˈxa] ‘peles’
[ˈpʲitʲ] ‘beber’	[pʲɪtʲɪˈvoj] ‘potável’
[ˈʃum] ‘barulho’	[ʃuˈmʲɛtʲ] ‘fazer barulho’

Para finalizar, distinga onde ocorre [ɐ] e onde ocorre [ə]. Esses dados podem ser revistos após ter sido lido o capítulo 11, que trata do acento.

[ˈdom] ‘casa’	[dɛˈma] ‘casas’	[dɐmɛˈvoj] ‘doméstico’
[ˈgolɐvu] ‘cabeça’ (AC)	[gɛˈlofkə] ‘cabecinha’	[gɐlɛˈva] ‘cabeça’ (NOM)

[ˈkamˈinʲ] ‘pedra’ [kəmˈiˈnʲej] ‘de pedras’ [kəmˈiˈnʲistɨj] ‘rochoso’
 [ˈdalʲiji] ‘mais longe’ [dɐˈlʲokʲij] ‘distante’ [dəlʲiˈkoj] ‘distante’

Obs: AC = acusativo; NOM = nominativo.

REFERÊNCIAS E LEITURAS COMPLEMENTARES

Cândido, Gláucia Vieira (1998). *Aspectos fonológicos da língua shanenawá (pano)*.

UNICAMP: Dissertação de Mestrado.

Cristófaró Silva, Thaïs; Cantoni, Maria (2011). On the emergence of contrast.

Dum-Tragut. Armenian.

Hayes

Nogueira, Antonia Fernanda de Souza (2011). *Wayoro êmêto: fonologia segmental e morfossintaxe verbal*. USP: Dissertação de Mestrado.

Mattoso Câmara Jr., Joaquim (1981). Dicionário de lingüística e gramática: referente à língua portuguesa. 10ª ed. Vozes: Petrópolis.

Odden

Paula, Aldir Santos de (2004). *A língua dos índios yananawá do Acre*. UNICAMP: Tese de Doutorado.

Silva, Adelaide

Tavares, Petronila da Silva (2005). *A Grammar of Wayana*. Rice University: Tese de Doutorado.

Trubetzkoy, Nikolai (1939). *Grundzüge der Phonologie*. Praga: Travaux du cercle linguistique de Prague.

Lass, Roger (1984). *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press. 125-168.

CAPÍTULO 5

FONÉTICA

“Phonetics can provide some of the explanations for the sound patterns discovered by phonologists.

But the benefit is symmetrical: phonology, by studying the behavior of speech sounds in language – sound changes, patterns of sound sequences in words and morphemes, allophonic variation, the structure of phoneme inventories – can help phonetics to focus on those articulatory, acoustic, and perceptual factors that principally serve the function of communication.”

(John Ohala (2005). The marriage of Phonetics and Phonology. In: *Journal of Acoustic Science and Technology* 26, 2, pp. 418-422)

Este capítulo trata das propriedades sonoras de consoantes e vogais, ou seja, não da perspectiva da fonética articulatória, como vimos nos capítulos 2 e 3, que examina que órgãos produzem os sons, mas da perspectiva da fonética acústica.

Trataremos inicialmente com algum detalhe da fonética articulatória. Depois veremos a fonética acústica e, ao final, algumas observações sobre a fonética auditiva, que tradicionalmente recebe muito menos atenção do que as outras, mas que é determinante em muitos processos e mudanças fonológicos. Nesse sentido, já foi apontado por autores como Boersma que a fonologia tem um viés indevidamente acentuado para a produção. É fundamental que se inclua também a questão da percepção. Boersma, por exemplo, fala de uma gramática da percepção, distinta da gramática da produção. É evidente que, se não se tem uma compreensão da percepção, nosso entendimento da fonologia fica incompleto.

5.1 FONÉTICA E FONOLOGIA

Não é óbvio que seja necessário haver duas disciplinas distintas que tratam dos aspectos sonoros da língua. Porém, desde Trubetzkoy (1939), passou a ser tradicional fazer uma distinção entre duas disciplinas: a fonética e a fonologia. Pode-se considerar que as duas tratam dos mesmos fenômenos, mas com focos distintos.

A primeira, a fonética, se aproxima das ciências biológicas e das exatas, levando em consideração, por um lado, questões como o processo de respiração e suas diferenças quando estamos falando ou não, o que ocorre com a corrente de ar quando estamos produzindo determinado som, e, por outro lado, quais frequências em hertz são encontradas nos sons produzidos, a duração desses sons em milissegundos, e assim por diante. A segunda, a fonologia, se ocupa de questões diferentes, essencialmente com o aspecto gramatical dos sons linguísticos, envolvendo questões tais como a distintividade dos sons, em que posições da palavra eles podem ocorrer, etc.

Apesar de serem disciplinas correlatas, desde Trubetzkoy as duas seguiram caminhos opostos e durante décadas tenderam a olhar com desconfiança o que era feito na outra disciplina. Durante boa parte do século XX, as duas se diferenciaram cada vez mais, chegando a um ponto em que pesquisadores de uma das duas áreas chegavam simplesmente a ignorar o que era feito na outra. Para retratar a visão que os que trabalham numa disciplina tinham da outra, podemos citar o trecho de Trubetzkoy (1939), em que ele afirma que a fonologia está para a economia assim como a fonética está para a numismática.

De acordo com Beckman e Kingston (1990), esse ponto de vista e o consequente desenvolvimento da fonologia alheio ao desenvolvimento da fonética geraram um preconceito correspondente nos foneticistas. Esses autores citam uma afirmação contrária, mais ferina ainda, segundo a qual a fonética está para a astronomia assim como a fonologia está para a astrologia. Os autores não citam quem fez essa afirmação. Deixo claro aqui que não subscrevo nenhuma dessas duas afirmações. Mas elas retratam muito bem o clima existente e talvez predominante pelo menos entre parte dos pesquisadores de ambas as áreas principalmente até o final da década de 1980. As rixas entre as duas áreas podem ser comparadas às que existem entre arquitetos e engenheiros, ou entre psicólogos e psiquiatras. No meu modo de ver, fonética e fonologia são fundamentais em si e para refinar as pesquisas feitas na área complementar.

Por questão de espaço, não é possível neste capítulo falar da fonética em termos extremamente aprofundados, mas pretendo apresentar nele concisamente suas bases e em capítulos posteriores analisar fenômenos fonológicos considerando os aspectos fonéticos. Em consonância com a epígrafe de Ohala no início deste capítulo, meu intuito é ilustrar como a fonética e a fonologia se beneficiam das pesquisas realizada na área complementar.

5.2 PRODUÇÃO E PERCEPÇÃO

É comum em manuais de fonologia que incluem um capítulo sobre fonética que ele seja essencialmente uma parte introdutória pela qual se tem que passar (às vezes quase a contragosto) para chegar ao que supostamente interessa, que é a fonologia propriamente dita. Não é esse o caso neste livro. Pretendo apresentar os principais conceitos da fonética e retomá-los nas análises, buscando também determinar quais suas consequências para os padrões encontrados na fonologia, ou seja, tentando determinar em que medida os fatos fonológicos têm explicações plausíveis na fonética.

Frequentemente em fonologia se busca determinar qual a representação mental de determinados segmentos ou determinada estrutura fonológica, considerando que isso por si só explicará os fenômenos fonológicos. Nem tudo na fonologia é determinado diretamente pela fonética, mas se já de início a excluirmos como uma fonte de explicações possíveis, correremos o risco de nos privar de parte considerável da explicação do que ocorre em fonologia. Se houver uma explicação clara em termos fonéticos para determinada característica da fonologia, uma explicação com base em algum recurso puramente formal ou se torna desnecessária ou apenas reformula o que a fonética faz. É possível ainda manter a ênfase nos aspectos formais, mas considero fundamental nos perguntarmos a que correspondem as restrições formais foneticamente e, se se preferir, cognitivamente, e não mantê-las num vácuo, como meras representações gráficas que se supõe terem relação com o que se passa em nossa mente. Considerar que traços ou restrições são objetos meramente formais diminui nossa capacidade de explicar os padrões relacionados a eles.

5.3 A PRODUÇÃO DOS SONS DA FALA

Já vimos nos capítulos 2 e 3 informações básicas sobre a produção dos sons da fala, mas neste capítulo examinamos em mais detalhe o que ocorre no nosso trato vocal para que consigamos produzi-los. Uma condição essencial para que isso aconteça é que haja uma corrente de ar. É o ar em movimento que vai possibilitar a produção dos sons da fala.

A situação esmagadoramente mais comum é aquela em que a corrente de ar tem sua origem nos pulmões e se dirige para o exterior do nosso corpo. Essa situação é a mais comum tanto entre as línguas como em cada língua. Para começar, esse é o único tipo de corrente de ar que toda língua usa. Em todas elas há corrente pulmonar, que em termos distintivos é sempre egressiva, de dentro para fora. Apenas se usa a corrente pulmonar ingressiva em contextos extremamente específicos em algumas línguas, como, por exemplo, no final de enunciados, quanto justamente os pulmões já se encontram com volume reduzido de ar. Em cada língua também, mesmo nas que utilizam com frequência outros tipos de corrente, a corrente mais comum é sempre a pulmonar.

Só a passagem do ar pelo trato vocal, no entanto, não é suficiente para produzir som, como podemos verificar se respirarmos normalmente, mas de boca aberta. Para ser produzido um som, é necessário que haja vibração em algum ponto, ocasionada pela existência de algum obstáculo.

O primeiro obstáculo que o ar que vem dos pulmões pode encontrar são as *pregas vocais*, mais antigamente chamadas *cordas vocais*. Como já vimos na seção 2.2, elas se localizam na região da laringe, ou pomo de adão, ou gogó, uma estrutura composta de algumas cartilagens, que é mais visível nos homens.

Quando o ar passa pelas pregas vocais, elas podem estar em configurações variadas, que produzem resultados sonoros distintos. Para o português, são primordiais duas configurações. A primeira é a encontrada na produção dos sons surdos, desvozeados ou não vozeados. Na produção desses sons, como vimos na seção 2.2, as pregas vocais se encontram afastadas (abduzidas), e o ar passa por elas sem produzir som audível. É justamente a manutenção do afastamento das pregas vocais que impossibilita que se gere som na passagem do ar por elas.

Na produção dos sons sonoros ou vozeados, por outro lado, as pregas vocais vibram. Como se dá essa vibração? Ela ocorre em ciclos. As pregas vocais são aduzidas, aproximadas, fechando, interrompendo por completo brevemente a passagem do ar. Se a corrente de ar está sendo impelida dos pulmões para fora, ao encontrar as pregas vocais fechadas, ela força a passagem e acaba abrindo a glote (a abertura entre as pregas vocais). Esse ciclo se repete dezenas ou centenas de vezes por segundo na produção dos sons sonoros.

Além de a corrente de ar usada na produção da fala poder ser iniciada nos pulmões, há duas outras possibilidades encontradas nas línguas em geral, como vimos na seção 2.6: uma relativamente comum, embora não tanto quanto a corrente pulmonar, e outra muito rara. O tipo relativamente comum de corrente de ar não pulmonar é a corrente glotal.

Quando tratamos da corrente pulmonar, vimos que os sons linguísticos produzidos com esse tipo de corrente de ar são unicamente egressivos, com o ar saindo do nosso organismo. A corrente glotal é a única das três que apresenta as duas possibilidades de direção do deslocamento do ar, podendo ser tanto ingressiva quanto egressiva. Em ambas as situações, a corrente glotal só é possível com o fechamento momentâneo da glote (embora a rigor, nem sempre isso se aplique às implosivas, como veremos mais adiante nesta seção). Quando isso acontece, a laringe pode ser elevada ou abaixada. Se ela for elevada, obviamente o ar será expelido para fora do nosso organismo e teremos uma corrente egressiva. Se ela for abaixada, por sua vez, teremos uma corrente ingressiva.

As consoantes produzidas com corrente glotal egressiva são chamadas ejetivas e foram mencionadas já na seção 2.6.1.1. Elas contêm uma fase de abertura que soa mais brusca do que as plosivas correspondentes. Lembro aqui que o termo plosiva é reservado para as oclusivas de corrente pulmonar. Na transcrição fonética, as ejetivas são representadas por um apóstrofo. Assim, temos consoantes correspondentes a nossas oclusivas surdas só que ejetivas: [p'], [t'] e [k']. Podemos citar os exemplos do aimara vistos no cap. 2:

- | | | |
|-----|---------------------------------|----------------------------|
| (1) | k'ata [k'ata] 'veloz' | p'itata [p'itata] 'tecido' |
| | q'ala [q'ala] 'tudo, terminado' | t'ula [t'ula] 'lenha' |

São encontradas também africadas ejetivas, tais como [tʃ̥], além de fricativas como [s̥], sendo estas mais raras. Seguem alguns exemplos de africadas ejetivas:

- (2) do aimara: [tʃ̥uspa] 'chuspa' (bolsa de coca)
- (3) do georgiano: [ts̥'igni] 'livro'
- (4) do xhosa: [kx̥'ola] 'inscrever'

Alguns exemplos de fricativas ejetivas:

- (5) do tlingit: [ʈ̥u:k̥h] 'peixe' [s̥'i:k̥h] 'urso' [x̥'á:x̥] 'maçã'
- (6) do lakota: [s̥'a] 'geralmente' [χ̥'o.k̥xa] 'tamboreiro'

A outra possibilidade de corrente glotal é aquela em que a laringe é abaixada, fazendo com que o ar entre em nossa cavidade oral, gerando consoantes implosivas. Elas normalmente são sonoras, e o símbolo usado para transcrevê-las se baseia na consoante pulmonar correspondente acrescentando a ele uma curva descendente na parte superior. Assim temos símbolos como para representar respectivamente as oclusivas labial, coronal e dorsal implosivas: [ɓ ɗ ɠ].

