PRECLINICAL SPEECH SCIENCE

PRECLINICAL SPEECH SCIENCE

Anatomy, Physiology, Acoustics, and Perception

THIRD EDITION

Thomas J. Hixon
Gary Weismer
Jeannette D. Hoit

PLURAL PUBLISHING INC.

PLURAL PUBLISHING
INC.

5521 Ruffin Road
San Diego, CA 92123

e-mail: info@pluralpublishing.com
website: http://www.pluralpublishing.com

Typeset in 10/12 Palatino by Flanagan's Publishing Services, Inc.
Printed in South Korea through Four Colour Print Group

Library of Congress Cataloging-in-Publication Data

Names: Hixon, Thomas J., 1940-2009, author. | Weismer, Gary, author. | Hoit, Jeannette D. (Jeannette Dee), 1954- author.
Title: Preclinical speech science : anatomy, physiology, acoustics, and perception / Thomas J. Hixon, Gary Weismer, Jeannette D. Hoit.
Description: Third edition. | San Diego, CA : Plural Publishing, [2020] | Includes bibliographical references and index.
Identifiers: LCCN 2018018596 | ISBN 9781635500615 (alk. paper) | ISBN 1635500613 (alk. paper)
Subjects: | MESH: Speech--physiology | Speech Perception | Speech Disorders | Respiratory System--anatomy & histology
Classification: LCC QP306 | NLM WV 501 | DDC 612.7/8--dc23
LC record available at https://lccn.loc.gov/2018018596

Contents

Preface

The third edition of *Preclinical Speech Science* is a carefully revised and expanded version of the second edition of the textbook. The revised parts include line-by-line edits of all chapters from the second edition for greater clarity, removal of certain sections (several of which are available as supplementary materials on the textbook companion website, including the scenarios of the previous edition), and addition of new material to chapters from the second edition, including text, figures, and recent references from the research literature.

This new edition also contains three new chapters, including Chapter 6 ("Speech Physiology Measurement and Analysis"), Chapter 13 ("Auditory Anatomy and Physiology"), and Chapter 14 ("Auditory Psychophysics"). Chapter 6 was added to complement Chapter 10 ("Speech Acoustic Measurement and Analysis") and Chapters 13 and 14 were added in response to suggestions made by colleagues and students, that this textbook would benefit from chapter-length material on Hearing Science. With the inclusion of these two chapters on hearing science, perhaps a more accurate title for the textbook would be *Preclinical Speech and Hearing Science*. Because this is the third edition of the text, we have chosen to retain the original title to be consistent with the previous editions.

The Workbook accompanying the third edition of this textbook has also been updated with complete sets of problems and exercises for the three new chapters, and revised exercises for all other chapters. The Workbook is a self-study resource, complete with answers to the problems and exercises.

A PluralPlus companion website also accompanies this new edition of *Preclinical Speech Science*. The website has supplementary text and figures, sound files, study guides, and instructor lecture slides.

