

# Fisiologia da percepção e processamento da fala

Aniela Improta França

Uma das habilidades linguísticas mais complexas e admiráveis é a percepção e produção dos sons da língua nativa. Não nos damos conta disso a não ser quando verificamos as discrepâncias na fala de estrangeiros com variados níveis de proficiência. Geralmente, mesmo quando se trata de uma pessoa com alto nível de proficiência, a pronúncia raramente é perfeita e impõe traços da língua nativa do falante na língua adicional que ele adquiriu. Fica claro que, como nativos, conhecemos uma grande quantidade de propriedades sonoras e de distinções significativas entre sons de nossa língua (fonemas) que são mais difíceis de perceber por quem não é falante nativo. Por exemplo, os fonemas correspondentes ao último <o> em *vovô* e *vovó* costumam ser desafiadores para os não nativos do português, assim como a nasalização em <não>.

Dois ramos da linguística se ocupam especificamente dos sons da língua: **a fonética e a fonologia**. Vamos estudar sobre os dois neste curso.

O que acontece quando alguém fala conosco? Depois que as ondas sonoras chegam às orelhas (Figura 1), elas são levadas pelo nervo auditivo para o lobo temporal, especificamente para o córtex auditivo primário que recebe indistintamente toda a gama de sons captados do exterior. De buzina à fala. Lá os sons de fala são separados pela frequência de onda e são levados para o lobo temporal posterior onde está a área de Wernicke. A informação que lá chega é pareada com as unidades de representação de som, ou fonemas, que estão guardados (representados) no lobo temporal.

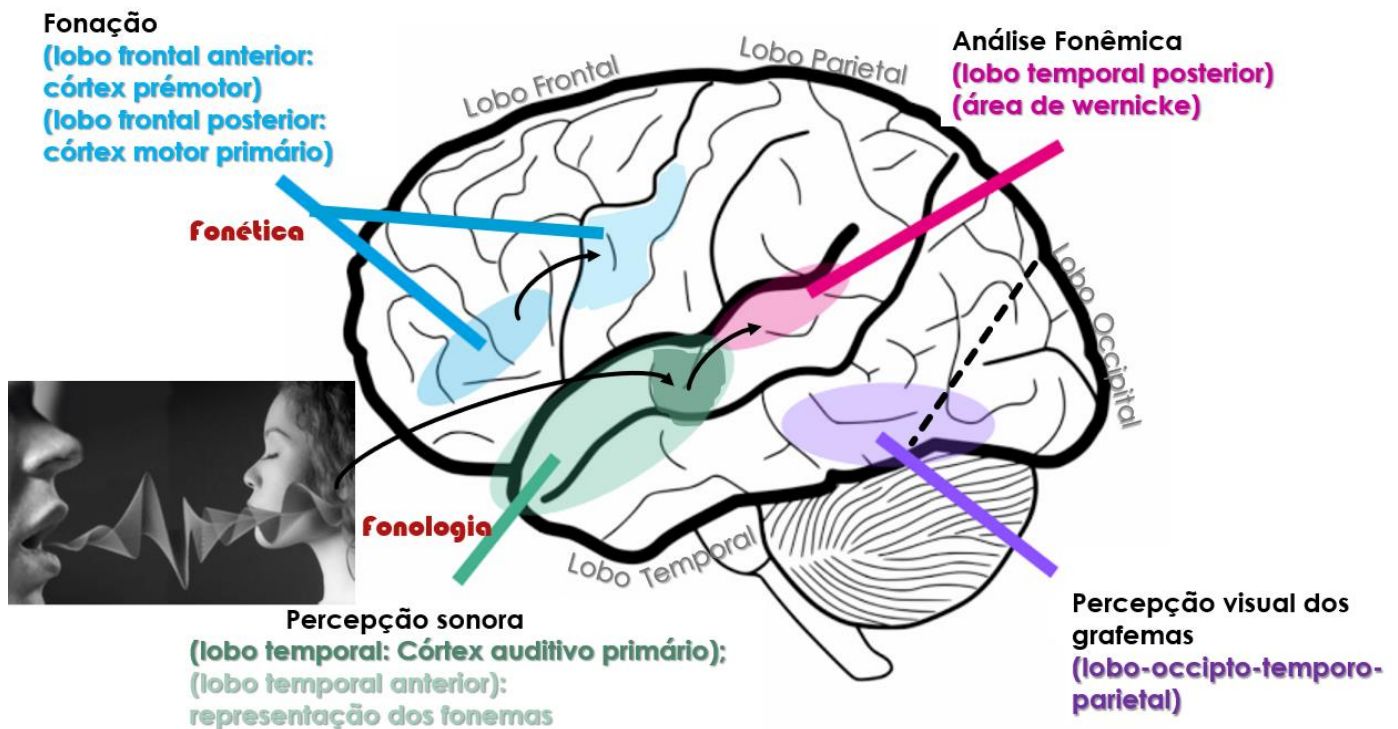


Figura 1: neurofisiologia da percepção sonora

**Este pareamento ou reconhecimento entre o som e as representações mentais é objeto de estudo da fonologia.** A unidade da fonologia, o fonema, geralmente é transcrito entre barras / /.