- (7) cambodjano (khmer): [ɓa:raŋ] França, francês [ɗap] 'dez'
- (8) síndi: [ɓa:ru] criança [ɗaru] 'fenda' [ɠəu:n] 'vaca'

De acordo com Ashby (2011: 73-74), as etapas na produção das implosivas são:

1. no início, com a passagem para a cavidade nasal fechada, há uma corrente egressiva pulmonar. Forma-se, então, uma oclusão na cavidade oral.
2. ao mesmo tempo, a laringe se abaixa, aumentando o volume da cavidade faringal.
3. esse aumento de volume pode causar uma diminuição da pressão intraoral.
4. ao final, a oclusão da cavidade oral é desfeita e a pressão intraoral e pressão ambiente se igualam.

Essa descrição não corresponde exatamente à definição dada logo antes do início da seção 2.6.1.1. Como aponta Clements (2002: 307), o abaixamento da laringe e outras estratégias usadas para manter o vozeamento não são exclusividade das implosivas. As oclusivas sonoras também apresentam essas características. Na página 308, Clements afirma que nem o fechamento da glote, nem o fluxo ingressivo de ar, nem a pressão do ar negativa (rarefação), nem o abaixamento da laringe são critérios que distinguem claramente as implosivas das oclusivas sonoras. Embora tradicionalmente se pensasse que nas implosivas há fluxo ingressivo de ar e pressão do ar negativa (rarefação), nenhuma das duas necessariamente ocorre. Como há vibração das pregas vocais, também não há fechamento da glote. Clements defende, então, que o que distingue as implosivas das oclusivas é que nas primeiras não há aumento da pressão do ar na cavidade oral. Pode haver diminuição ou não, mas não há aumento.

Há também, como foi mencionado na seção 2.6.2, a corrente velar, com a qual são produzidos os chamados cliques. Retomo os exemplos vistos lá para facilitar a leitura. Lembro que, como os cliques sempre têm uma oclusão velar, eles podem ser transcritos precedidos de um [k]. Seguem os exemplos da seção 2.6.2. Com o clique bilabial, um beijo estalado, temos um exemplo da língua ʼamkoe (Botswana): [ʔoa] ou [kʔoa] ‘dois’. Com o clique dental ou alveolar, o som usado para expressar a negação, grafado *tsk* em inglês, no temos o exemplo zulu *icici* [i:lí:li] ou [i:klí:kli] ‘brinco (substantivo)’. Com o clique lateral, equivalente ao som utilizado para incitar um cavalo, há o verbo *xoxa* em zulu, pronunciado [lló:lla] ou [klló:klla]. Com o clique palatal, em que toda a coroa da língua entra em contato com o céu da boca, temos o nome da língua ʰaba [ʰabá] ou [kʰabá] (Botswana). Por fim, com o clique pós-alveolar, produzido com a língua formando uma concavidade e sendo puxada para baixo, e não para trás como nos outros cliques, temos o *iqanda* ‘ovo’ [i!anda] ou [ik!anda] em xhosa. Esses fones são quase exclusividade do sudoeste e sul da África como segmentos usados linguisticamente, podendo ser extremamente frequentes em algumas línguas da região. Segundo Ladefoged e Traill (1994), mais de 70% das palavras de um dicionário de !xóõ começam com algum tipo de clique. De acordo com Ladefoged e Traill (1994: 34-35), os cliques são produzidos sempre com corrente de ar velar ingressiva. Como eles sempre envolvem o fechamento total em dois pontos da cavidade oral, todos os cliques são oclusivas ou africadas. Ladefoged e Traill descrevem a sequência de fases na produção de um clique como sendo a seguinte:

1. Uma quantidade de ar fica contida numa cavidade fechada formada pela elevação do ápice ou da lâmina da língua na frente da cavidade oral e ao mesmo tempo pela elevação do dorso da língua, fechando a passagem do ar no véu palatino ou na úvula.
2. O ar que está nessa cavidade fechada se rarefaz como consequência do movimento para baixo do centro da língua enquanto a oclusão anterior e a posterior são mantidas.
3. O ápice, a lâmina ou os lados da língua se deslocam para baixo, abrindo a oclusão existente na parte anterior da cavidade oral, o que gera o ruído do clique.
4. Abre-se a oclusão posterior.

Logo, todo clique tem uma ação da coroa da língua, que determina o tipo de clique, e uma ação velar ou uvular. Ladefoged e Traill denominam acompanhamentos do clique: a ação velar ou uvular; a atividade da laringe; o possível envolvimento da cavidade nasal. Para cada tipo de clique, há a possibilidade de ele ser surdo, sonoro, aspirado, nasal, e combinações como nasal murmurado. Para informações detalhadas, recomendo a leitura do artigo de Ladefoged e Traill, que traz inclusive ilustrações dos cinco tipos de cliques no momento do início da oclusão e no momento logo antes de seu estalo.

5.4 FRICATIVAS LATERAIS

Quando vimos os modos de articulação na seção 2.2, mencionei a existência das fricativas laterais, inexistentes no português. Para tratar delas num contexto em que a língua predominante é o português, é útil pensar em termos fonoaudiológicos. As duas fricativas coronais anteriores do português pronunciadas apenas com o deslocamento vertical da coroa da língua em direção aos alvéolos, [s] e [z], apresentam certos desvios na pronúncia na fala de alguns falantes. Se em vez de utilizar essas duas fricativas, que são alveolares, o falante aproximar a língua dos dentes superiores produzindo fricativas dentais, transcritas [θ ð], temos o desvio de fala denominado ceceo, ou também ceceo anterior. Por outro lado, se o desvio de pronúncia consistir na aproximação mais acentuada do corpo da língua com relação ao céu da boca, além de haver contato da língua com os alvéolos, isso pode fazer com que o ar passe produzindo ruído ao escapar pelos lados da língua. Isso dá origem ao chamado ceceo lateral: as fricativas com essas características são transcritas [ʃ], a surda, e [ʒ], a sonora. No final do capítulo há links em que se pode ouvir as duas pronúncias.

Um exemplo disso seria a pronúncia da expressão *cinco vezes*, que na pronúncia não desviante tem as fricativas alveolares [s] e [z]. Na pronúncia com ceceo, tem os sons [θ] e [ð]. E na pronúncia com ceceo lateral, tem os sons [ʃ] e [ʒ]. É importante ressaltar que esses chamados desvios dependem do que é comum na língua. Em línguas que possuem [θ] e [ð], e nas que têm os sons [ʃ] e [ʒ], não necessariamente temos ceceo anterior ou lateral, se esses sons forem usados na pronúncia de palavras que os contêm, mas apenas se substituírem o [s] e o [z] nas palavras que deveriam ser pronunciadas com esses dois últimos sons. Na prática, o que se deve observar é que os chamados desvios de pronúncia podem ser nada mais do que a utilização de mecanismos que estão à disposição do falante nas línguas em geral, mas que naquela língua específica não são usados.

5.5 ONDAS SIMPLES, COMPLEXAS E HARMÔNICOS

A vibração das pregas vocais na produção da voz, nos sons sonoros ou vozeados, é aproximadamente periódica, ocorrendo em ciclos de abertura e fechamento. O número de vezes que esse ciclo se repete num determinado intervalo de tempo nos dá a frequência de vibração das pregas vocais. Essa frequência na fala normal varia numa voz masculina adulta entre cerca de 80 e 150 vezes por segundo, ou entre 80 e 150 Hz (hertz). Numa voz feminina adulta fica em geral entre 150 e 250 Hz. E numa voz infantil de 300 Hz para cima, podendo chegar a 500 Hz ou mais. Considerando que nossos olhos captam imagens 20 vezes por segundo, mesmo na voz masculina, mais grave, a frequência de vibração das pregas vocais não é adequadamente perceptível a olho nu, só podendo ser vista com câmera muito lenta.

O tipo mais simples de onda são as ondas periódicas simples. Nelas temos um padrão que se repete ciclicamente, e em cada ciclo há apenas um pico e um vale, ou seja, em termos do deslocamento das partículas do ar, se atinge o afastamento máximo em um sentido e o deslocamento das partículas se inverte até atingir o afastamento máximo no outro sentido, sem inversão do movimento durante o

deslocamento entre esses extremos. Esse tipo de onda é raro no dia a dia, não sendo talvez encontrado na natureza. Uma das poucas fontes desse tipo de som é o diapasão, usado na música.

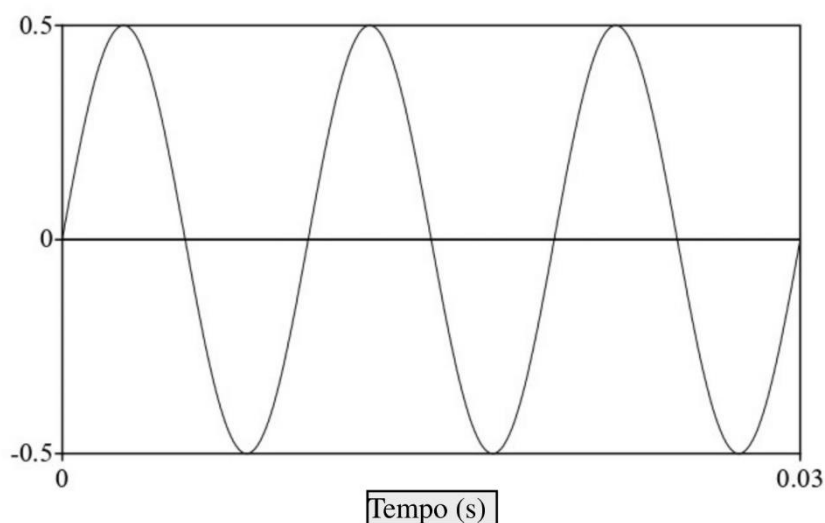


Figura 27. Oscilograma de onda de 100 Hz produzido no praat.

Essa figura feita no programa *praat* representa uma onda de 100 Hz. O tempo vai de 0 a 0,03 segundos, ou de 0 a 30 milissegundos. Nesses 30 milissegundos, o ciclo da onda se repete três vezes, o que dá seu período de 10 milissegundos, que corresponde a uma frequência de 100 Hz. A onda está na posição 0 no tempo 0. A cada 0,01 segundos ela passa duas vezes pela mesma posição 0, mas em uma delas o deslocamento é no sentido contrário ao inicial. Para completar o ciclo ela precisa passar uma segunda vez para se encontrar na mesma fase do momento inicial. 0,5 é a amplitude dessa onda, o deslocamento máximo que ela atinge com relação ao ponto neutro. Ela oscila até atingir um máximo de 0,5 unidades (não especificadas aqui) de distância da posição de repouso em cada uma das direções.

ONDAS COMPLEXAS

Na realidade, o quadro apresentado no parágrafo acima é bem simplificado, se comparado à vibração das pregas vocais. Embora elas vibrem de maneira aproximadamente periódica, sua vibração é semelhante à de uma corda de violão. Ela não é uma onda simples como a apresentada na figura acima, mas uma onda complexa, formada de várias ondas simples sobrepostas (concomitantes). Observe a figura 28.

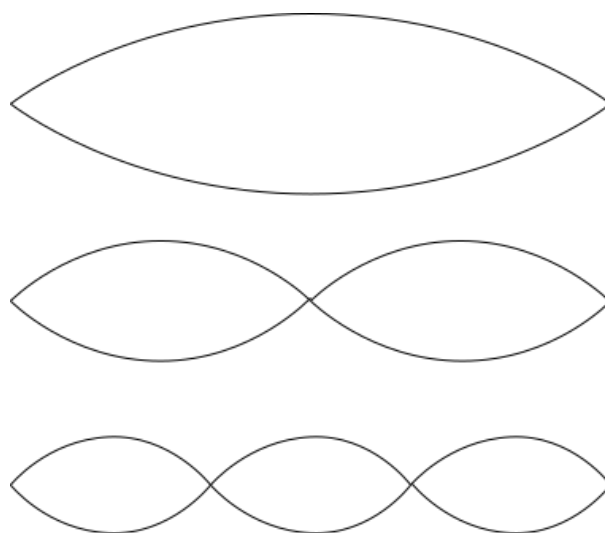


Figura 28. Os três primeiros modos de vibração de uma onda estacionária.

Na figura vemos os três primeiros modos ou padrões de vibração das pregas vocais, o que é idêntico ao que ocorre com cordas de um violão. Tanto as pregas vocais quanto as cordas de um violão são presas nas duas extremidades. Sendo assim, o modo mais simples de elas vibrarem é o que tem apenas dois nós, ou pontos de amplitude zero, que permanecem imóveis o tempo todo, um em cada extremidade. No meio de cada prega vocal ou da corda de violão, por outro lado, temos um ponto de amplitude máxima, que é chamado de antinó, entre outras denominações.

No segundo padrão de vibração, há um nó e um antinó a mais, e o mesmo acontece no terceiro com relação ao segundo. O número de padrões de vibração é em tese infinito. Podemos ver que o segundo padrão tem um ciclo completo, enquanto o primeiro tem meio ciclo, e o terceiro tem um ciclo e meio. Sendo assim, há uma relação matemática simples entre os diversos modos de vibração: as diversas ondas simples apresentam uma relação de frequência constante entre si. A mais grave, a que tem a frequência mais baixa, é chamada de frequência fundamental, ou F_0 (efe zero). As demais frequências são os múltiplos inteiros dessa frequência: seu dobro, seu triplo, seu quádruplo etc. Assim, se a frequência fundamental de vibração das pregas vocais num dado momento for de 100 Hz, ao mesmo tempo elas estarão vibrando a 200 Hz, 300 Hz, 400 Hz e assim por diante. Cada frequência dessas é denominada de harmônico.