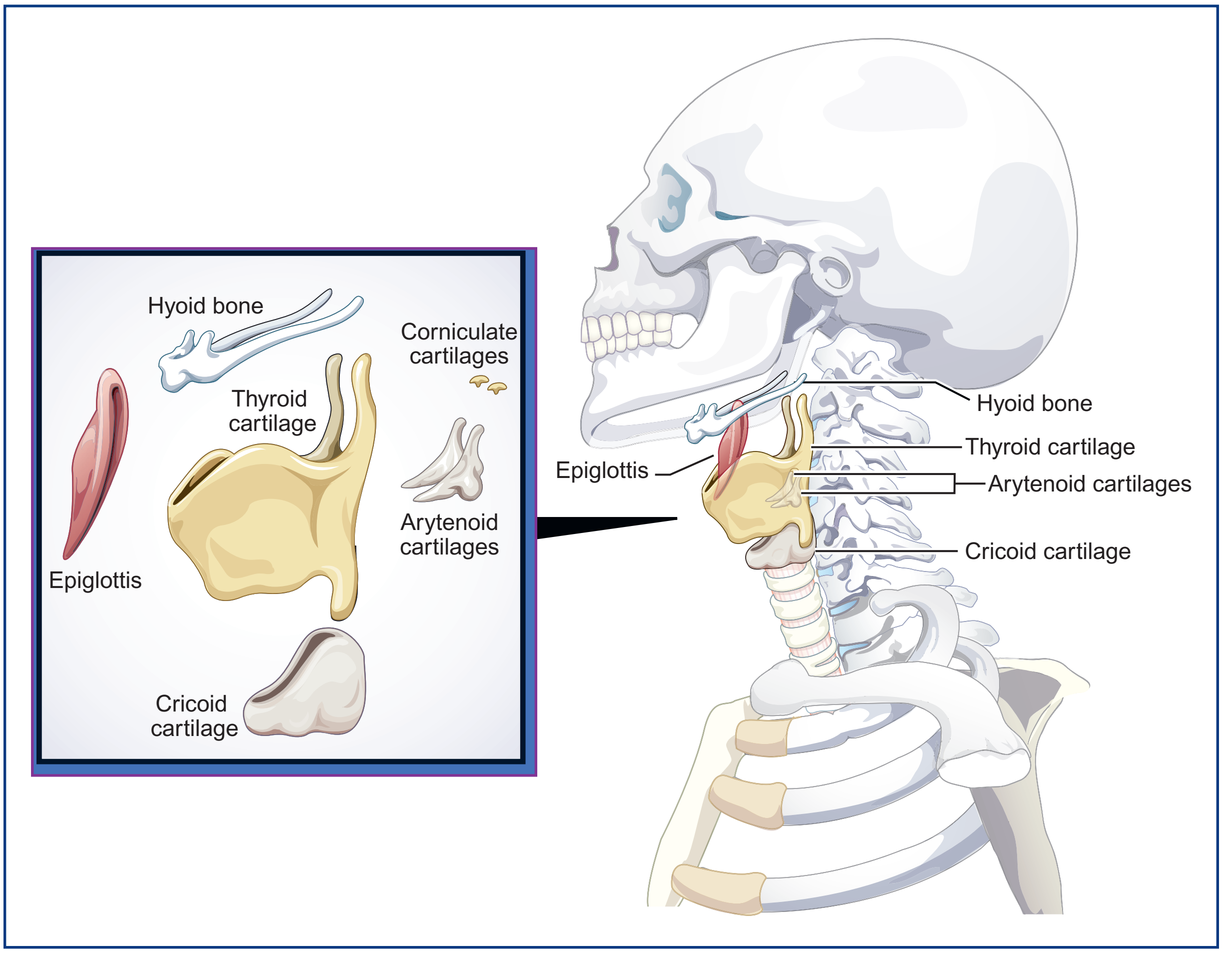


Figure 3–1. Skeletal framework of the laryngeal apparatus. This framework is composed of two paired cartilages (arytenoid cartilages and corniculate cartilages), three unpaired cartilages (thyroid cartilage, cricoid cartilage, and epiglottis), and one bone (hyoid bone).

the thyroid cartilage and diverge widely (more so in women than in men) toward the back. The configuration of the two thyroid laminae resembles the bow of a ship. The line of fusion between the two plates is called the angle of the thyroid. The upper part of the structure contains a prominent V-shaped depression termed the thyroid notch that can be palpated at the front of the neck. This notch is located just above the most forward projection of the cartilage, an outward jutting called the thyroid prominence or Adam's apple.

The back edges of the thyroid laminae extend upward into two long horns, called the superior cornua, and downward into two short horns, called the inferior cornua. The superior cornua are coupled to the hyoid bone. The inferior cornua have facets (areas where other structures join) on their lower inside surfaces that form joints with the cricoid cartilage. The inferior cornua straddle the cricoid cartilage like a pair of legs (see Figure 3–1).

Cricoid Cartilage

The cricoid cartilage forms the lower part of the laryngeal skeleton. It is a ring-shaped structure located above the trachea. As shown in Figure 3–3, the cricoid cartilage has a thick plate at the back, the posterior quadrate lamina, which resembles a signet on a finger ring. A semicircular structure, called the anterior arch, forms the front of the cricoid cartilage and is akin to a band on a finger ring.

Four facets are located on the cricoid cartilage. The lower two facets, one on each side at the same level, are positioned near the junction of the posterior quadrate lamina and anterior arch. Each of these facets articulates with a facet on one of the inferior cornua

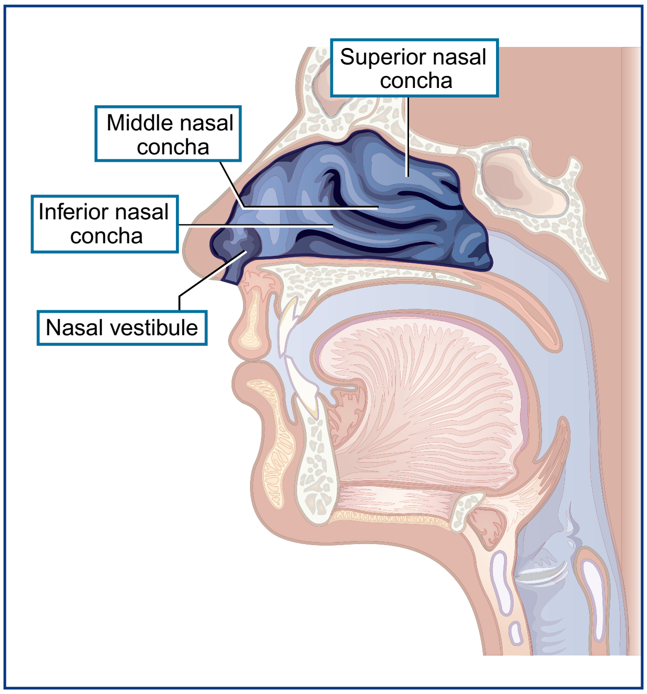


Figure 4–7. Superior, middle, and inferior nasal conchae (also called nasal turbinates). These conchae contain many nooks and crannies and create a large surface area to the inner nose.

provides a large surface area to the inner nose and has a rich blood supply. Near the front of each nasal cavity is the nasal vestibule, a modest dilation just inside the aperture of the anterior naris.