Temos estocado na mente um número menor de fonemas do que o número de fones (unidades de pronúncia) que existem na totalidade das línguas naturais. Ocorre que as representações mentais que introjetamos durante o período crítico da aquisição de linguagem são as correspondentes aos sons selecionados por nossa língua mãe. Estes fonemas representam sons que perfazem apenas um subconjunto da totalidade de sons da fala pronunciada.

Como se dá a representação dos fonemas? Quando nasce o bebê, as informações sonoras contidas nos *dados primários*<sup>1</sup> vão acionar no cérebro a formação do sistema sonoro da língua-mãe, que conterà todos os sons distintivos na língua, como, por exemplo, /b/ e /p/ que no português são fonemas que nos permitem contrastar facilmente *bote* de *pote* apesar destas duas palavras terem uma sequência de fonemas bem parecida. Constituem o que chamamos de par mínimo: todos os fonemas iguais exceto um. Além disso, guardamos especificações estruturais para licenciar as combinações possíveis entre estes sons, como por exemplo *pote*, OK, mas *pokte*, não.

Durante o período de aquisição representamos todos os fonemas da nossa língua e nos tornamos aptos a pronunciar qualquer palavra dentro deste sistema, incluindo palavras novas e até não palavras que tenham a *fonotática*<sup>2</sup> do português. Só que, para pronunciar, é necessário uma mobilização muscular complexa e precisa. O bebê, por mais que comece a montar o conjunto de representação dos fonemas desde que nasce, demora a ter controle fino do aparelho fonador para implementar a fala. E como vimos a região cortical que representa os fonemas não tem relação com a região que implementa a fala.

Quando estruturamos as cadeias de fonemas, seus traços abstratos servirão de instruções para acionar os órgãos da fala que vão produzir sons no mundo físico. Aí é que entra a área estudada pela fonética.

A fonética estuda a natureza dos sons como realidade física implementada e como eles são produzidos pelo aparelho fonador e percebidos no ar até chegar na orelha do interlocutor. Sua unidade de sonora é o fone, que deve ser transcrito entre colchetes [ ]. Esta ciência investiga os eventos físicos que ocorrem desde que o cérebro recebe ordem de articular o som de uma forma determinada, até que ele chegue ao ouvido do interlocutor.

A fonação ou a produção dos sons da fala mobilizam várias partes do cérebro, em ambos os hemisférios. Mas vamos nos focar neste curso da região do córtex motor primário que fica na divisão entre os lobos frontal e parietal. Nessa porção cortical se dá o controle preciso de todas as fibras musculares que comandam os movimentos voluntários do corpo, incluindo as fibras dos músculos que estruturam os órgãos do aparelho fonador

Você já viu que em termos cognitivos os fonemas, objetos de estudo da fonologia, vem primeiro do que os fones, objetos da fonética. Mas por razões didáticas, geralmente os cursos começam pela fonética, e nós faremos o mesmo.

A fonética desenvolve métodos para a descrição, classificação e transcrição de todos os sons de fala de todas as línguas naturais. Ela faz isso em três especialidades que enfocam aspectos específicos do som:

1. Fonética Articulatória: estuda os sons do ponto de vista fisiológico e articulatório.
2. Fonética Auditiva: estuda a percepção sonora da fala.

---

<sup>1</sup> Dados primários são dados frequentes ao redor do bebê que são absorvidos e interagem com a formação de circuitos cerebrais no cérebro do bebê.

<sup>2</sup> Fonotática é a combinação de fonemas mais lógica e usual em dada língua. O um /f/ em português dificilmente será seguido de um /k/. Então dizemos que a sequência /fk/ não tem boa fonotática no português. Mas se introduzirmos vogais, obteremos boa fonotática, Por exemplo /fiku/.