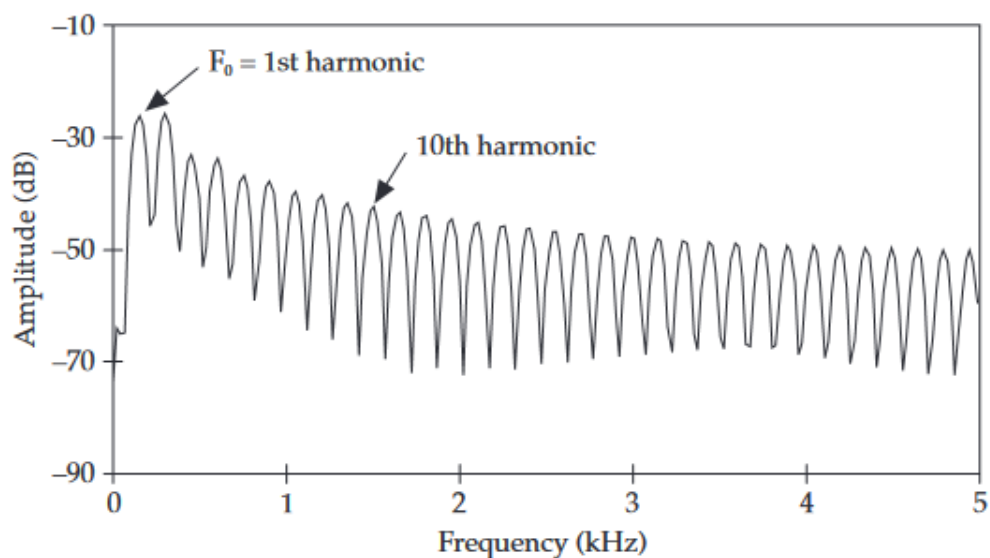


Figure 2.2 A power spectrum of the vocal cord vibration shown in figure 2.1. The first harmonic (which has the same frequency as the fundamental frequency of voicing) occurs at 150 Hz; therefore the tenth harmonic occurs at 1,500 Hz.

Figura 29. Espectro da vibração das pregas vocais.

Examinemos agora a relação entre a amplitude e a frequência dos harmônicos. Como podemos ver no espectro acima, os dois primeiros harmônicos são os mais intensos (os que possuem maior amplitude). Cada harmônico seguinte tem uma amplitude ligeiramente menor. Dependendo do tipo de voz empregado, os dois primeiros harmônicos têm intensidades diferentes, então esse decréscimo é o caso geral, mas pode não ocorrer com relação aos dois primeiros. Falarei brevemente dos tipos de voz no final deste capítulo, na seção 5.12.

O *pitch*, ou sensação de altura, está associado predominantemente a variações em F_0 . Quando F_0 sobe, temos um som mais agudo. Quando desce, um som mais grave. Em línguas como o português, uma das principais funções da variação de F_0 no decorrer de um enunciado é a entoação (ver cap. 13). Através dela podemos sinalizar, por exemplo, que o enunciado é uma afirmação ou uma pergunta.

5.6 FORMANTES

Se a voz produzida nas pregas vocais fosse ouvida diretamente sem alterações, não ouviríamos os sons da fala, mas simplesmente perceberíamos uma melodia, que corresponderia à variação da frequência de vibração das pregas vocais.

As pregas vocais funcionam como a fonte de som periódico de nosso aparelho fonador (Fant 1960). O som produzido por elas é modificado pelo trato vocal, a extensão que vai das pregas vocais até os lábios e narinas. O trato vocal recebe as ondas provenientes das pregas vocais, e as modifica de acordo com sua configuração, sua extensão etc. Uma das consequências da configuração do tubo pode ser entendida se comparada a um instrumento musical. Quanto maior o instrumento, mais grave o

som produzido. Um violino tem o som bem mais agudo que um contrabaixo, e um flautim, um som mais agudo que uma flauta. Analogamente, uma criança tem voz bem mais aguda que um adulto.

Depois que esse conjunto de ondas atravessa nosso trato vocal, boa parte dos harmônicos tem sua intensidade abafada, amortecida, mas alguns tem sua intensidade amplificada. O trato vocal à frente das pregas vocais modifica as propriedades da onda complexa produzida, ou seja, filtra essa onda, de forma que certas faixas de frequência são amplificadas, enquanto outras são abafadas ou amortecidas. É isso o que diz a teoria de fonte e filtro da produção da fala de Fant.

O que determina o efeito de abafamento ou amplificação de cada frequência específica são as propriedades da cavidade em que o som produzido pela fonte reverbera. Dependendo do formato específico, dimensões e material da cavidade, mudam suas propriedades de ressonância. O prefixo *re-* traz aqui essa implicação de que não estamos falando da fonte, que soa, ou produz um som, mas do filtro, que recebe um som e ressoa. Ou seja, a fonte soa, o filtro ressoa. O resultado é que a onda complexa terá picos em frequências que não guardam entre si a relação de múltiplos inteiros. Esses picos são denominados formantes. O pico de frequência mais baixa é denominado F1, o segundo, F2, e assim por diante. A identificação das vogais é feita principalmente com base em F1 e F2, embora F3 também possa contribuir.

As frequências que são amplificadas são as chamadas frequências naturais de vibração de um corpo. Um exemplo cotidiano bem acessível é empurrar um balanço em que uma criança está sentada. Se os empurrões são sincronizados com o vaivém do balanço, a amplitude da oscilação do balanço vai ficando cada vez maior. Há também os famosos casos de alguém quebrando uma taça de cristal com a voz. Para que isso aconteça a onda sonora produzida precisa ter determinada frequência e intensidade, além de ser sustentada por um certo tempo, como se vê no vídeo cujo *link* está ao final do capítulo.

Um exemplo impressionante do efeito da ressonância é a queda da Ponte de Tacoma Narrows. No final do capítulo há um *link* para que se possa ver o que aconteceu e outro que explica em certo detalhe o que causou a queda. O vento entrou em ressonância com a ponte e fez com que ela vibrasse com amplitude cada vez maior, até que seu vão livre se rompeu e a ponte desabou. A engenharia aprendeu muito com a queda dessa ponte, e isso ilustra bem a frase que veio num biscoitinho da sorte quando pedimos comida chinesa em casa: “O fracasso é a mãe do sucesso.”

Observemos agora tabelas com os valores médios dos formantes de vozes femininas e masculinas de falantes do português brasileiro e do europeu (respectivamente PB e PE nas legendas):

Formantes	[i]	[e]	[ɛ]	[a]	[ɔ]	[o]	[u]
F_1	307	425	646	910	681	442	337
F_2	2676	2468	2271	1627	1054	893	812

Tabela 1. Valores médios típicos de formantes de falantes femininas do PB

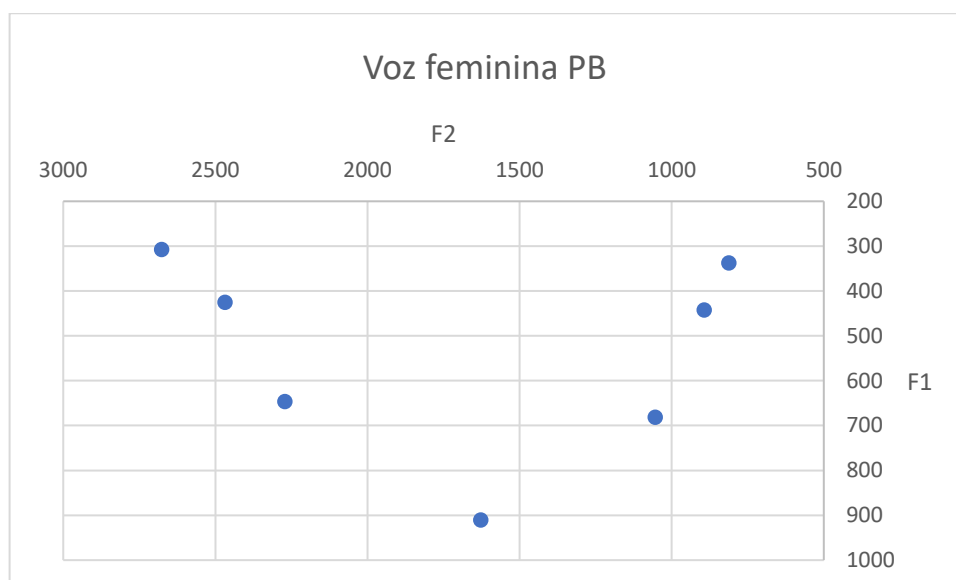


Gráfico 1. Valores médios típicos de formantes de falantes femininas do PB

Formantes	[i]	[e]	[ɛ]	[a]	[ɔ]	[o]	[u]
F_1	313	402	511	781	592	422	335
F_2	2760	2508	2360	1662	1118	921	862

Tabela 2. Valores médios típicos de formantes de falantes femininas do PE

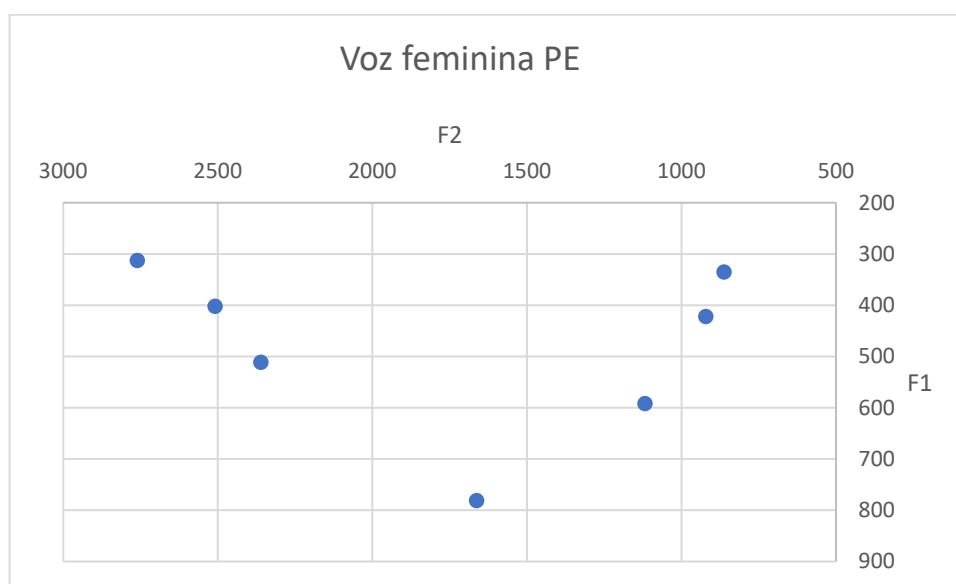


Gráfico 2. Valores médios típicos de formantes de falantes femininas do PE

Formantes	[i]	[e]	[ɛ]	[a]	[ɔ]	[o]	[u]
F_1	285	357	518	683	532	372	310
F_2	2198	2028	1831	1329	927	804	761

Tabela 3. Valores médios típicos de formantes de falantes masculinos do PB

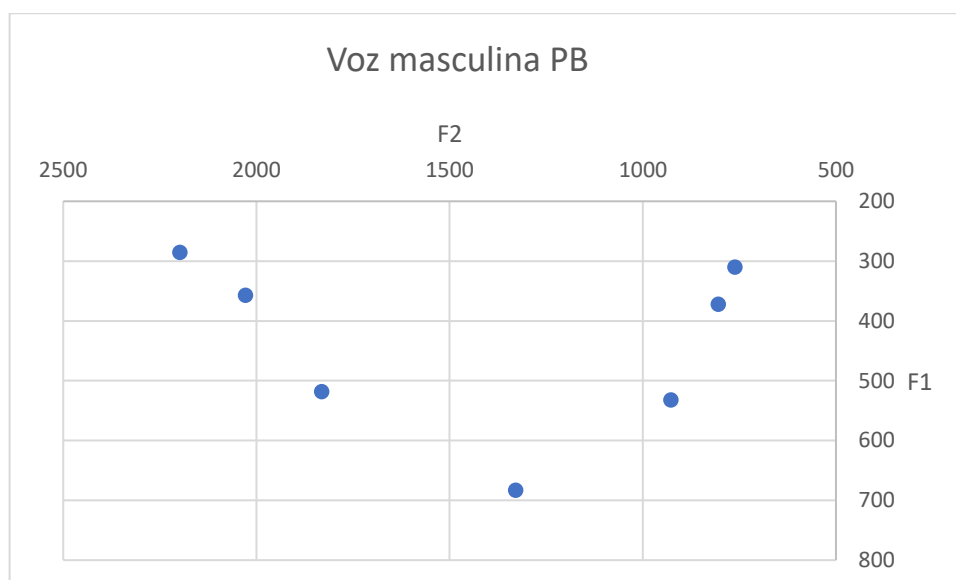


Gráfico 3. Valores médios típicos de formantes de falantes masculinos do PB

Formantes	[i]	[e]	[ɛ]	[a]	[ɔ]	[o]	[u]
F_1	284	355	455	661	491	363	303
F_2	2161	1987	1836	1365	934	843	814

Tabela 4. Valores médios típicos de formantes de falantes masculinas do PE

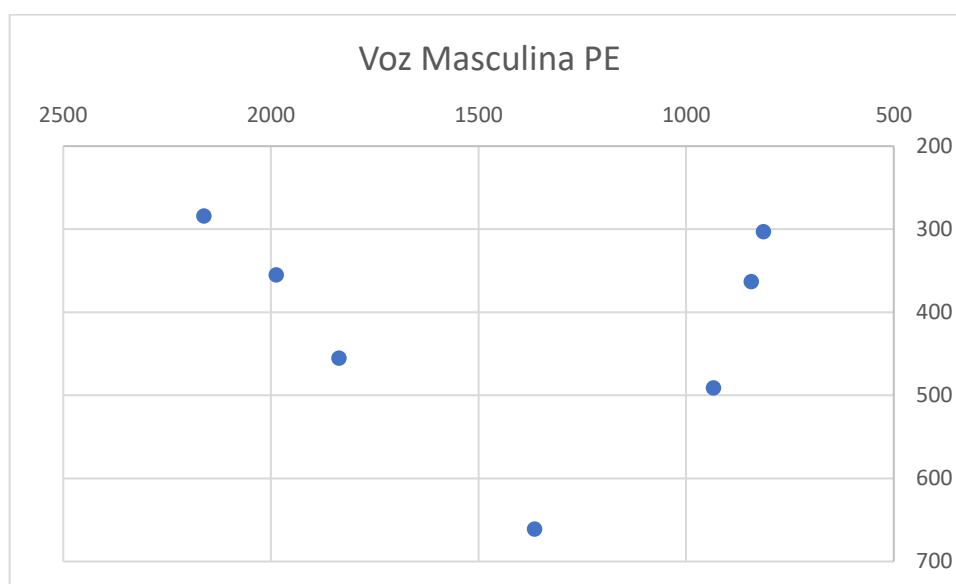


Tabela 4. Valores médios típicos de formantes de falantes masculinas do PE

Fonte dos dados das tabelas: Escudero et al. (2009)

Gráficos de elaboração própria com os dados de Escudero et al. (2009)

Como vemos pelas tabelas e gráficos, para cada um dos 4 subgrupos de falantes, as vogais podem ser divididas em quatro grupos com base no seu F1, sendo que nas duas últimas o PB e o PE parecem se diferenciar:

- as que têm F1 mais baixo (abaixo de 350 Hz nas vozes femininas e por volta de 300 Hz nas vozes masculinas, [i u]);
- as que têm F1 medianamente baixo (entre 400 e 450 Hz nas vozes femininas e entre 350 e 375 Hz nas vozes masculinas), [e o];
- as que têm F1 medianamente alto (entre 650 e 700 Hz nas vozes femininas do PB, entre 500 e 600 Hz nas vozes femininas do PE, e pouco acima ou abaixo de 500 Hz nas vozes masculinas), [ɛ ɔ];
- e a que tem F1 alto (acima de 900 Hz nas vozes femininas do PB, perto de 800 Hz nas vozes femininas do PE e entre 650 e 700 Hz nas vozes masculinas), o [a].