There are four sinuses (hollows) that surround the nasal cavities. Called the paranasal sinuses, they include the maxillary, frontal, ethmoid, and sphenoid sinuses, each located within the bone of corresponding name. Three of these are shown in Figure 4–8. The sphenoid, not pictured, is located behind and above the superior nasal conchae within the sphenoid bone. They are usually air filled but can become liquid filled when infected. Their relevance to speech is primarily related to their effects on the resonance characteristics of the acoustic signal during nasal sound production (see Chapter 9).

Outer Nose

Unlike the other components of the velopharyngeal-nasal apparatus, the outer nose is familiar to everyone. The outer nose is hard to ignore because it is in the center of the face and projects outward and downward conspicuously. The more prominent surface features of the outer nose include the root, bridge, dorsum, apex,

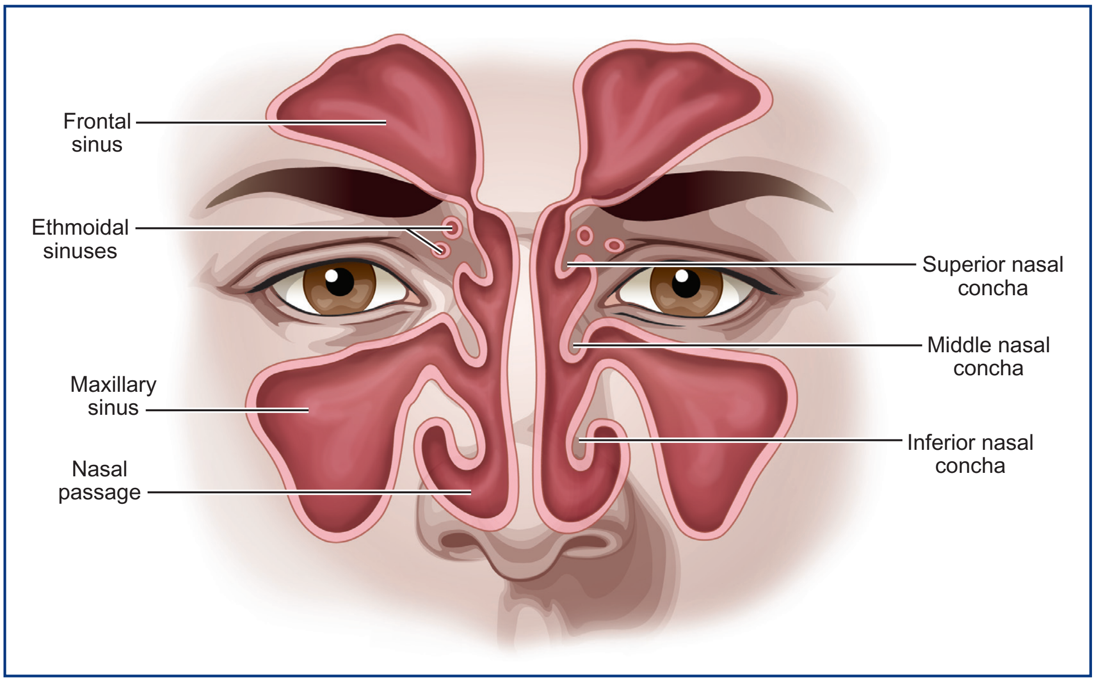


Figure 4–8. The paranasal sinuses. Shown in this figure are the maxillary, frontal, and ethmoid sinuses. Not shown are the paired sphenoid sinuses, which are located behind and above the superior nasal conchae.