3. Fonética acústica: estuda as características físicas da propagação do som depois que ele sai da boca do falante.

Observação Importante: Existe uma diferença fundamental entre os conceitos de fone e o de grafema. Observem:

- mesa | sopa – um único grafema, por exemplo “s”, pode representar dois fones distintos: [s], [z];
- uma sequência de dois grafemas, como em chuva, pode ser representada por um único fone [ʃ].
- um único grafema, como “x” em *léxico*, pode representar dois fones [ks]
- o mesmo fone, como [ʃ], pode ser representado por grafemas diversos: *xuxu*, *chuva*
- há grafemas sem fones correspondentes, por exemplo o “h” em *homem*.

É interessante notar que os órgãos físicos envolvidos na produção de fala não dão para ela dedicação exclusiva nem mesmo função principal. Os órgãos que em conjunto formam o Aparelho Fonador têm como funções principais a respiração, a deglutição, a mastigação etc, mas também se prestam para executar a fonação.

Para fins didáticos convencionou-se dividir o Aparelho Fonador em três sistemas: o respiratório ou subglotal, o fonatório ou laringeal, e o articulatorio ou supralaringeal.

**O sistema respiratório ou subglotal** (Figura 2) consiste no diafragma, pulmões com seus lobos, brônquios, árvores brônquicas e traqueia. Além de produzir a respiração, este sistema tem a função secundária de fazer um fole, comandado pelo músculo diafragma, que dá início à corrente de ar que se move através dos brônquios e traquéia, levando o ar para a laringe.

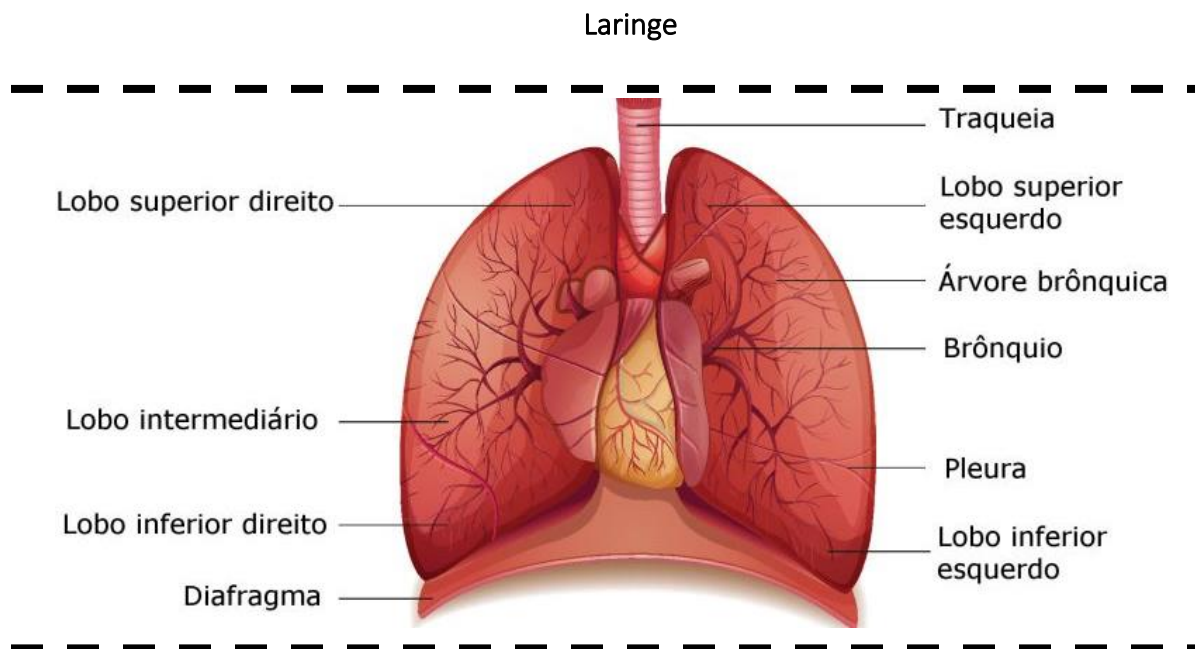


Figura 2: Sistema respiratório ou subglotal

Depois que o sistema respiratório ou subglotal conduz o ar pela traqueia, ele chega ao **Sistema fonatório ou laringeal**. O sistema fonatório é constituído pela laringe (Figura 3), um anel cartilaginoso situado na parte superior da traquéia. Dentro da laringe encontra-se dois órgãos



fundamentais à fonação - as cordas ou pregas vocais e a glote, que é um órgão funcional, só existe quando as pregas vocais se abrem.