O primeiro grupo é o das vogais fechadas ou altas; o segundo, o das médias fechadas (ou médias altas); o terceiro, o das médias abertas (ou médias baixas); e a última é a vogal aberta (ou baixa). Podemos ver claramente como o F1 está relacionado diretamente com o grau de abertura da cavidade oral. Quanto mais aberta a vogal, maior o valor do F1.

Examinando agora o F2, podemos ver outra correlação. Do [u] até o [i], o valor de F2 sobe o tempo todo. Isso nos indica que F2 está relacionado à anterioridade ou posterioridade da vogal. Mais especificamente quanto mais posterior a vogal, mais baixo o valor de F2, e quanto mais alta a vogal, mais alto o valor de F2. Podemos facilmente observar que a diferença no valor de F2 entre as vogais posteriores é menor que a existente entre as vogais anteriores. Por isso utilizei um trapézio escaleno, que é assimétrico, para representar as vogais da seção 3.1 em diante.

Essa é a correlação mais comumente mencionada com relação ao F2, mas existe outra muito importante. O arredondamento dos lábios diminui consideravelmente o valor de F2, como pode ser visto comparando as vogais anteriores arredondadas e não arredondadas do francês. Vemos na tabela a seguir o F1 e o F2 médio das vogais anteriores do francês, segundo as medições reportadas em Georgetown et al. (2012).

Vogal	F1 médio	F2 médio
[i]	275	2585
[y]	276	2091
[e]	405	2553
[ø]	409	1599
[ɛ]	614	2306
[œ]	599	1678

Tabela 2. F1 e F2 das vogais anteriores do francês.

Considerando os dados da tabela, fica evidente que o arredondamento praticamente não afeta os valores de F1 das vogais anteriores com o mesmo grau de abertura, pois as duas vogais dos pares [i] e [y], [e] e [ø], [ɛ] e [œ] têm praticamente os mesmos valores desse formante. Só entre [ɛ] e [œ] há uma pequena diferença, mas que não passa de 2,5%. Quanto ao F2, é visível que o arredondamento

produz um grande efeito nele. Se comparamos as duas vogais dos três pares citados, constatamos que nos três casos há uma diferença considerável no valor de F2, indo de 23% no par [i] e [y], passando por 37% no par [ε] e [œ], e chegando a 60% no caso de [e] e [ø].

5.7 VOT (*VOICE ONSET TIME*)

O termo VOT, abreviatura de *voice onset time*, ou momento de início do vozeamento, proposto por Abramson e Lisker (1964), se refere ao momento numa sequência CV (consoante seguida de vogal) em que se inicia a vibração das pregas vocais. Isso pode ocorrer, por exemplo, já durante a consoante, ou só na passagem da consoante para a vogal, ou mesmo não ocorrer imediatamente após o início da articulação da vogal.

No quadro abaixo vemos três tipos de oclusivas com valores de VOT bem diferentes. São os três tipos mais comuns com relação a essa característica sonora, apresentando vibração ou não em cada porção dos segmentos. A vibração das pregas vocais está sendo representada por uma linha zigzagueada.

	oclusiva	vogal
pa		~~~~~
ba	~~~~~	~~~~~
p ^h a		~~~~~

Quadro 1. VOT em sequências CV

Como vemos no quadro idealizado, na produção de [pa], o vozeamento se inicia no instante em que se inicia a articulação da vogal. Se considerarmos que o tempo está avançando e o instante de início da vogal é o instante zero, o VOT de [pa] é aproximadamente igual a zero. Já na produção de [ba], o vozeamento se inicia durante a produção da oclusiva, sendo inferior a zero, pois se inicia antes de chegar o momento zero. Por fim, na produção de [p^ha], a vogal se inicia, mas há um atraso no início do vozeamento, caracterizando um VOT superior a zero. Em suma, temos o seguinte:

$$(9) \quad \text{VOT [pa]} = 0 \qquad \text{VOT [ba]} < 0 \qquad \text{VOT [p}^h\text{a]} > 0$$

5.8 CLASSES DE SONS

Nesta seção relacionaremos as características articulatorias de cada classe de som (vogais, oclusivas etc.) com suas características acústicas.

5.8.1 VOGAIS

Na produção das vogais, a corrente de ar passa sem grandes obstáculos pelo trato vocal. A língua pode avançar ou recuar, e assumir uma posição vertical bastante variável, desde bem distante do céu da boca, como na produção do [a], até bem próxima dele, como na produção do [u] e do [i]. Mas mesmo no caso destas duas vogais, a aproximação da língua ao céu da boca ainda é de um grau suficiente para que a corrente de ar passe bastante desimpedida. Além da posição da língua, os lábios também podem estar ativos na produção das vogais, ficando arredondados.

Como vimos na seção 5.6, são os formantes que distinguem as vogais umas das outras acusticamente. Para visualizarmos os formantes num programa como o Praat, usa-se o espectrograma, que é um gráfico que representa as propriedades do sinal acústico em duas dimensões. Embora sejam apenas duas as dimensões, ele é capaz de representar as três propriedades que caracterizam um sinal acústico: a frequência, a duração e a intensidade. A frequência é representada na vertical: quanto maior a distância de um sinal do eixo das abscissas, o horizontal, maior será sua frequência. A duração é representada na horizontal. A projeção do trecho sonoro sobre o eixo das abscissas representa a duração do som. Por fim, a intensidade do som é representada através de tons de cinza/preto. Quanto mais escuro for o som representado, mais intenso ele será. Dessa forma, os picos de energia acústica serão representados por regiões bem escuras.

Num espectrograma, os formantes das vogais são nitidamente perceptíveis, consistindo de linhas escuras que indicam as frequências em que se concentra a energia acústica. As frequências com a energia acústica mais baixa aparecem em cor mais clara, ao passo que aquela em que há mais energia acústica aparecem em cor mais escura.

Por não ter obstáculos, as vogais têm formantes bem nítidos no espectrograma, faixas escuras nas frequências em que a energia acústica se concentra, as quais se destacam num fundo mais claro.

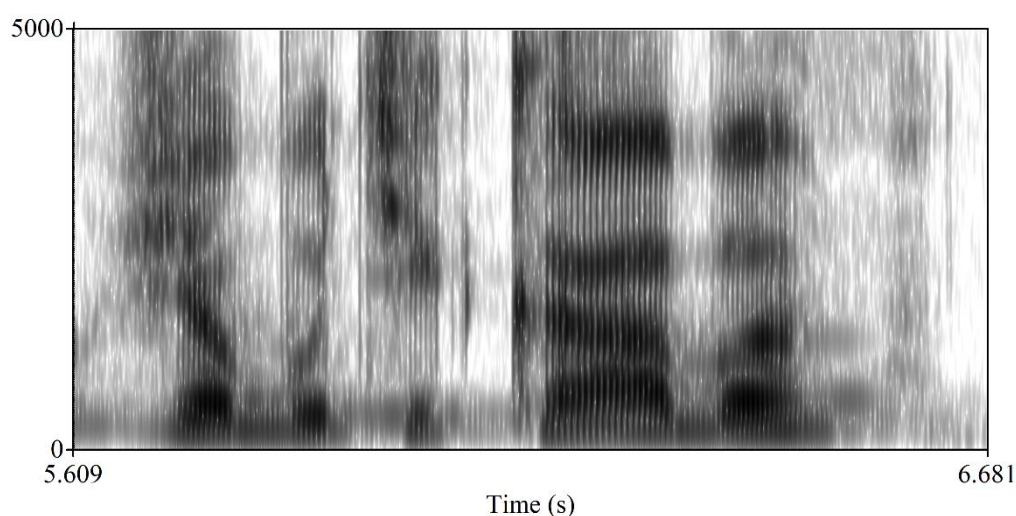


Figura 30. Espectrograma de *jabuticaba*. Fonte: Bia.

Os cinco trechos com linhas mais escuras no espectrograma são as vogais de *jabuticaba*. Observe como a vogal tônica [a], produzida pouco após a metade do trecho mostrado, é muito mais longa do que as outras.

5.8.2 OCLUSIVAS

Podemos pensar na produção das oclusivas como consistindo de três fases. A fase inicial é a de aproximação dos órgãos articuladores até se atingir a oclusão, isto é, o fechamento total do trato vocal. A segunda fase é aquela em que se mantém a oclusão. Por fim, a terceira fase é aquela em que a oclusão é desfeita, produzindo o som característico das oclusivas, o de uma breve explosão. Podemos chamar as três fases de aproximação, oclusão e afastamento. Numa oclusiva prototípica, durante as três fases, a pressão do ar dentro da cavidade oral é maior que a pressão externa, o que tenderia a fazer o ar sair do nosso trato vocal, se a cavidade estivesse aberta. Mas essa saída é impedida durante a fase de oclusão. Ocorre que, em alguns contextos, a fase de afastamento dos órgãos articuladores pode não produzir ruído. Há mais de uma razão para que isso aconteça. Uma primeira possibilidade é a de que na sequência seja produzida outra oclusiva. Quando isso ocorre, frequentemente a fase de afastamento dos órgãos articuladores da primeira oclusiva se sobrepõe à fase de oclusão da segunda, o que tem como resultado a não produção do ruído de explosão. Observe na figura 30 que o que seria a fase de aproximação da segunda oclusiva fica encoberta pela primeira, e a fase que seria de afastamento da primeira oclusiva fica encoberta pela segunda.

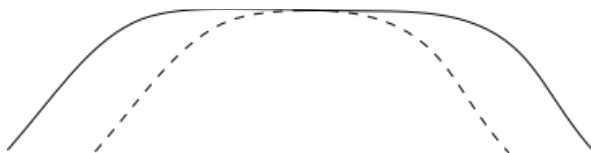


Fig. 31. Imagem esquemática de duas oclusivas em sequência

É bastante comum também, por outro lado, que em final de palavra numa língua específica as oclusivas sejam obrigatoriamente produzidas como não explodidas, ou *unreleased*. O termo *unreleased* do inglês contém na verdade uma certa imprecisão, pois não é que não haja a fase de afastamento na produção dessas oclusivas. O que ocorre é que, durante essa fase, a corrente de ar já foi interrompida ou diminuída em tal grau que não é produzido ruído. Para representar esse tipo de pronúncia, pode-se usar o diacrítico [']. Um termo que pode ser usado, embora possa parecer contraditório, é aplosivas. As plosivas prototípicas têm a explosão. Caso não tenham, são aplosivas. O inglês apresenta variação com relação à pronúncia das oclusivas em final de palavra. É o que ocorre na pronúncia de palavras como *cat*, *map* e *back*. Cada uma delas pode ser pronunciada com explosão audível ou não, isto é, com pronúncia plosiva ou aplosiva, como vemos na transcrição a seguir:

- (10) pronúncia plosiva: ['k^hæt] ['mæp] ['bæk]
pronúncia aplosiva: ['k^hæt̚] ['mæp̚] ['bæk̚]

No espectrograma, as oclusivas aparecem como uma faixa vertical clara, indicando a ausência ou quase ausência de energia acústica. Se a oclusiva for sonora ou vozeada, podemos perceber a chamada barra de sonoridade na parte inferior do espectrograma.

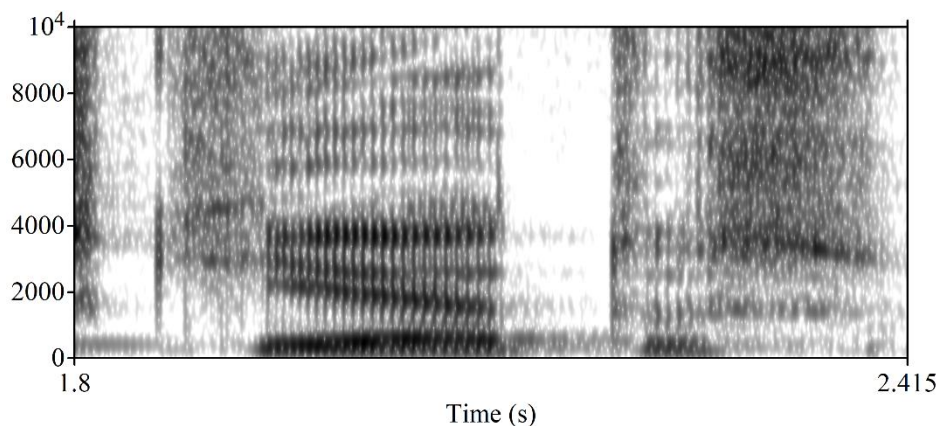


Figura 32. Espectrograma da palavra *quietos*. Fonte: Bia.