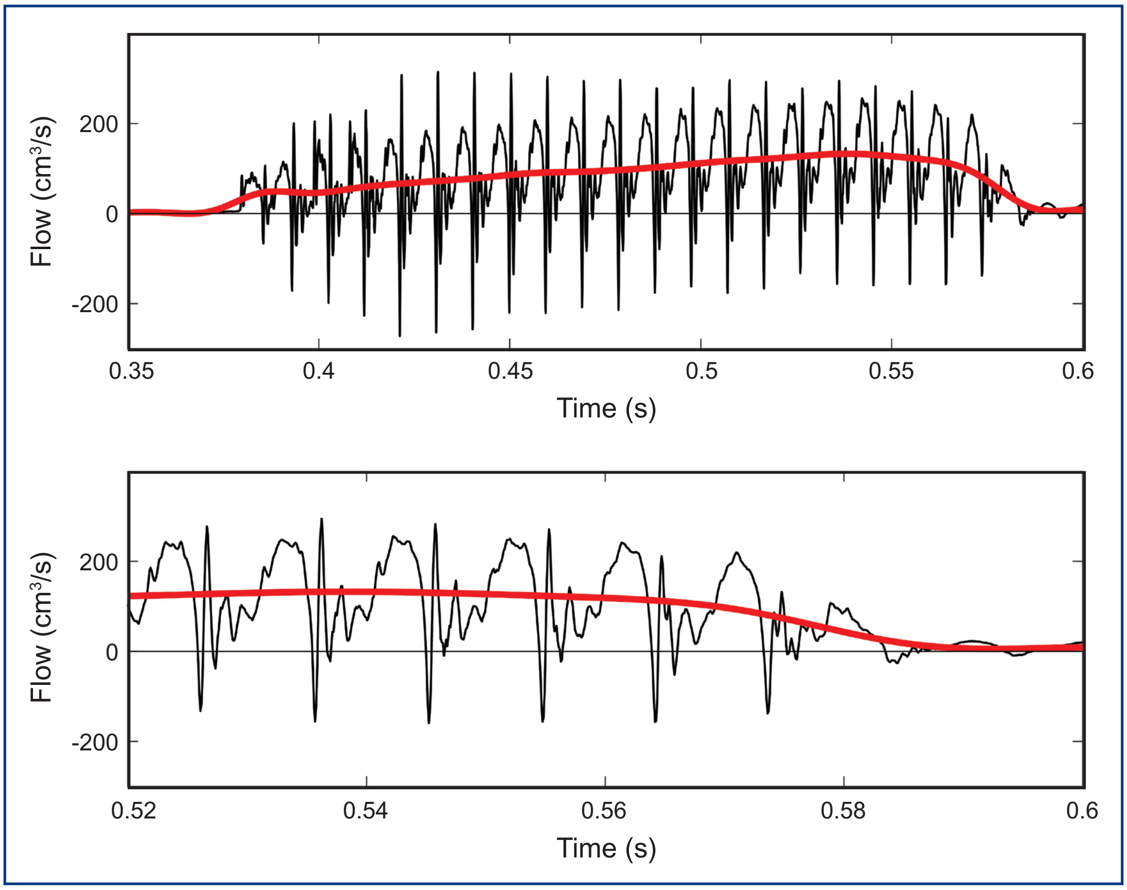


Figure 6–11. Airflow recorded at the airway opening during vowel production. The black tracings show the fast airflow events associated with each cycle of vocal fold vibration. The red tracings represent the average airflow obtained by low-pass filtering the black airflow signal (to filter out high-frequency airflow events). The bottom set of tracings are a zoomed-in image from the upper set of tracings. The fundamental frequency is about 100 Hz (courtesy of Brad Story).

airway-opening airflow to calculate laryngeal airway resistance. As shown in Figure 6–12, measurements are taken at moments that enable estimates to be made of the air pressure difference across the larynx and the airflow through it during vowel productions. Resistance is calculated by dividing the air pressure difference (estimated tracheal air pressure minus estimated pharyngeal air pressure) by the translaryngeal airflow (estimated from the airflow at the airway opening). Resistance values are typically expressed in cmH_2O/LPS (centimeters of water/liters per second) and can range from very low (wide open airway) to infinite (airtight closure of the airway). Such resistance values reflect the degree of opening of the laryngeal airway during voice production (Holmberg, Hillman, & Perkell, 1988, 1999; Leeper & Graves, 1984; Smitheran & Hixon, 1981).

Phonation threshold pressure is another aeromechanical measure that can provide information about laryngeal function, or more specifically, vocal fold function. Phonation threshold pressure is defined as the minimum tracheal pressure required to initiate vocal fold vibration and is understood to reflect the status of the vocal folds (viscosity and thickness) and their distance from one another (glottal width) (Titze, 1988). Although there are invasive ways to measure phonation threshold pressure, the most common way to estimate it is by using the noninvasive approach depicted in Figure 6–7, with the client producing the /p/-vowel syllable strings in the quietest voice possible (Verdolini-Marston, Titze, & Druker, 1990). The lower the peak oral pressures during /p/ productions (estimated tracheal pressure), while still maintaining voicing during the vowel segments, the lower the phonation threshold pressure. And the lower the phonation threshold pressure, the healthier vocal fold function is judged to be. Although this measure is relatively easy to obtain, it is not without its limitations. For example, it is common