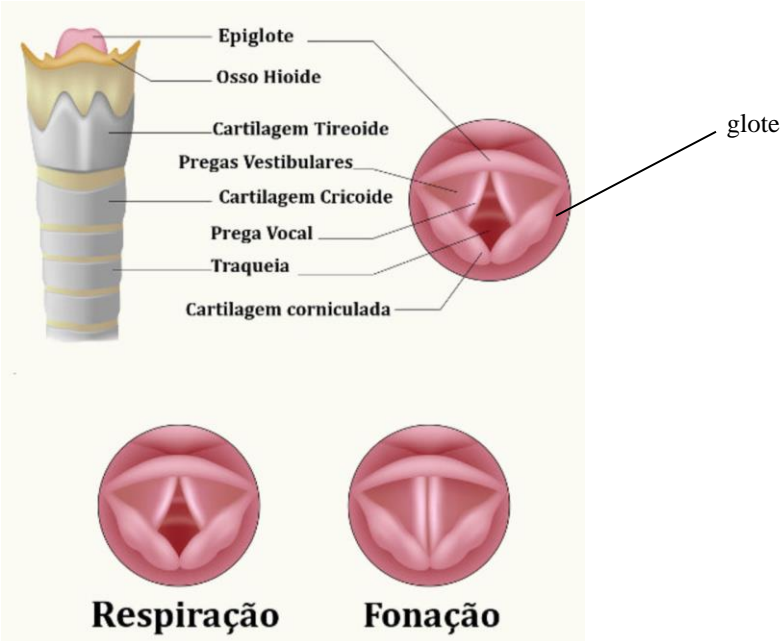


Figura3: Sistema fonatório

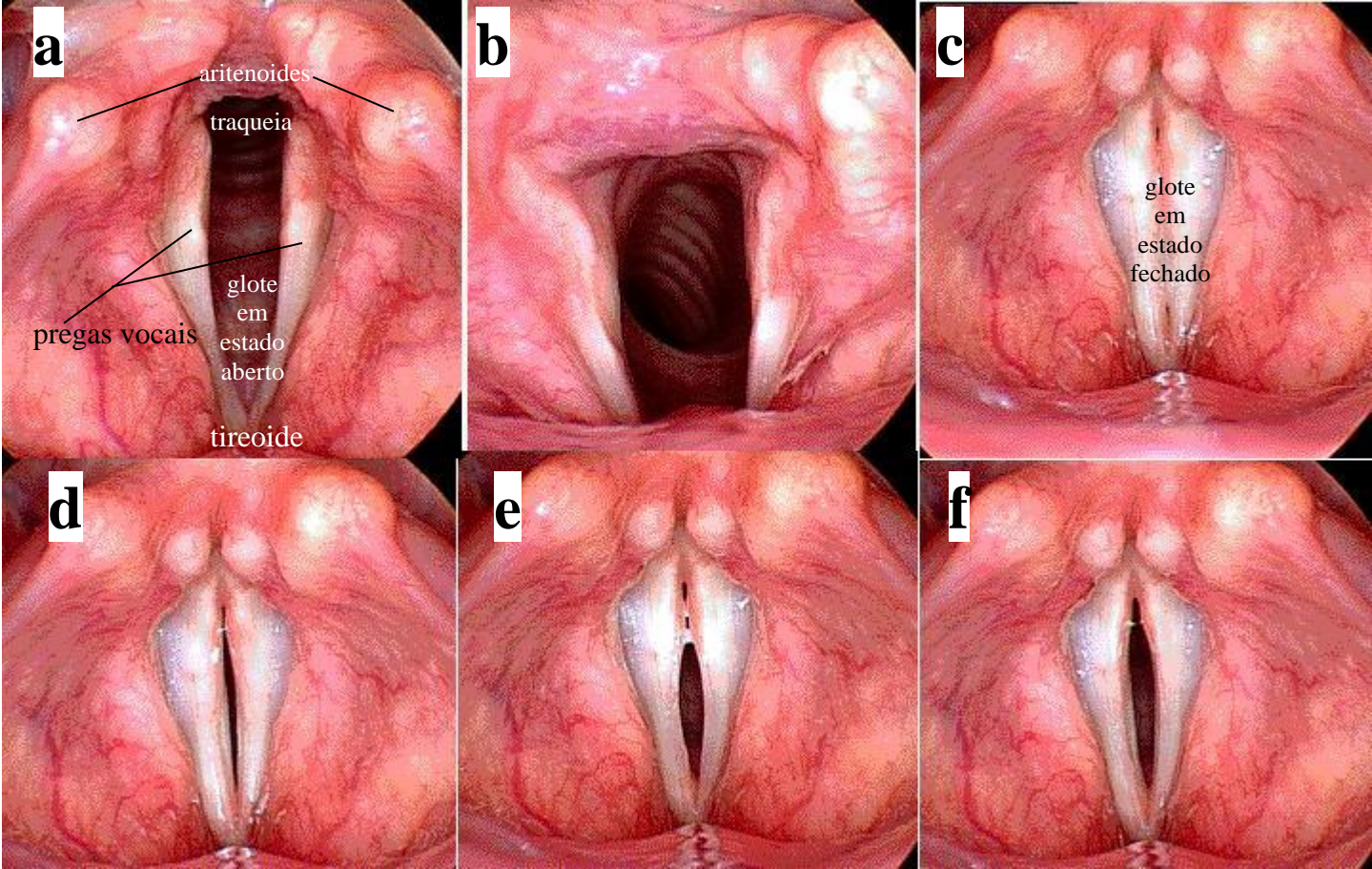


Figura 5: Laringe, pregas vocais, glote, traqueia, cartilagens aritenoide, corniculada e tireoide.

Figura 5a,b,c,d,e,f: Cordas ou pregas vocais, 5b glote aberta mostrando a traqueia na respiração; 5c glote fechada; 5a, 5d, 5e, 5f diferentes níveis de abertura da glote

As pregas vocais (Figura 5) são como dois lábios esticados horizontalmente ao longo da abertura central da laringe, no sentido frente-trás. Elas são altamente flexíveis porque são feitas de um tecido elástico, chamado ligamento, e do músculo estriado tireocricóide. Na frente, as cordas estão presas a tireóide (que são visíveis em homens como o nó no pescoço conhecido com pomo-de-adão). Atrás, as cordas vocais são ligadas às cartilagens aritenóides.