No espectrograma da figura 32 vemos a palavra *quietos*. Observe as duas faixas brancas verticais correspondentes às oclusivas surdas. O [k] não está tão claro **porque tem mais ruído**.

Uma característica acústica intrigante apresentada pelas oclusivas é a maneira como identificamos o seu ponto de articulação. Principalmente no caso das oclusivas surdas, que têm na sua fase de oclusão, a fase oclusiva propriamente dita, um silêncio total, fica claro que a identificação do ponto de articulação da oclusiva não pode se dar nessa fase. Como conseguimos, então, identificá-lo? Surpreendentemente, isso é feito na transição entre a oclusiva e os sons vizinhos. Se uma oclusiva estiver ladeada por duas vogais, é na transição dos formantes das vogais vizinhas que encontramos as pistas acústicas que possibilitam essa identificação. Isso ocorre mais especificamente principalmente na transição do F2, no chamado *locus* da oclusiva. Se na produção das vogais, o F2 está fortemente relacionado ao grau de anterioridade ou posterioridade de uma vogal, numa oclusiva ele está relacionado ao ponto de articulação, que se situa nessa mesma dimensão horizontal.

5.8.3 FRICATIVAS

Na produção das fricativas, como já vimos, o ar precisa passar numa obstrução bastante estreita, suficientemente estreita para produzir ruído. O fato de as partículas de ar terem que passar nessa constrição tão acentuada faz com que elas se choquem entre si, produzindo turbulência e, conseqüentemente, ruído. No espectrograma, as fricativas aparecem como faixas largas

verticalmente em que a energia acústica não fica concentrada em frequências distintas como nas vogais. A energia acústica fica distribuída numa ampla faixa de frequência.

Na figura abaixo, que traz um espectrograma da palavra *açaí*, com a frequência indo até 10.000 Hz, a fricativa fica bem nítida. Ela aparece no espectrograma como energia acústica espalhada numa faixa de frequência ampla, bem diferente do que ocorre com as vogais. O [s] tem uma concentração mais forte de energia acústica (um preto mais intenso) nas frequências mais altas. Repare que o trecho vocálico [a'i] tem faixas mais escuras só até por volta de 4.000 Hz, aproximadamente a frequência acima da qual temos a região escura da fricativa.

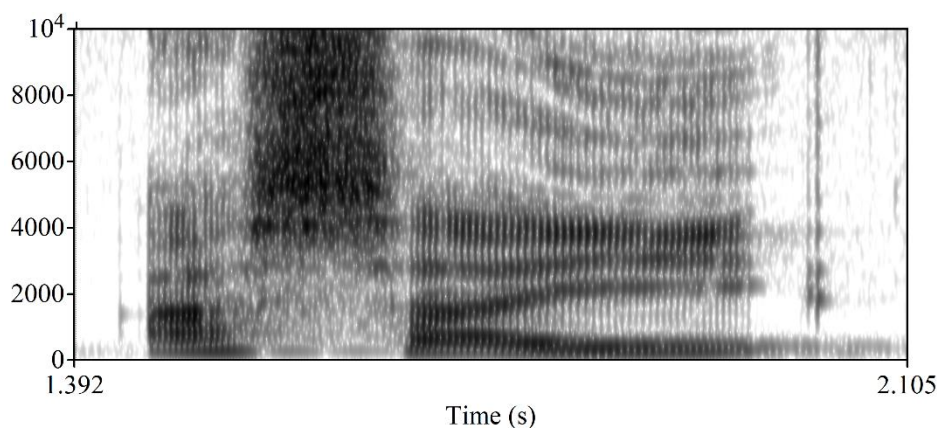


Figura 33. Espectrograma da palavra *açaí*. Fonte: Bia.

Como as fricativas são sons que se prolongam no tempo, diferentemente das oclusivas, nas quais há um ruído praticamente instantâneo, é bem mais fácil identificarmos qual a fricativa que está sendo produzida na duração dela própria. O ruído produzido no ponto de constrição máxima é turbulento e, portanto, aperiódico. Se a fricativa que está sendo produzida for surda, será esse o seu único componente sonoro, mas se ela for sonora ela terá também a vibração periódica das pregas vocais. Nesse caso, além da faixa larga escura em frequências mais altas, haverá também a barra de sonoridade na parte inferior do espectrograma.

Já o ponto de articulação da fricativa é identificável com base na faixa de frequências em que se encontra mais energia acústica. Há uma correlação entre o tamanho da cavidade à frente da constrição e essa faixa de frequência. Uma cavidade grande produz sons mais graves, e uma cavidade menor produz sons mais agudos. Dessa forma por exemplo, a energia acústica estará em faixas de frequência mais altas na produção de uma fricativa alveolar, como o [s], do que na produção de uma fricativa pós-alveolar, como o [ʃ].

5.8.4 AFRICADAS

Como as africadas consistem de duas fases, a primeira correspondente a uma oclusiva e a segunda correspondente a uma fricativa, seu espectrograma terá exatamente essas características: um trecho

de silêncio seguido de um trecho de ruído. Para que se tenha uma consoante africada, não basta que a fase inicial dela seja uma oclusiva e a final uma fricativa. É necessário ainda que o mesmo articulador ativo execute as duas fases. Isso faz parte de caracterizar uma única consoante com duas fases. Desse modo, [kʃ], [ps], [kf] e [px], por exemplo, não são africadas (confira por que cada uma dessas sequências não é). Já [kx], [ts], [pf] e [tʃ] podem ser, porque têm o mesmo articulador ativo tanto na fase de oclusiva quanto na de fricativa. Digo podem ser porque, dependendo da língua, essas sequências contam como duas consoantes, e em outras contam como africadas. É possível inclusive que a mesma língua admita as duas possibilidades. Um exemplo muito citado é o polonês *czy* [tʃɨ] ‘se (conjunção)’, que contrasta com *trzy* [tʃɨ] ‘três’. Observe que nesses casos é importante que na transcrição se use a ligadura nas africadas para distinguir quando se trata de uma consoante (tʃ) ou duas (tʃɨ).

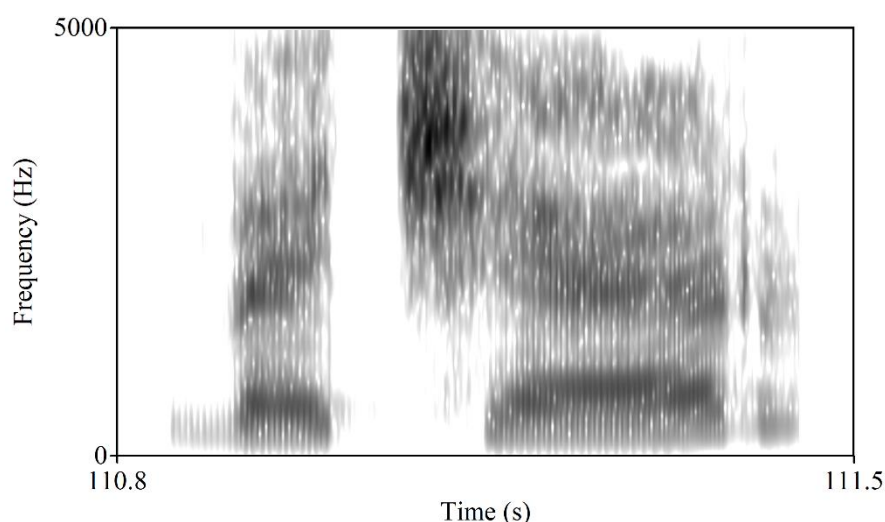


Figura 34. Espectrograma de *večer* [ˈvɛtʃɛr] em tcheco.

No espectrograma acima de [ˈvɛtʃɛr], podemos ver distintamente, logo após a primeira vogal, com seus formantes, a fase de silêncio seguida de uma fase de ruído em frequências relativamente elevadas. É a africada [tʃ].

5.8.5 APROXIMANTES

Embora o termo aproximante sugira que há apenas a aproximação de dois órgãos articuladores, uma definição acústica enfatiza o fato de que não é produzido ruído na sua articulação. Não é um termo perfeito, mas ficou consagrado assim. Como foi dito no capítulo 2, essa é distinção existente entre o [ð] do inglês *mother* e o [ð̞] do espanhol *nada*.

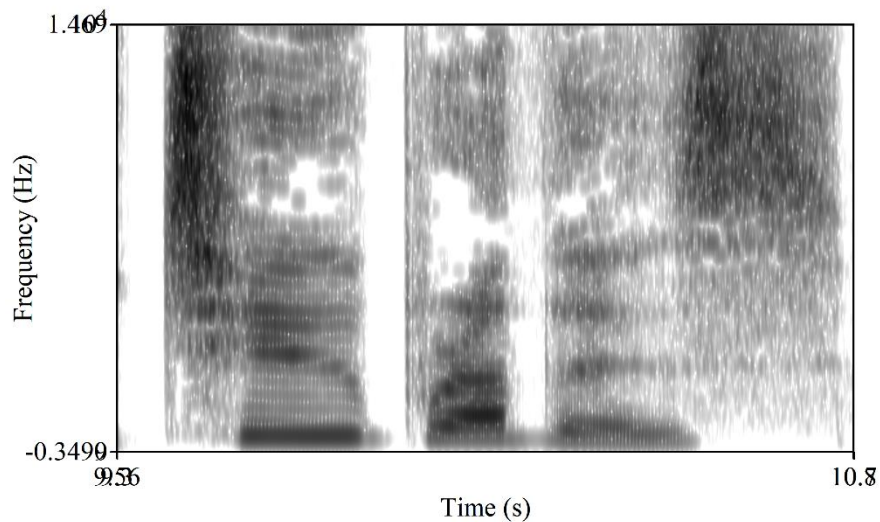


Figura 35. Espectrograma de *two brothers*

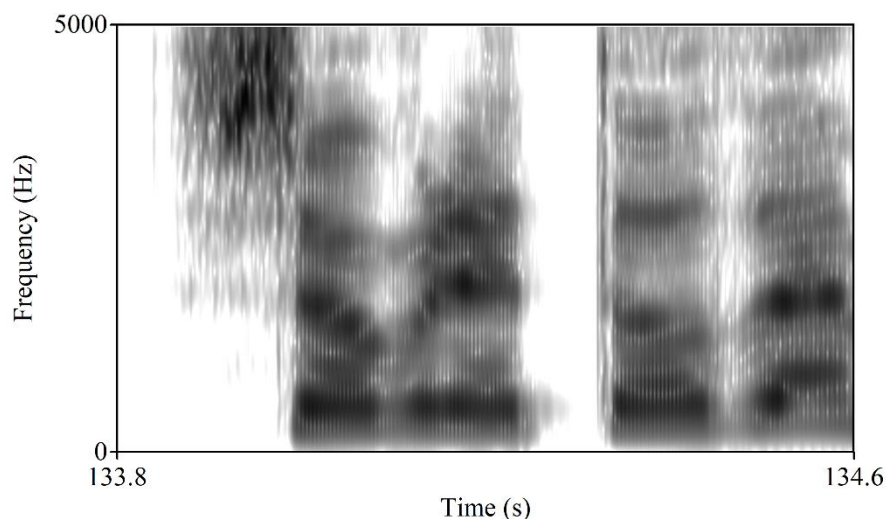


Figura 36. Espectrograma de *sobretudo* em espanhol.

Como podemos ver, no espectrograma de *two brothers* há duas oclusivas: uma logo no início, o [t], e outra pouco antes da metade, o [b]. A rigor, o [t] aparece africado, pois em inglês ele tem uma certa africacão, o que faz com que, numa transcrição mais estrita, ele seja representado como [tʰ].

A fricativa [ð] tem uma certa duração, não sendo quase instantânea, e aparece com um certo ruído, logo após o (s) da legenda embaixo, mas em termos de fricativa, ela é uma das menos ruidosas. Se examinamos agora o espectrograma de *sobretudo* em espanhol, temos o [t] logo após o *Time (s)* da legenda. Já o [ð] do espanhol aparece como uma perturbação muito breve do fluxo de ar presente nas vogais vizinhas, não produzindo ruído. Isso é consequência do fato de no espanhol não haver uma aproximação acentuada que cause turbulência, mas sim um estreitamento parcial e de pouca duração.

5.8.6 NASAIS

As consoantes nasais são produzidas com fechamento total da cavidade oral e com abertura concomitante da cavidade nasal. Isso faz com que o som que provém da faringe ressoe na cavidade oral fechada e ao mesmo tempo na cavidade nasal aberta. Como a ressonância em cada uma dessas cavidades terá características diferentes, e as duas estão acopladas ou conectadas, acaba havendo interferência da ressonância das duas cavidades, o que dá origem aos chamados antiformantes, regiões com queda acentuada da energia acústica. No espectrograma as nasais têm características semelhantes às das vogais, mas atenuadas. Os formantes das nasais têm ainda a característica de terem uma largura de banda muito acentuada, devido ao amortecimento. A identificação do ponto de articulação das nasais é feita com base nos formantes, como nas vogais, e nas suas transições, como nas oclusivas. Na seção seguinte há um espectrograma que mostra consoantes nasais.

5.8.7 LATERAIS

A produção das laterais tem características semelhantes à produção das nasais, pois em ambos os casos há duas cavidades acopladas, o que dá origem a antiformantes. Assim como nas nasais, há uma cavidade com interrupção da saída do ar e outra em que a corrente se mantém ininterrupta. A cavidade em que há interrupção da saída do ar nas laterais fica acima da língua, na parte central. Como o próprio nome indica, no entanto, dos lados dessa obstrução total, a corrente passa desimpedida. Assim como as consoantes nasais, então, as laterais aparecem no espectrograma como se fossem vogais atenuadas, com formantes menos nítidos.