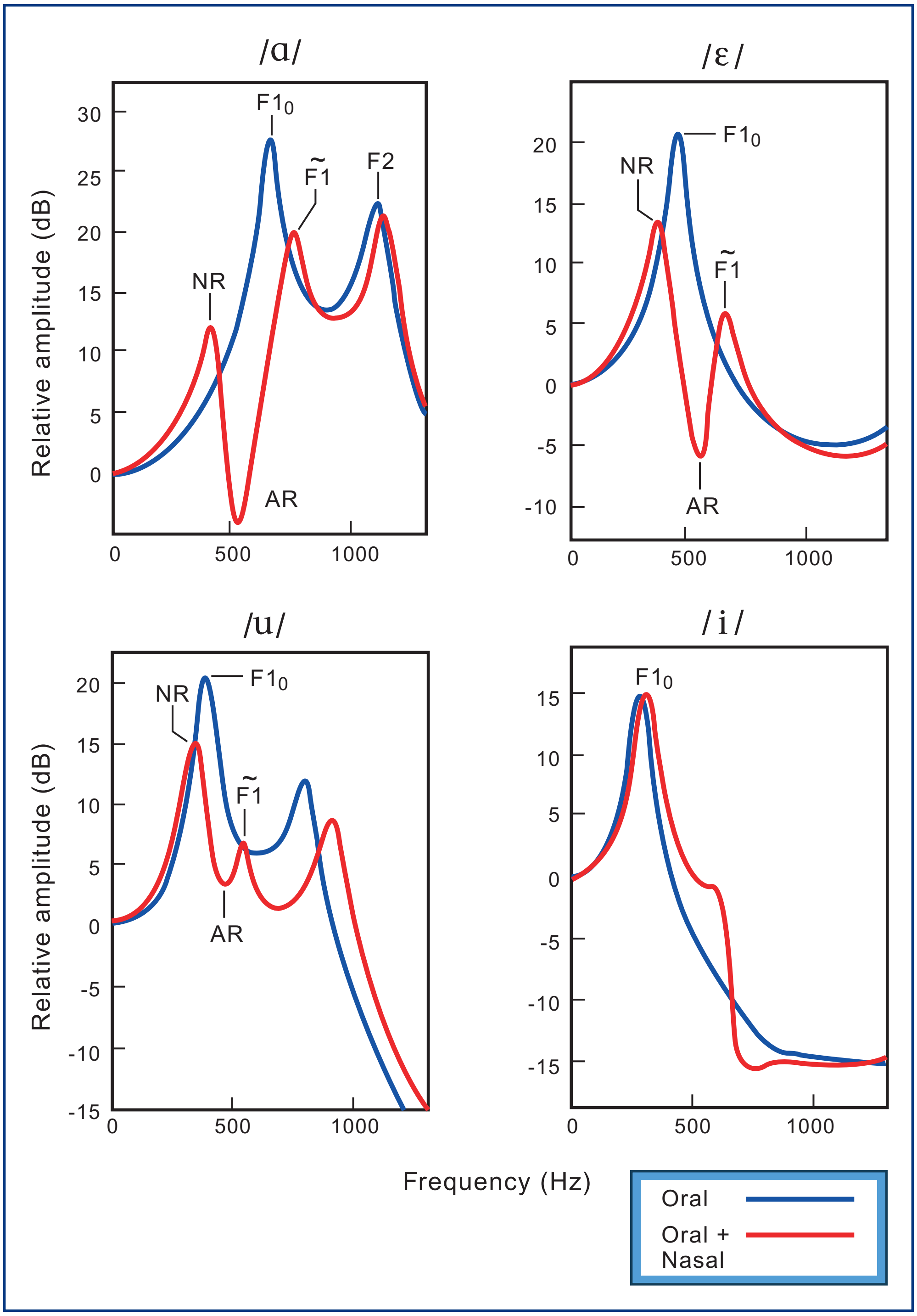


Figure 9–4. Spectra for the vowels /ɑ/, /ɛ/, /u/, and /i/ for non-nasalized (*blue curves*) and nasalized (*red curves*) productions. Frequency is plotted between 0 and 1300 Hz on the *x*-axis and relative amplitude, in dB, is plotted on the ordinate. NR = nasal resonance. AR = antiresonance. $F1_o$ = F1 of non-nasalized vowel. $\tilde{F1}$ = F1 of nasalized vowel. For each vowel except /i/, there is a nasal resonance-antiresonance-F1 pattern in the nasalized spectra. In the case of /i/, the nasal resonance is canceled by the antiresonance because of the small coupling (small velopharyngeal port opening) between the oral and nasal cavities. From "Some acoustical and perceptual correlates of nasal vowels," by K. Stevens, G. Fant, and S. Hawkins in *In Honor of Ilse Lehiste* (p. 246), edited by R. Channon and L. Shockey, 1987, Dordrecht, Netherlands: Foris. Copyright 1987 by Foris. Modified and reproduced with permission.