As cartilagens aritenóides são multiarticuladas e são as responsáveis por uma mobilização sutil das pregas vocais tendo como resultado diferentes aberturas que resultam em diferenças de vibração, que por sua vez ressoam de diferentes formas ao seguirem o caminho de expulsão do ar pelo trato vocal. Em repouso, as duas pregas vocais são separadas uma da outra.

Quando elas estão separadas exibem um espaço entre elas chamado de glote. A glote (Figuras 5 e 6) é um órgão funcional porque só existe quando as pregas cais estão afastadas. Com *estado da glote aberto*, o ar passa livremente e as cordas vocais não vibram. Quando isso acontece o *estado da glote é aberto e o som é desvozeado* Mas, em função da mobilidade das aritenóides, as cordas vocais podem se aproximar. Neste caso, para sair, o ar tem de forçar sua passagem fazendo as cordas vibrarem. Quando isso acontece o *estado da glote é fechado e o som é vozeado*.

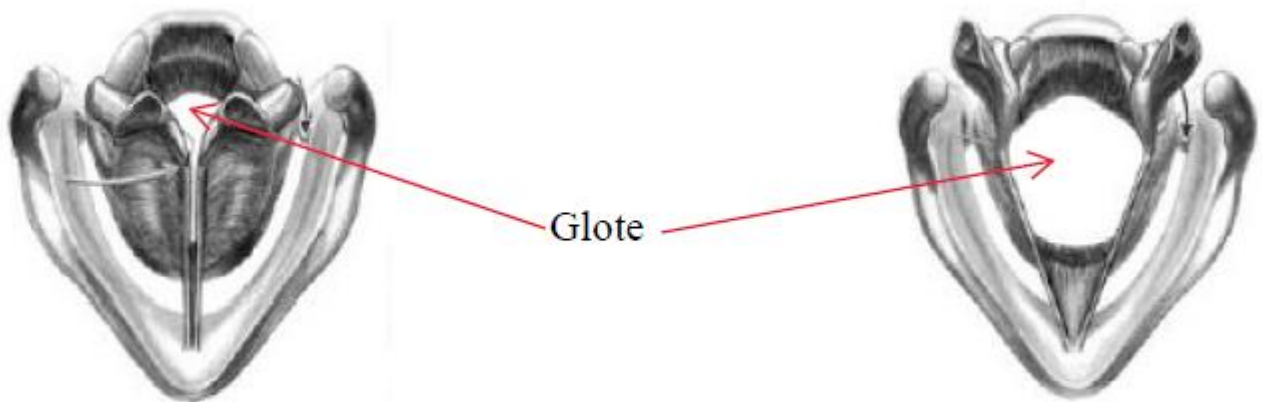


Figura 6: Estrutura laringal, exibindo a epiglote, as pregas vocais e glote, à esquerda em estado fechado e à direita em estado aberto.