As laterais são normalmente classificadas como aproximantes. Apesar de essa denominação ter o inconveniente de sugerir que não há contato, mas só aproximação de órgãos articuladores, é o termo usado tradicionalmente. Como já foi dito, o que caracteriza uma aproximante, em oposição às fricativas, é que aquelas não envolvem a produção de ruído, enquanto que estas envolvem. Se na produção de uma lateral, a língua se aproximar tanto do ponto de contato que o ar acabe passando numa cavidade muito estreita que gera a produção de ruído, teremos uma fricativa lateral, como vimos na seção 5.4. A identificação do ponto de articulação das laterais é feita com base nos seus formantes.

No espectrograma abaixo de *as lâminas*, vemos as fricativas com concentração de energia acústica nas frequências mais altas, e as vogais, com faixas escuras em frequências abaixo de 4.000 Hz. Os trechos intercalados em que os formantes aparecerem mais apagados são, na sequência, o [l], o [m] e o [n] de *lâminas*. Repare também que mesmo em frequências mais baixas há regiões brancas. Elas correspondem aos antiformantes.

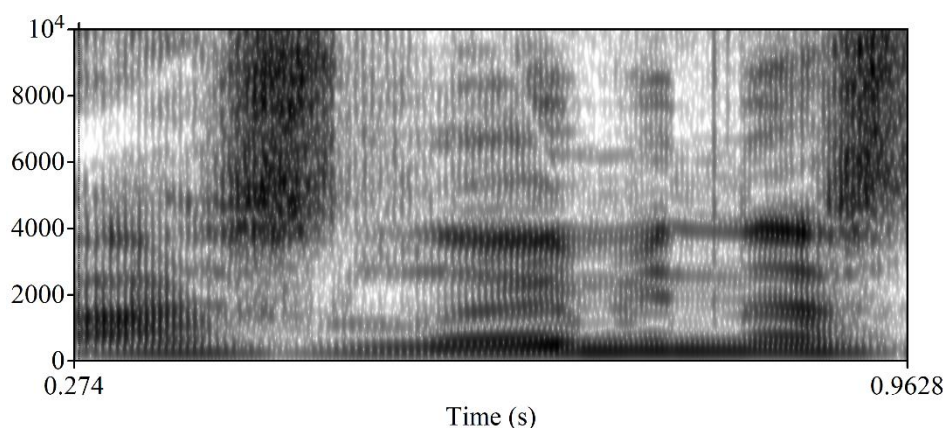


Figura 37. Espectrograma do trecho *as lâminas*. Fonte: Bia

5.8.8 RÓTICOS

Na seção 2.4, falei dos erres do português e comentei que eles podem ser analisados como constituindo uma classe com semelhança de família, sem que necessariamente haja algum traço comum a todos eles. Um termo que é utilizado para se referir a esses “erres” é róticos. Esse termo tem origem no nome da letra rô do grego, <ρ>, correspondente ao <r> do alfabeto latino.

Os róticos prototípicos são o tepe e a vibrante alveolar. Tratarei dos dois nesta seção, propondo também uma terminologia vernácula para nos referirmos a *taps* e ao *flaps*, entre outros pontos.

Se consultarmos a tabela de consoantes pulmonares do alfabeto fonético internacional, veremos que na linha correspondente aparecem dois nomes: *tap* ou *flap*. Ambos podem ser descritos como um tipo de colisão contra um obstáculo. O movimento executado para produzir um *tap* ou tepe é essencialmente apenas uma elevação vertical da coroa da língua, como uma oclusiva extremamente breve, em que a coroa colide com os alvéolos e se afasta em seguida, voltando à posição original. Já num *flap* ou flepe, há, por assim dizer, uma colisão lateral, ou, melhor ainda, tangencial, como um carro que passa raspando num muro. Fonologicamente, é extremamente raro alguma língua contrastar os dois tipos de articulação, mas eles não são iguais foneticamente, embora seja incomum os textos distinguirem os dois tipos de segmentos. Uma língua em que a literatura aponta contraste é o rikbaktsa (Mato Grosso), em que há, por exemplo, o par mínimo [a'ɾoʔ] ‘teu pai’, com flepe retroflexo, e [a'roʔ] ‘papagaio’.

Para propor uma nomenclatura com termos em português para *tap* e *flap*, preciso incluir também a vibrante. Em termos de produção, a vibrante é bem diferente do tepe porque, se neste há o arremesso da coroa contra os alvéolos, na vibrante é produzido um obstáculo total que é ativamente mantido, sendo, todavia, interrompido de forma reiterada pela corrente de ar forte, de forma análoga ao que ocorre quando vibram as pregas vocais.

É comum na literatura se referir a tepes e vibrantes como membros da mesma categoria, incluindo ambos na categoria das vibrantes, mas distinguindo-os com os adjetivos *simples* e *múltiplo*. Nesse caso, o tepe é denominado *vibrante simples* e a vibrante propriamente dita, *vibrante múltipla*.

A meu ver, contudo, a terminologia usada em disciplinas diversas, quando possível, deve ser uniforme. O termo *vibrante* está obviamente relacionado ao termo *vibração*, usado para falarmos das pregas vocais, por exemplo, mas também em física, engenharia, música etc. Em todas essas disciplinas, o termo *vibração* implica algo que se repete, como vemos na citação de Singiresu (2009: 6), um livro de engenharia acústica: “Qualquer movimento que se repita após um intervalo de tempo é denominado *vibração* ou *oscilação*.” Em Roederer (1998: 38), sobre a física e a psicofísica da música, por sua vez, lemos: “Há um certo tipo de movimento em que o ponto material segue um padrão temporal que se repete sempre. Este é o chamado movimento periódico simples, ou *vibração*. É o tipo de movimento de maior importância na física da música.” Se empurro um livro e ele se desloca sobre a mesa, isso não constitui uma *vibração*. Se arremesso uma bola contra a parede, ela bate e volta para minhas mãos, isso também não constitui uma *vibração*. Esse segundo movimento é semelhante ao do tepe, e não se justifica caracterizá-lo como vibrante.

Em vista de definições como essas, não vejo vantagem nenhuma, ao contrário, vejo prejuízo em usar o termo *vibrante (simples)* para nos referirmos ao tepe. Ele é executado através de uma única batida, um único contato da coroa da língua com os alvéolos.

Se se quiser manter a caracterização do tepe como simples e da vibrante como múltipla, proponho que se use o termo *percussiva*, relacionado ao termo percussão. Se pensarmos em instrumentos de percussão, eles se dividem em instrumentos de percussão direta, que envolvem batidas, pancadas, e percussão indireta, os que envolvem agitação ou fricção. Utilizo o termo *percussão* aqui no sentido de percussão direta, em que algo bate em outra coisa.

O tepe envolve uma batida, podendo ser definido como uma percussiva simples. A vibrante envolve uma repetição, podendo ser definido como uma percussiva múltipla. Esses usos obviamente permitem agrupar os dois tipos de fones numa categoria só. Se não se tem esse objetivo, pode-se manter o termo vibrante e utilizar o termo percussiva apenas para o tepe. Por fim, o flepe, pode ser definido como uma consoante tangencial.

Proponho duas terminologias alternativas, entre as quais se pode escolher com base, por exemplo, no inventário da língua de que se estiver tratando. Para simplificar, trato aqui apenas de sons coronais.

	Simples x múltipla	Percussiva x vibrante
<i>tap</i>	percussiva simples	percussiva
vibrante	percussiva múltipla	vibrante
<i>flap</i>	tangencial	tangencial

Quadro 2. Terminologias propostas para as consoante róticas.

A primeira tem a vantagem de englobar tepe e vibrante numa classe. A segunda tem a vantagem de usar termos compostos por uma única palavra.

No espectrograma, as percussivas (simples) aparecem como se fossem oclusivas extremamente breves; as vibrantes, aparecem como se fossem uma sucessão de percussivas, como vemos, respectivamente, nas figuras 38 e 39. Compare na figura 38 as duas oclusivas, os dois trechos com faixas verticais brancas ou quase brancas largas, o [g] logo no início e o [t] pouco depois da metade, com a percussiva, muito breve, também uma faixa vertical branca, mas muito mais estreita, situada um pouco antes da metade da figura.

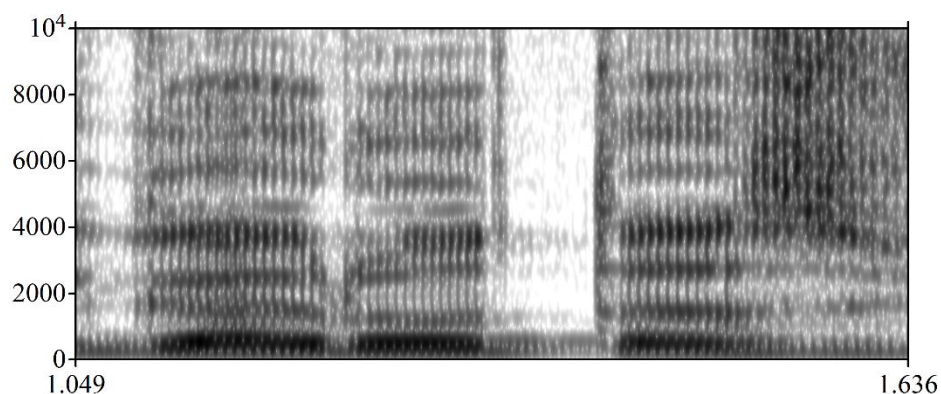


Figura 38. Espectrograma de *garotas*. Fonte: Bia

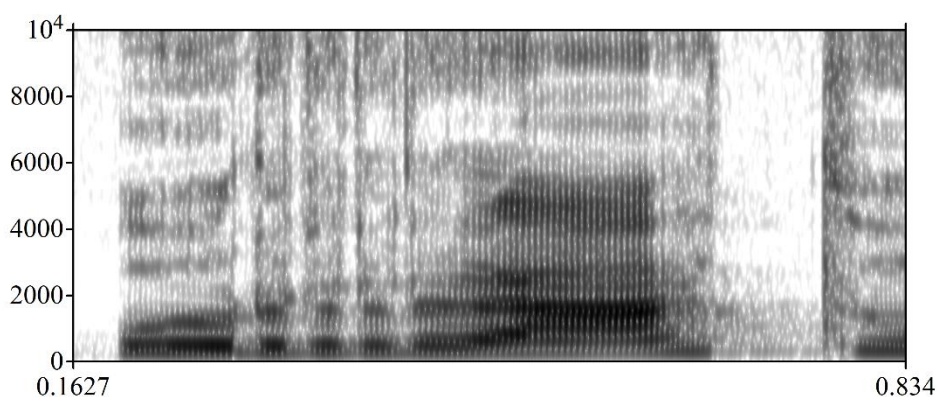


Figura 39. Espectrograma de *o rato* (com vibrante). Fonte: Bia

Na figura 39 vemos 4 faixas verticais estreitas ao ser pronunciado *o rato* com uma vibrante.

5.9 FONÉTICA AUDITIVA

Como disse na introdução deste capítulo, sem termos alguma compreensão de como funciona nossa percepção da fala, nossa concepção da fonologia está longe de ser abrangente. Para termos uma compreensão básica de como funciona nossa audição, examinemos agora a estrutura do sistema auditivo. Esta seção e a seguinte se baseiam no capítulo 4 de Johnson (2012).

Para deixar claro o quanto o lado auditivo da fonética é importante, e como a base corpórea de nossa cognição, como defende a linguística cognitiva, cumpre ressaltar já de início que a definição de som se baseia naquilo que é captável pelo sistema auditivo humano. Se uma vibração estiver fora dos

limites definidos por nossa audição, não é um som, embora possa ter várias propriedades semelhantes às de um som.

Concentrando-nos na frequência das vibrações existentes no ar, uma vibração que tem uma frequência acima de 20.000 Hz (ou 20 kHz) está além do que o ouvido humano consegue captar. São as frequências que definem o ultrassom. Um cachorro consegue captar até 45 kHz (na verdade, isso depende muito da raça), e um gato até cerca de 65 kHz. Um morcego capta frequências que estão acima de 100 kHz, e talvez possam atingir 200 kHz. Entre os animais aquáticos, a toninha ou boto, parente dos golfinhos, pode atingir até cerca de 160 kHz. Já uma galinha ouve somente até pouco mais de 7 kHz. Um canário até cerca de 8 kHz.

Por outro lado, as galinhas conseguem ouvir frequências mais graves do que o ser humano. Para nós, o limiar inferior da audição se situa em 16 Hz. Para as galinhas, em 9 Hz. Para os camundongos, em 900 Hz. As galinhas conseguem captar o que é denominado de infrassom: uma vibração com frequência abaixo dos 16 Hz. O rinoceronte de Sumatra consegue captar frequências de apenas 3 Hz! Isso ilustra o quanto a definição do que é audível depende de cada espécie, envolvendo tanto características fisiológicas quanto mentais da percepção auditiva. Como nosso ouvido capta até 20.000 Hz, e as frequências linguisticamente relevantes mais agudas estão por volta de 10.000 Hz, mesmo uma perda de boa parte das frequências mais agudas não chega a prejudicar a percepção dos sons da fala.

A linguística cognitiva (Langacker, Lakoff *datas*) defende que a linguagem é corporeada em nós, sendo **constrained** por sua estrutura e seu funcionamento. Lakoff (1987: xiv) afirma que “as estruturas nas quais se fundamenta nosso sistema conceptual se desenvolvem com base em nossa experiência corporal; além disso, o âmago de nossos sistemas conceptuais é diretamente alicerçado na percepção, no movimento corporal e na experiência de uma natureza física e social”. Pena que, por enquanto, a linguística cognitiva ainda não tenha se dado conta do quanto considerar a fonologia daria peso a sua argumentação de que nosso conhecimento linguístico é profundamente baseado em como nosso corpo funciona.