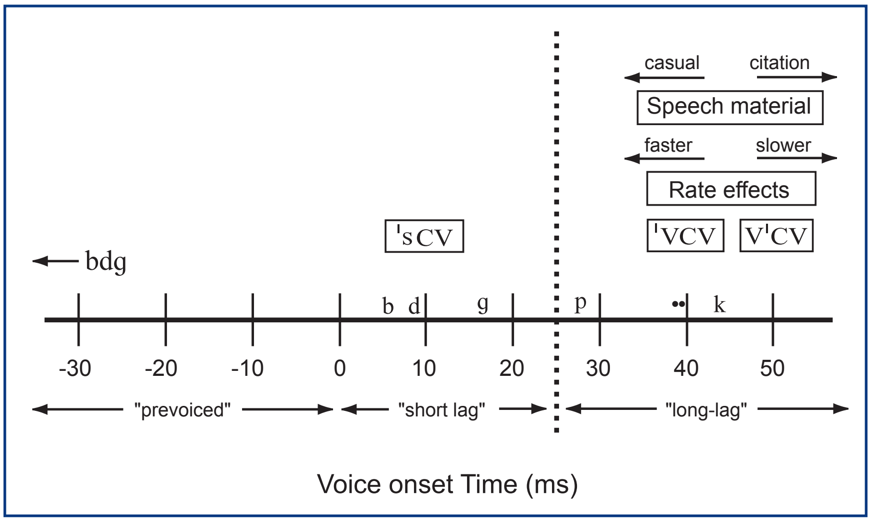


Figure 11–30. Graphic summary of VOT data from English speakers. A VOT continuum ranging from −30 to +50 ms is shown, and effects are indicated by the phonetic symbols and boxes above the continuum line. See text for additional detail.

glottal pulse of the following vowel, negative VOT values represent the time by which glottal pulses within the closure interval *precede* the burst.

Both positive and negative VOT values are common in voiced stop production. As noted above, negative VOTs are associated with stops produced in the utterance-initial position (no speech sounds preceding the stop); intervocalic voiced stops often have glottal pulses during the closure interval, but these are not considered prevoiced. When voiced stops have glottal pulses that are not continuous throughout the closure interval, the VOT is often positive, but very short, as shown in Figure 11–30. The prevoiced, voiced stops reported by Lisker and Abramson all had VOTs more negative than −30 ms, the last negative value on the continuum shown in Figure 11–30.

Figure 11–30 shows a vertical dotted line at 25 ms along the VOT continuum. This line designates a boundary between typical positive VOTs for voiced and voiceless stops. Voiceless stops can be expected to have VOTs exceeding 25 ms (*long-lag* VOTs), whereas voiced stops have VOTs less than 25 ms (*short-lag* VOTs) (Weismer, 2006).

The boxes above the VOT continuum and to the right of the 25 ms boundary identify factors that cause VOT to vary in systematic ways. These boxes are in the long-lag range of the VOT continuum because the effects are most prominent for voiceless stops, with much smaller effects on the short-lag VOTs of voiced stops. VOT is affected by the position of a voiceless stop relative to a stressed vowel. Longer VOTs are measured when the stop precedes, compared with follows, a stressed vowel. The box containing the V'CV frame has been placed to the right (longer VOTs) of the 'VCV box to indicate this effect. In fact, VOTs for voiceless stops in 'VCV frames may be so short as to place them in the short-lag range (Umeda, 1977). The effect of speaking rate on VOT, indicated in Figure 11–30 by the box and arrows immediately above the stress effects, are predictable from the direction of rate change. Slower rates produce longer VOTs for voiceless stops (shown by the arrow pointing to the right), and faster rates produce shorter VOTs (left-pointing arrow) (Kessinger & Blumstein, 1997). The reduction (shortening) of long-lag VOTs at very fast speaking rates is rarely so dramatic as to encroach on the short-lag range (Kessinger & Blumstein, 1997; Summerfield, 1975). Finally, the topmost box indicates that speaking style affects the value of long-lag VOTs. Longer VOTs for voiceless stops are produced in more formal speaking styles, sometimes referred to as citation form or "clear" speech (Krause & Braida, 2004; Smiljanić & Bradlow, 2005). Casual speech styles yield shorter VOTs. The difference between formal and casual speaking styles is likely to involve a difference in speaking rate. Formal speaking styles typically have slower rates than casual styles (Picheny et al., 1986).

A special case of VOT modification for voiceless stops is indicated by the "sCV" box above the short-lag range. "sCV" stands for prestressed s + stop clusters, in words such as "stop," "skate," "speech," "astounding." Voiceless stops in s + stop clusters have short-lag cochlea is oriented in the head as if the tip is pointing along the horizontal axis. The back half of the cochlea is shown in this view. On either side of the center of the slice, two "triplets" of ducts are seen, one triplet at the base (labeled "basal turn" in the figure), the other just above it (labeled "middle turn"). The top triplet of ducts is at the apical turn of the cochlea, at the very tip of which the two outside ducts—the scala vestibuli and scala tympani—are connected. The center "core" section of the cut is called the modiolus (not labeled in Figure 13–16). The turns of the bony cochlea wrap around this center core as they spiral to the apex. The modiolus contains the nerve fibers that innervate the hair cells. It also contains ganglion cells where fibers emerging from the cochlea make their first synapse before continuing to the internal auditory meatus as the auditory part of the auditory-vestibular nerve.