Para proteger toda esta estrutura complexa, acima das pregas vocais há uma dobra de tecido muscular estriado formando uma válvula localizada acima da laringe e antes do final da língua na garganta. É a *epiglote*. Ela evita que durante a alimentação entre comida ou líquido pela laringe e traquéia e acabe chegando até aos pulmões, onde estes corpos estranhos causam sérios danos à saúde.

**O Sistema Articulatório ou supralaríngeo** é formado pelas cavidades ou caixas de ressonância oral e nasal e pelos órgãos farínge, língua, palato (dividido em crista alveolar palato duro medial e palato duro final, palato mole e úvula), nariz, dentes e lábios. Mobilização de estruturas neste sistema cria constrições na caixa de ressonância que modifica o fluxo do ar e conseqüentemente o som resultante (Figura 7).



# O APARELHO FONADOR

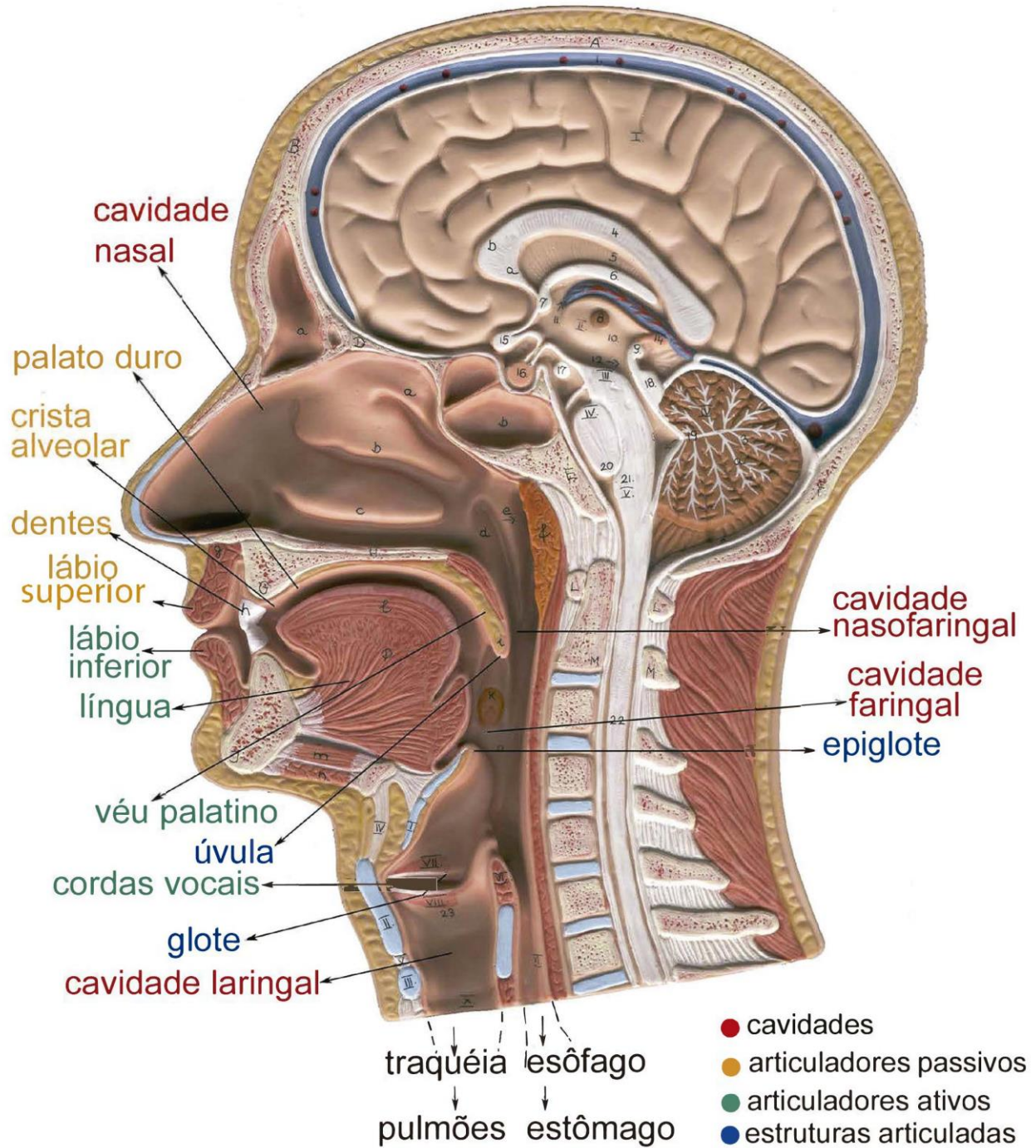


Figura 7: Aparelho Fonador

O caminho para respiração começa no sistema respiratório onde o ar é impelido dos pulmões em direção para fora do corpo (corrente de ar egressiva). Depois de passar pelos brônquios, o ar chega ao sistema fonatório, onde o estado da glote vai determinar o vozeamento do som: estado da glote aberto

deixa o ar passar livremente sem fazer vibrar as cordas vocais, resultando em um ruído de consoante desvozeada:

[p t k f s ...]