Um caso interessante relacionado à frequência é do telefone, que tradicionalmente transmite frequências entre 300 e 3.400 Hz apenas. Para boa parte dos sons da fala isso já produz um resultado bastante bom, principalmente para as vogais. O mesmo não se aplica às fricativas. Se olharmos o espectrograma de um [f] vemos que não há frequências que se destacam de maneira acentuada. O [f] tem um espectro bastante plano (*flat*), sem grandes picos e vales. Já o [s] tem uma faixa de frequência que se destaca muito: a faixa entre 4.000 e 10.000 Hz. Se o telefone corta as frequências acima de 3.400 Hz, essa faixa de frequência primordial é eliminada. Isso acaba fazendo com que seja impossível distinguir um [s] de um [f] ao telefone, e precisamos dizer “s de sapo” e “f de faca”. A telefonia de banda larga, que transmite frequências até 7.000 Hz, atenua muito esse problema.

Embora em medicina no Brasil se tenha convencionado utilizar o termo *orelha* para fazer referência às regiões externa, média e interna do aparelho auditivo, isso parece fazer violência à terminologia

tradicional em língua portuguesa, em que *orelha* se refere somente à parte mais externa do sistema auditivo e *ouvido* a sua parte interna. Como a orelha na terminologia vernácula apenas direciona as vibrações e é a parte interna que ouve propriamente, me parece um contrassenso adotar o termo *orelha*. Uso, portanto, o termo *ouvido* aqui, como se faz em Portugal. Não considero o fato de que *ear* em inglês ou *oreille* em francês se referem a todo o aparelho auditivo seja um argumento válido para o português. Adoto uma terminologia de base vernácula.

O sistema auditivo se subdivide em central e periférico. O sistema auditivo periférico é o que se localiza no ouvido. O central abrange os nervos que transmitem os estímulos auditivos para o cérebro e o próprio cérebro.

O sistema auditivo periférico é a parte do sistema auditivo que não está contida no cérebro. Sua tarefa é converter sinais acústicos em sinais neurais. As ondas sonoras incidem no ouvido externo, passam pelo canal auditivo até o tímpano, o qual é uma membrana fina que responde às flutuações de pressão do ar. Um conjunto de três pequenos ossos (estribo, bigorna e martelo) transmite esses movimentos para o ouvido interno. No ouvido interno está contida a cóclea, uma estrutura em espiral que contém um fluido. No interior da cóclea se localiza a membrana basilar, que percorre a cóclea do início ao fim.

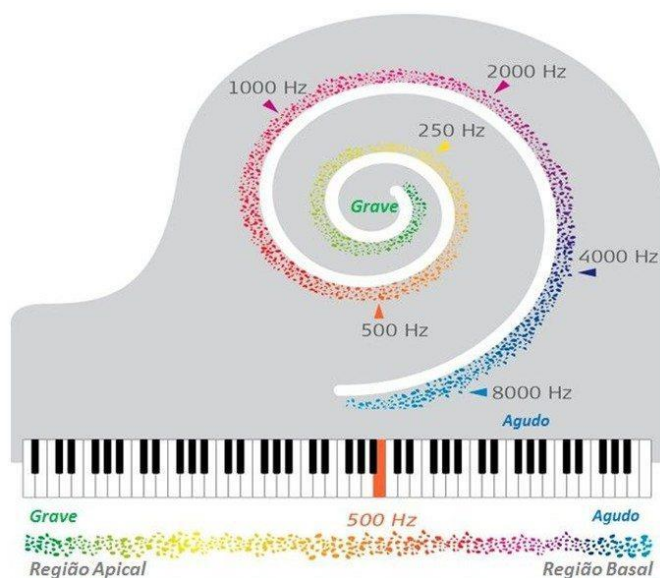


Figura 40. Ilustração da cóclea e da membrana basilar. Fonte: biosom.com.br.

Uma propriedade importante da membrana basilar é o fato de que ela não tem a mesma espessura em toda a sua extensão. A extremidade mais fina se localiza perto do conjunto de ossos (estribo, bigorna e martelo). Um ponto crucial é que essa extremidade responde a componentes de uma onda que tenham as frequências mais altas.

Cada nervo auditivo enerva uma parte da membrana basilar, sendo assim responsável por transmitir informações sobre uma determinada porção da gama de frequência. O que se pode concluir é que o ouvido interno realiza uma decomposição de uma onda complexa em seus componentes de frequências distintas, o equivalente a uma análise de Fourier, a análise matemática que tem o objetivo

de decompor uma onda complexa em ondas simples. Podemos pensar na cóclea como um piano às avessas. Se no piano cada tecla produz uma frequência, na cóclea cada nervo capta uma faixa de frequência.

Diferentemente da decomposição realizada por uma análise de Fourier matemática, que decompõe ondas complexas em ondas simples, a que o aparelho auditivo realiza não é linear. Essa não linearidade da percepção humana da frequência se deve ao fato de que ela é determinada pela estrutura da membrana basilar. Como se pode ver da figura adaptada de Johnson, a extensão da membrana responsável por captar frequências mais baixas é maior do que a responsável por captar frequências mais altas. Se nossa percepção auditiva vai de 16 Hz a 20.000 Hz, o ponto médio fica bem próximo de 10.000 Hz. Apesar disso, a porção da membrana basilar que capta frequências acima de 10.000 Hz é visivelmente muito menor do que a porção que capta frequências abaixo desse valor.

Com isso, vemos que a percepção da frequência pela mente humana não é algo linear (repetido?), mas moldado pela forma como funciona nosso sistema auditivo, sendo, portanto, não algo estritamente mental, mas alicerçado no funcionamento de nosso corpo físico.

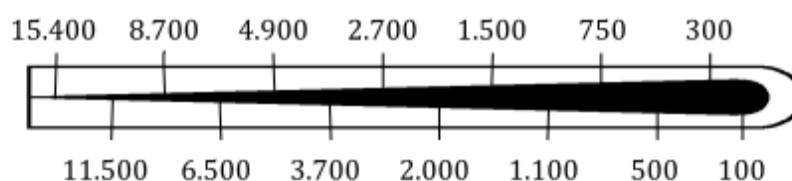


Figura 41. Localização de algumas frequências na membrana basilar.

Ilustração baseada em Johnson (2012: 89),
com pontos em vez de vírgulas para separar os milhares.

Detalhando ainda mais, se observarmos a parte inferior da figura acima, abaixo da linha preta de espessura variável, veremos que a extensão do trecho da membrana basilar responsável por detectar frequências de até 500 Hz tem uma dimensão muito próxima ou equivalente às outras divisões marcadas abaixo da linha. Apesar de essas subdivisões terem uma dimensão equivalente, elas não são responsáveis por detectar faixas de frequência equivalentes. Denominemos as subseções demarcadas na parte inferior da figura da membrana basilar com as letras A, B, C, D, E e F, indo da direita para a esquerda. A é responsável por detectar uma faixa de frequência de 500 Hz; B, uma faixa de 600 Hz (1.100 – 500); C, uma faixa de 900 Hz (2.000 – 1.100); D, uma faixa de 1.700 Hz (3.700 – 2.000); E, uma faixa de 2.800 Hz (6.500 – 3.700) e finalmente F, uma faixa de 5.000 Hz (11.500 – 6.500). Podemos verificar, dessa forma, que a percepção de faixas de frequência mais baixa é muito mais privilegiada do que a de faixas de frequência mais altas. Isso faz com que as vogais, cujos

componentes acústicos mais intensos estão em frequências mais baixas, sejam os sons mais perceptíveis para nós.

Na prática, isso equivale a dizer que o espectro de um som da fala tanto em termos de intensidade quanto em termos de frequência medido com auxílio de instrumentos não tem uma correspondência exata com aquilo que é percebido pelo ouvido humano. A discrepância entre as duas medições é mais acentuada com relação a sons que envolvem faixas mais altas de frequência, principalmente ruídos de soltura de oclusivas e fricativas.

Como as escalas de frequência acústica e auditiva são diferentes, os picos se localizam em pontos diferentes nas duas representações, isso apesar de os dois espectros cobrirem a mesma gama de frequências, de 0 a 10.000 Hz. A porção do espectro auditivo que abrange as frequências entre 0 e 1.500 Hz representa praticamente metade da escala, ao passo que a porção correspondente no espectro acústico não chega a 20 por cento. Percebe-se, então, que são as frequências mais baixas que dominam o espectro auditivo.

5.10 REPRESENTAÇÕES AUDITIVAS

Um espectrograma tradicional é um espectrograma acústico, representando as propriedades que ele tem enquanto está sendo transmitido pelo ar, não apresentando uma correspondência muito próxima com as propriedades auditivas do som. Uma forma de retratar as propriedades auditivas de maneira mais próxima de nossa percepção são os chamados espectrogramas auditivos ou cocleagramas. Embora a mesma faixa de frequência seja coberta nos dois diagramas, a que vai de 0 a 10 kHz, as informações sonoras das frequências mais baixas recebem muito mais destaque num espectrograma auditivo do que num acústico.

A escala de Bark, proposta por Zwicker em 1961, é uma escala que representa muito melhor a sensação de altura dos nossos ouvidos do que a escala em hertz, que é algo medido por aparelhos. Ela é usada no espectrograma auditivo (ou cocleograma) da fig. 43, correspondente ao espectrograma da fig. 42, produzidos no praat.

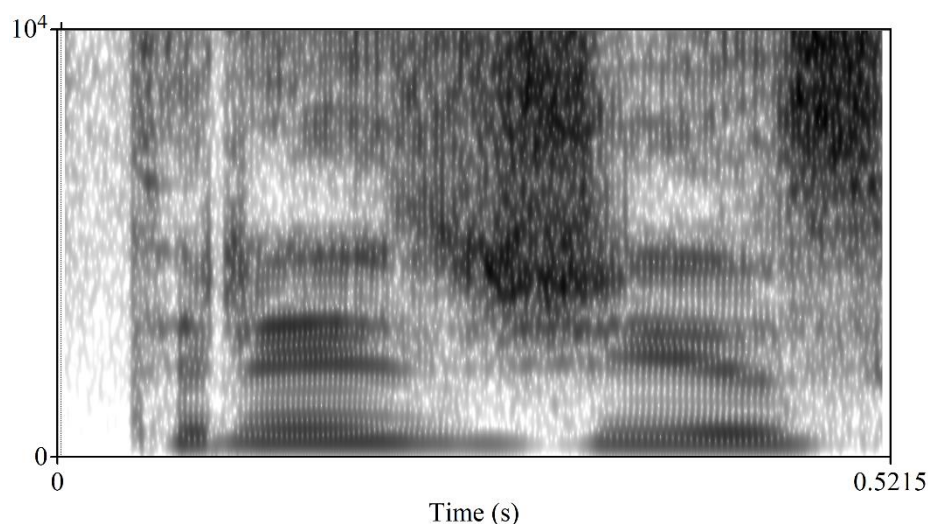


Figura 42. Espectrograma da expressão *três dias*.
 Fonte do arquivo sonoro: Raposo de Medeiros (2020).

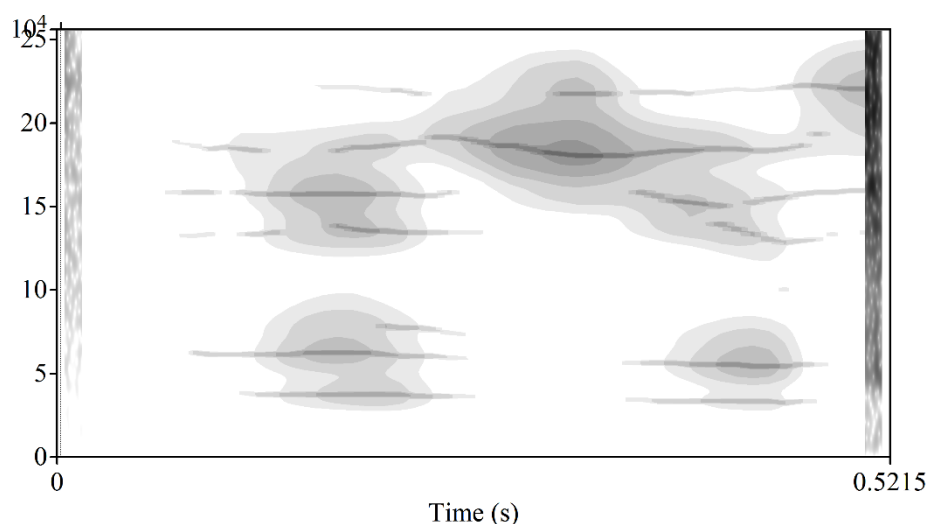


Figura 43. Cocleograma da expressão *três dias*.
 Fonte do arquivo sonoro: Raposo de Medeiros (2020).

Comparando as duas representações gráficas, vemos que os dois primeiros formantes se localizam na região inferior do espectrograma corresponde a cerca de um quarto da extensão vertical, enquanto no cocleograma eles se localizam na metade inferior. Se pensarmos de maneira complementar, o que ocupa 75% do espaço vertical no espectrograma ocupa apenas metade no cocleograma. Isso dá uma ideia do quanto são mais nítidas as diferenças nas frequências mais baixas. A forma de representar a frequência nos dois tipos de espectrograma é a mesma. No entanto, a escala de frequência auditiva dá destaque às frequências mais baixas (como nas vogais) e suas alterações, *warps* (altera, contorce ou redimensiona) o sinal acústico, dando mais destaque a essas frequências.