From base to tip, the modiolus sends out two bony shelves toward the outer edges of the spiraling cochlea. These shelves are called the spiral lamina, whose bony extensions serve as the divider between the two outer ducts—the scala vestibuli and scala tympani (labeled only for the basal turns in Figure 13–16). The spiral lamina does not extend to the lateral, bony border of the cochlea. Rather, as described below, membranes extending from the end of the bony lamina to the inside of the lateral border of the cochlea create the third duct sitting between the scala vestibuli and scala tympani. This third duct is called the scala media, or alternately the cochlear duct. All three ducts are filled with fluid.

The second way to appreciate the structure of the cochlea is by studying a zoomed view of the ducts in the cochlea. The zoomed view of the bony cochlea in Figure 13–17 is from its basal turn. From top to bottom the ducts are the scala vestibuli, scala media, and scala tympani. At the beginning of the basal turn of the scala vestibuli, near the section shown in the figure, is the oval window. The termination of the basal turn of the scala tympani is the round window. The two membranes that extend from shelves of the spiral lamina to the outer edge of the cochlea, and enclose the scala media, are called Reissner's membrane (dividing the scala vestibuli from the scala media) and the basilar membrane (dividing the scala tympani from the scala

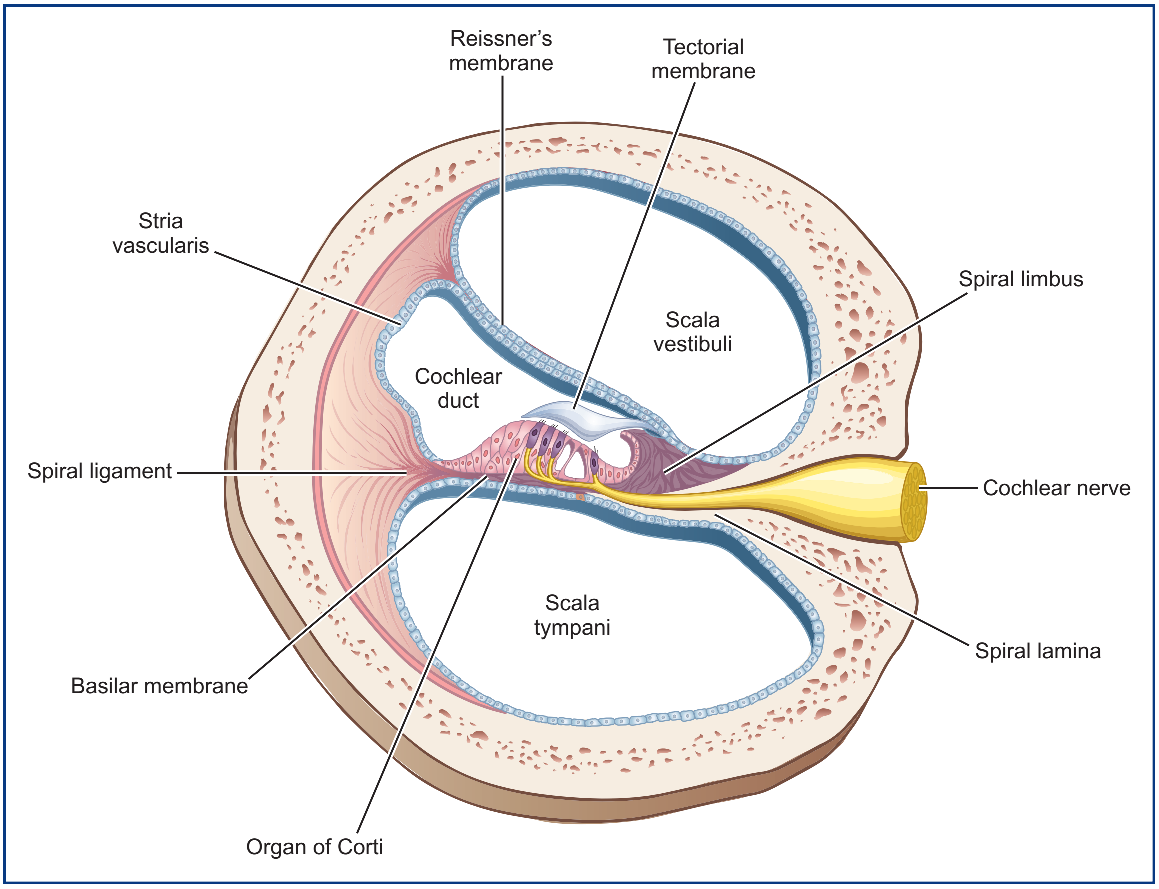


Figure 13–17. Zoom view of cochlear scalae from the basal turn of the cochlea.