Quando o estado da glote é **fechado** as pregas vocais estão mais unidas e esta aproximação cria uma barreira dinâmica ao ar. O impedimento da saída do ar faz as pregas **vibrarem**. O som então é vozeado: [b,d,g, v,z, m, n]

Depois que o ar passa para o sistema suprafaríngeo temos muitos articuladores ativos, estruturas que se movem em direção a outros articuladores passivos com os quais estabelecem articulação.

“Uma grande variedade de sons pode ser produzida pelo aparelho fonador por meio do trabalho conjunto dos três sistemas. Por exemplo, se as cordas vocais estão retesadas, a ponta da língua encosta na região atrás dos dentes (arcada alveolar) e o ar que vem dos pulmões é solto pela boca todo de uma vez (acumula-se na cavidade oral e é solto quando a ponta da língua se descola da arcada alveolar), produzimos o [d]; se tudo se repete, mas as cordas vocais não estão retesadas (e, portanto, não vibram com a passagem do ar que vem dos pulmões), produzimos um [t]. Essa variedade de sons, entretanto, é necessariamente limitada pelas possibilidades oferecidas pelo aparelho fonador. Ou seja, não é possível, por exemplo, produzir um som em que a ponta da língua encoste na glote, o um em que o lábio inferior tenha que cobrir o nariz. Contudo, não havendo problemas no aparelho fonador, uma pessoa quando nasce pode produzir qualquer dos sons encontrados em qualquer língua conhecida sem esforço. Isso é possível porque nossa espécie compartilha as mesmas configurações anatômicas do aparelho fonador, com pequenas diferenças individuais que não afetam o desempenho nessa função”. (Aula 1- Apostila do Alessandro Boechat- UFRJ)

Contudo a partir da adolescência alguns fonemas usados em outras línguas não-nativas podem ser um problema para nós, porque ao aprendermos a mobilização correta para emitir os sons da nossa língua, “esquecemos como fazer outros sons que poderíamos produzir ao nascer, mas que percebemos não serem relevantes para a língua que estamos adquirindo como nativos. Um biólogo italiano muito importante para os estudos linguísticos, Piatelli Palmarini, uma vez proferiu uma frase que ficou célebre a este respeito: A aprendizagem da língua nativa se opera pela desaprendizagem. Vale a pena vocês verem um vídeo mais geral sobre percepção de sons no Youtube (<https://www.youtube.com/watch?v=1b2FbY3z6w0>). Depois vocês devem assistir ao vídeo da grande pesquisadora Janet Werker (<https://www.youtube.com/watch?v=CSMjKDZvNWA&t=27s>) sobre como um bebê desaprende coisas quando fica mais forte na sua própria língua. Os vídeos estão em inglês mas é possível mudar as legendas para português. Eles vão fazer com que vocês lapidem o conhecimento sobre percepção da fala. Não percam.

A relação entre articuladores pode ser explorada através do quadro do Alfabeto Fonético Internacional (Figura 8) que dá uma boa base para a classificação dos fones das línguas naturais:

THE INTERNATIONAL PHONETIC ALPHABET (revised to 1993)

CONSONANTS (PULMONIC)

|                     | Bilabial | Labiodental | Dental | Alveolar | Postalveolar | Retroflex | Palatal | Velar | Uvular | Pharyngeal | Glottal |
|---------------------|----------|-------------|--------|----------|--------------|-----------|---------|-------|--------|------------|---------|
| Oclusiva            | p b      |             |        | t d      |              | ʈ ɖ       | c ɟ     | k ɡ   | q ɢ    |            | ʔ       |
| Nasal               | m        | ɱ           |        | n        |              | ɳ         | ɲ       | ŋ     | ɴ      |            |         |
| Vibrante múltipla   |          |             |        | r        |              |           |         |       | ʀ      |            |         |
| Vibrante            |          |             |        | 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     |
| Lateral fricative   |          |             |        | ɬ ɮ      |              |           |         |       |        |            |         |
| Approximant         |          | ʋ           |        | ɹ        |              | ɻ         | j       | ɰ     |        |            |         |
| Lateral approximant |          |             |        | l        |              | ɭ         | ʎ       | ʟ     |        |            |         |