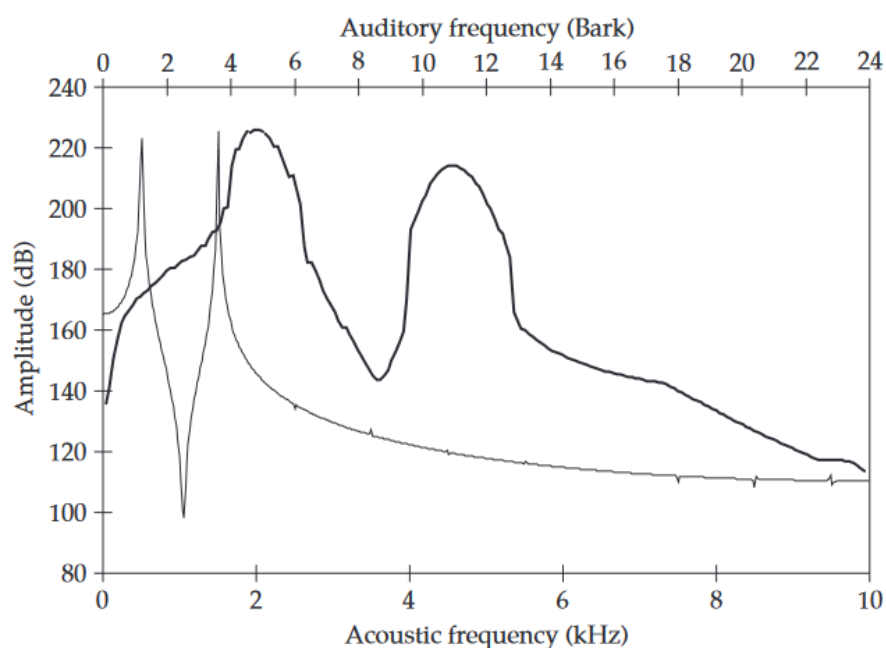


Figure 4.7 Comparison of acoustic (light line) and auditory (heavy line) spectra of a complex wave composed of sine waves at 500 and 1,500 Hz. Both spectra extend from 0 to 10 kHz, although on different frequency scales. The auditory spectrum was calculated from the acoustic spectrum using the model described in Johnson (1989).

Figura 44. Comparação de um espectro acústico e auditivo.

Fonte: Jonhson (2012: 93).

Uma última ilustração dessa diferença de representações é a da figura 44, de Johnson (2012), a qual apresenta a escala de frequência acústica abaixo do gráfico e o espectro acústico de uma onda com ondas senoidais de 500 e 1500 Hz com linhas finas, e a escala de frequência auditiva acima do gráfico e o espectro auditivo dessa onda complexa com linhas mais espessas. A proporção exata no espaço acústico varia de acordo com o modelo empregado. Como vemos, no modelo de Johnson, as frequências mais baixas se destacam ainda mais do que no modelo usado no Praat.

5.11 COARTICULAÇÃO

A descrição mais tradicional dos sons da fala nos dá a impressão que temos uma sequência de fones claramente distintos, nitidamente separados uns dos outros. Nada poderia estar mais distante da realidade. Quando articulamos um determinado fone em posição não inicial nem final num enunciado, estamos carregando ainda algumas propriedades do som precedente e já estamos nos preparando para articular o som seguinte.

Um ótimo exemplo é a palavra *suco*. Ela pode ser representada numa transcrição lata como ['suku]. Essa é uma transcrição correta. Mas ela poderia representar mais detalhes da pronúncia. Se considerada literalmente, a transcrição sugere que em ['suku] temos dois segmentos pronunciados

com arredondamento dos lábios: o [u] da sílaba inicial e o [u] da sílaba final. Poderíamos deduzir, então, que o [s] e o [k] são pronunciados sem arredondamento. Se prestarmos bem atenção, contudo, ou, melhor ainda, se pronunciarmos essa palavra olhando num espelho, veremos que o arredondamento dos lábios acaba abrangendo a palavra inteira. Isso poderia ser retratado numa transcrição estrita como [s^wuk^wu]. O [s] inicial fica arredondado por coarticulação antecipatória. Como o fone seguinte é arredondado, ele acaba ficando arredondado também. Mais motivo ainda tem o [k] para ficar arredondado, já que ele está cercado por vogais arredondadas.

Uma analogia útil é a da escrita. Como estamos mais acostumados a prestar atenção aos detalhes da forma escrita da língua, talvez já tenhamos percebido que algumas letras têm sua forma ligeiramente alterada, dependendo das letras que estão vizinhas a elas, o que é ilustrado com a letra <c> nas duas palavras da figura. Em *casa*, o traçado do <c> começa na parte superior direita, ao passo que na palavra *acaso*, como o <a> que o antecede se une a ele, seu traçado começa na parte inferior esquerda.

casa acaso

Figura 45.

5.12 TIPOS DE VOZ OU FONACÃO

O padrão de vibração das pregas vocais descrito até agora é o mais comum nas línguas em geral, a chamada voz modal. Na produção da voz modal, as pregas vocais fecham a glote em toda a sua extensão e vibram de maneira aproximadamente periódica. Há outras possibilidades, no entanto, cada uma produzida por um tipo de configuração das pregas vocais.

Os estados da glote podem ser distinguidos através de três parâmetros: tensão adutiva, tensão longitudinal e compressão medial. Observemos a figura 45.

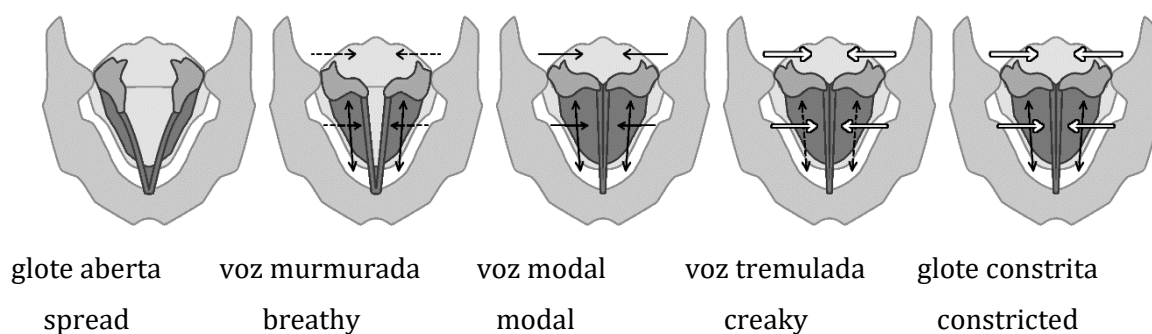


Figura 46. Configurações da glote.

Ilustrações baseadas em Wright et al. (2019),
com acréscimo das setas baseado em Marasek (1997).

Na ilustração vemos cinco configurações. Nelas, as setas inteiriças indicam forças intermediárias, as setas mais espessas, forças maiores e as tracejadas, forças menores. Há duas configurações em que não há vozeamento: na primeira, com as pregas vocais abduzidas ao máximo, o ar passa livremente por elas; na segunda, elas estão cerradas firmemente, e o ar não passa. Nas demais, há algum tipo de fonação.

Na segunda configuração, temos a voz soprosa ou murmurada (*breathy voice*), que se caracteriza por uma abertura glotal maior, tensão longitudinal e compressão medial menores. Há presença de turbulência, frequência menor de vibração e um declínio, declive ou inclinação espectral (a diminuição da amplitude à medida que a frequência aumenta) mais acentuado. Segundo Johnson (2012), esse tipo de voz tem vibração das pregas vocais mas sem muito contato entre elas. Na terceira configuração, temos a voz modal, caracterizada por sua compatibilidade com uma ampla gama de frequências, ausência de ruído turbulento, e um declínio espectral aproximadamente linear. Nela há o fechamento pleno e vibração plena das pregas vocais. Essa é a configuração do vozeamento que encontramos no português. Na quarta configuração, as pregas vocais são mantidas aduzidas, mas de forma frouxa, o que permite vibração. É a voz tremulada, crepitosa, rangida, laringalizada ou glotalizada (*creaky voice, vocal fry*), normalmente produzida em frequências abaixo do limite normal da nossa voz. Na ilustração aparece a configuração da voz tremulada prototípica, caracterizada por tensão adutiva e compressão medial maiores do que na voz modal, mas tensão longitudinal menor. A onda sonora resultante é irregular e tem frequências bem baixas. Já o declínio espectral é menos acentuado que na voz modal.

Na última configuração, por fim, temos as pregas vocais aduzidas de forma constrita, o que impede a passagem do ar e a vibração das pregas vocais. É o que ocorre na parada glotal. (conferir só se o texto bate com as ilustrações e tabela)

Modal: três setas normais

Murmurada: duas setas tracejadas: no meio e longitudinais

Tremulada: seta longitudinal tracejada e as outras duas espessas

Resumo essas características numa tabela.

Configuração	Abertura glotal	Tensão longitudinal	Compressão medial
Glote aberta	grande	não há	não há
Voz murmurada	intermediária	baixa	intermediária
Voz modal	fechada	intermediária	intermediária
Voz tremulada	fechamento frouxo	baixa	alta
Glote constrita	sem abertura	alta	alta

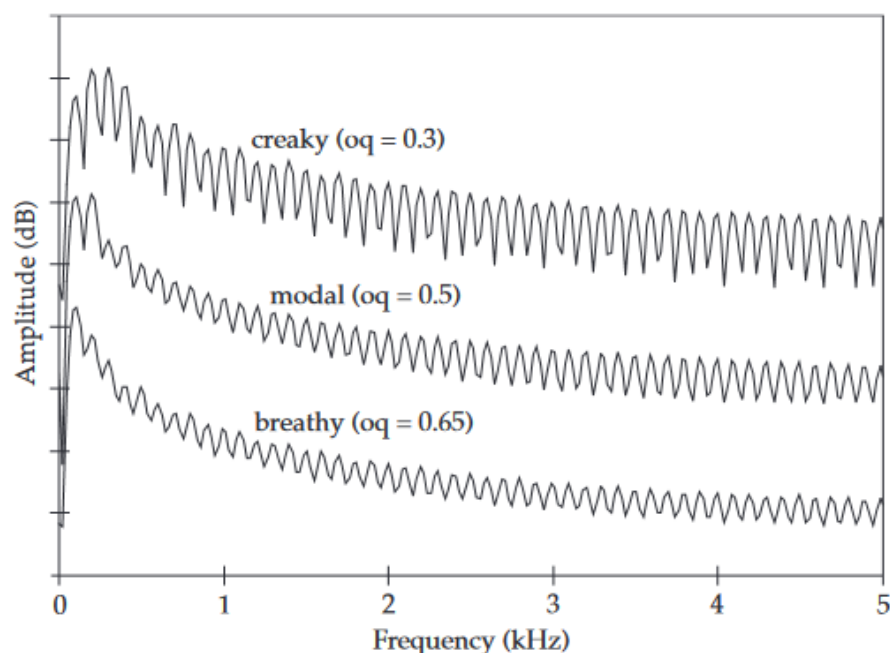


Figure 8.3 Power spectra of the synthetic glottal waveforms shown in figure 8.2. These spectra are offset on the amplitude axis to make them easier to distinguish. The interval between ticks on the vertical axis is 10 dB.

Figura 46. Espectro de três tipos de voz.

Fonte: Jonhson (2012: 172).

REFERÊNCIAS E LEITURAS COMPLEMENTARES

Boersma. On the Need of a Separate Perception Grammar.

Cristófaró Silva, Thaïs; Seara, Izabel; Silva, Adelaide; Rauber, Andreia Schurt; Cantoni, Maria (2019).

Fonética Acústica: Os Sons do Português Brasileiro. São Paulo: Contexto.

Georgeton, Laurianne et al. (2012). Analyse formantique des voyelles orales du français en contexte isolé: à la recherche d'une référence pour les apprenants de FLE. In: *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012*, vol. 1, p. 145-152.

Johnson. Acoustic and Auditory Phonetics.

Kent e Read. The Acoustic Analysis of Speech.

Ladefoged, Peter (coord.) (2007). *The UCLA Phonetics Lab Archive*. Los Angeles, CA: UCLA Department of Linguistics. (<http://archive.phonetics.ucla.edu>)

Ladefoged, Peter; Maddieson, Ian (1996). *The Sounds of the World's Languages*. Londres: Blackwell.

Ladefoged, Peter; Traill, Anthony (1994). Clicks and their accompaniments. *Journal of Phonetics* 22, p. 33-64.

Marasek, Krzysztof (1997). Egg and voice quality (tutorial online). Universität Stuttgart. [Site <https://www2.ims.uni-stuttgart.de/EGG/> visitado em 4/9/2023]

Raposo de Medeiros, B. (2020) LAFALIN Dataset I. [disponível em <https://lafalin.fflch.usp.br/bancodedados-I>. Visitado em 4/9/2023]

Singiresu, Rao (2009). *Vibrações Mecânicas*. 4ª ed. São Paulo: Pearson Prentice Hall.

Wright, Richard; Mansfield, Courtney; Panfili, Laura (2019). Voice quality types and uses in North American English. *Anglophonia* [Online], 27., connection on 04 September 2023. [Artigo visitado em <http://journals.openedition.org/anglophonia/1952> em 4/9/2023. DOI: <https://doi.org/10.4000/anglophonia.1952Zsiga>]

LINKS:

Canal Ubuntu Bridge tem vários vídeos com cliques:

<https://www.youtube.com/user/UbuntuBridge>

Ceceo anterior e lateral (em português):

<https://www.youtube.com/watch?v=2FhLK8JzCOI>

Ceceo lateral (em inglês): <https://www.youtube.com/watch?v=zOpXzFUyF4E>

Cliques em xhosa: <https://www.youtube.com/watch?v=VKOQ1pw3AJQ>

Cliques em zulu: <https://www.youtube.com/watch?v=I6AjEWP-vTY>

Fricativas laterais em zulu:

<https://www.youtube.com/watch?v=K8mbMJ4S-RU>

- sonora em 3:20 e a surda em 8:45

Sons do tlingit com palavras (contém exemplos de oclusivas, fricativas e africadas glotalizadas):

<https://www.youtube.com/watch?v=gr-x6EL39PY>

Tangencial labiodental em mono:

https://scholars.sil.org/kenneth_s_olson/current_research/the_labiodental_flap

Queda da Ponte de Tacoma Narrows: <https://www.youtube.com/watch?v=XggxeuFDaDU>

Explicação de por que a Ponte de Tacoma Narrows caiu (em inglês):

<https://www.youtube.com/watch?v=mXTSnZgrfxM>

Vídeo sobre quebrar uma taça com a voz (em português):

<https://www.youtube.com/watch?v=uTqKbiahKA4>