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

| Clicks | Voiced implosives | Ejectives             |
|--------|-------------------|-----------------------|
| ◌ ɸ    | ɓ                 | ʼ as in:              |
| ◌ ɸ    | ɗ                 | p' Bilabial           |
| ◌ ɸ    | ɟ                 | t' Dental/alveolar    |
| ◌ ɸ    | ɡ                 | k' Velar              |
| ◌ ɸ    | ɠ                 | s' Alveolar fricative |

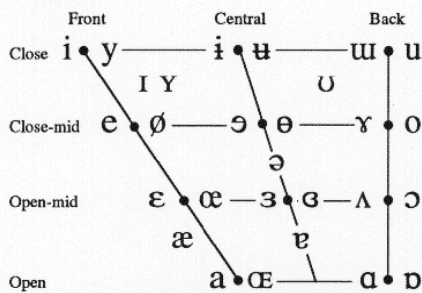
SUPRASEGMENTALS

|    |                              |             |
|----|------------------------------|-------------|
| ˈ  | Primary stress               | fəʊnəˈtɪʃən |
| ˌ  | Secondary stress             |             |
| ː  | Long                         | eː          |
| ˑ  | Half-long                    | eˑ          |
| ◌̥ | Extra-short                  | ɛ̥          |
| ◌̩ | Syllable break               | i.ækt       |
| ◌̯ | Minor (foot) group           |             |
| ◌̰ | Major (intonation) group     |             |
| ◌̱ | Linking (absence of a break) |             |

TONES & WORD ACCENTS

| LEVEL  | CONTOUR          |
|--------|------------------|
| é or ˥ | Extra high       |
| ē      | High             |
| ē      | Mid              |
| è      | Low              |
| ẽ      | Extra low        |
| ↓      | Downstep         |
| ↑      | Upstep           |
| ↗      | Rising           |
| ↘      | Falling          |
| ↗↘     | High rising      |
| ↘↗     | Low rising       |
| ↗↘↗    | Rising-falling   |
| ↗↘↗↘   | Global rise etc. |
| ↘↗↘↗   | Global fall      |

VOWELS



Where symbols appear in pairs, the one to the right represents a rounded vowel.

OTHER SYMBOLS

|   |                                   |   |                            |
|---|-----------------------------------|---|----------------------------|
| ɱ | Voiceless labial-velar fricative  | ɕ ʑ   | Alveolo-palatal fricatives |
| ʋ | Voiced labial-velar approximant   | ɺ   | Alveolar lateral flap      |
| ɰ | Voiced labial-palatal approximant | ɥ   | Simultaneous ʃ and x       |
| ħ | Voiceless epiglottal fricative    | Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary. |                            |
| ʕ | Voiced epiglottal fricative       | k͡p t͡s   |                            |
| ʔ | Epiglottal plosive                |   |                            |

DIACRITICS

Diacritics may be placed above a symbol with a descender, e.g. ɲ̥̊

|    |                 |    |                             |    |                                    |
|----|-----------------|----|-----------------------------|----|------------------------------------|
| ◌̥ | Voiceless       | ◌̊ | Breathy voiced              | ◌̚ | Dental                             |
| ◌̦ | Voiced          | ◌̜ | Creaky voiced               | ◌̜ | Apical                             |
| ◌̨ | Aspirated       | ◌̜ | Linguolabial                | ◌̜ | Laminal                            |
| ◌̙ | More rounded    | ◌̜ | Labialized                  | ◌̜ | Nasalized                          |
| ◌̘ | Less rounded    | ◌̜ | Palatalized                 | ◌̜ | Nasal release                      |
| ◌̙ | Advanced        | ◌̜ | Velarized                   | ◌̜ | Lateral release                    |
| ◌̙ | Retracted       | ◌̜ | Pharyngealized              | ◌̜ | No audible release                 |
| ◌̙ | Centralized     | ◌̜ | Velarized or pharyngealized |    |                                    |
| ◌̙ | Mid-centralized | ◌̜ | Raised                      | ◌̜ | (ɹ̥ = voiced alveolar fricative)   |
| ◌̙ | Syllabic        | ◌̜ | Lowered                     | ◌̜ | (β̥ = voiced bilabial approximant) |
| ◌̙ | Non-syllabic    | ◌̜ | Advanced Tongue Root        | ◌̜ |                                    |
| ◌̙ | Rhoticity       | ◌̜ | Retracted Tongue Root       | ◌̜ |                                    |

Vibrante Multipla  
Vibrante