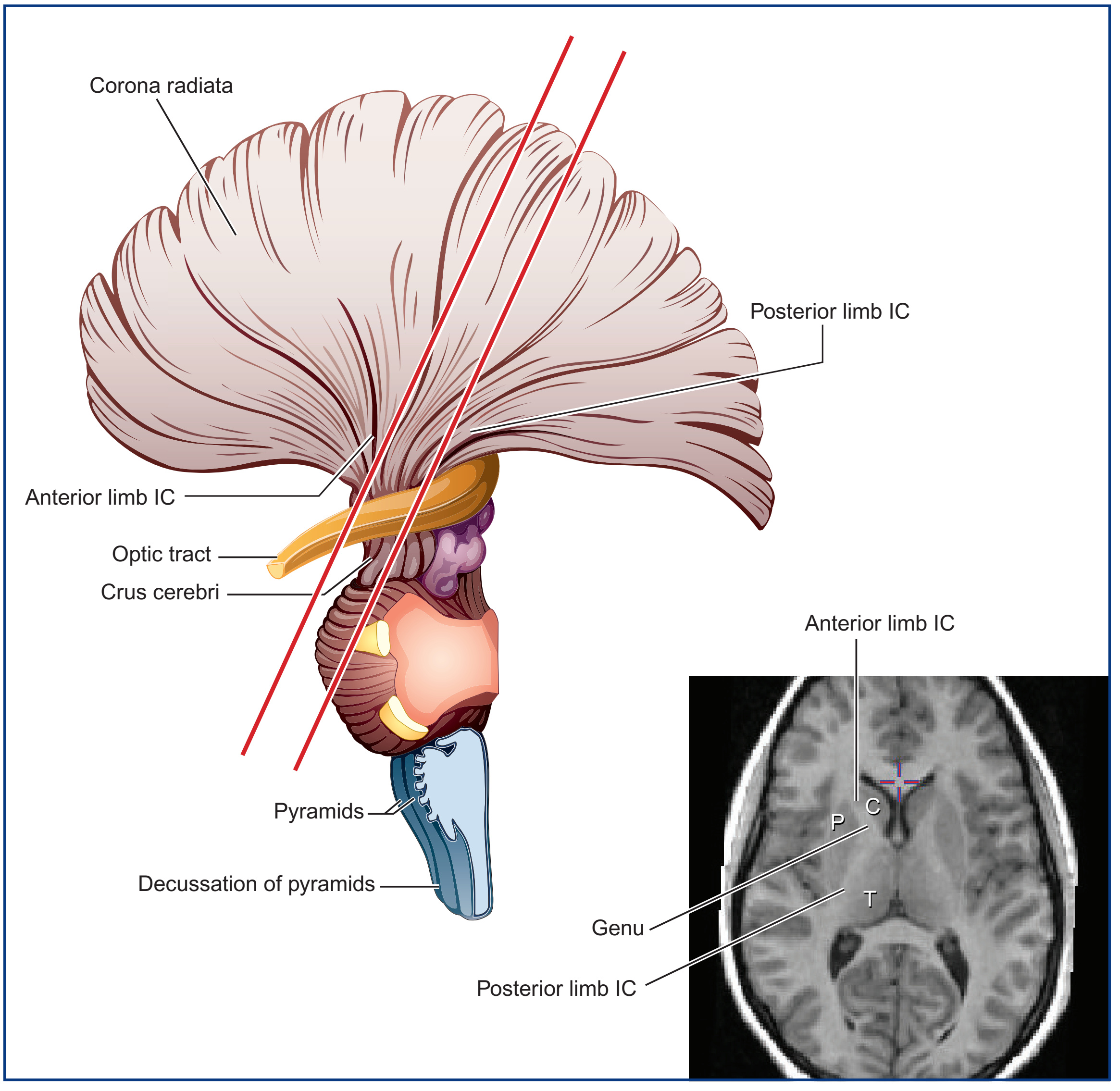


Figure 15–12. *Upper left*, view of fibers of the corona radiata descending in the cerebral hemispheres and gathering into a narrow bundle called the internal capsule (IC), which passes between several subcortical nuclei en route to the brainstem. *Lower right*, horizontal section of cerebral hemispheres showing the "boomerang" shape of the internal capsule. The anterior and posterior limbs plus the genu of the internal capsule are labeled. C = caudate nucleus; P = putamen; T = thalamus.

to reveal the fibers of the corona radiata and internal capsule. Even though the internal capsule is the tightly gathered merger of the many fibers of the corona radiata, the internal capsule has an anterior, middle, and posterior part (IC = internal capsule in Figure 15–12, upper image). The precise location of a coronal slice therefore determines which part of the internal capsule is displayed. Like so many other parts of the brain, the internal capsule is not a random jumble of fibers, but is arranged systematically based on the cortical origin of the fibers. In a horizontal (axial) slice (inset, lower right of Figure 15–12; the anterior part of the brain is toward the top of the image) the internal capsule in each hemisphere has a boomerang shape with the "angle" of the boomerang most medial and the two arms extending away from this angle anterolaterally and posterolaterally. To provide a rough idea of the systematic arrangement of fibers within the internal capsule, most corticobulbar fibers associated with control of facial, jaw, tongue, velopharyngeal, and laryngeal muscles run through a compact bundle close to or within the angle (called the genu) of the internal capsule. Fibers descending to motor neurons in the spinal cord are mostly located in the posterior arm (called the

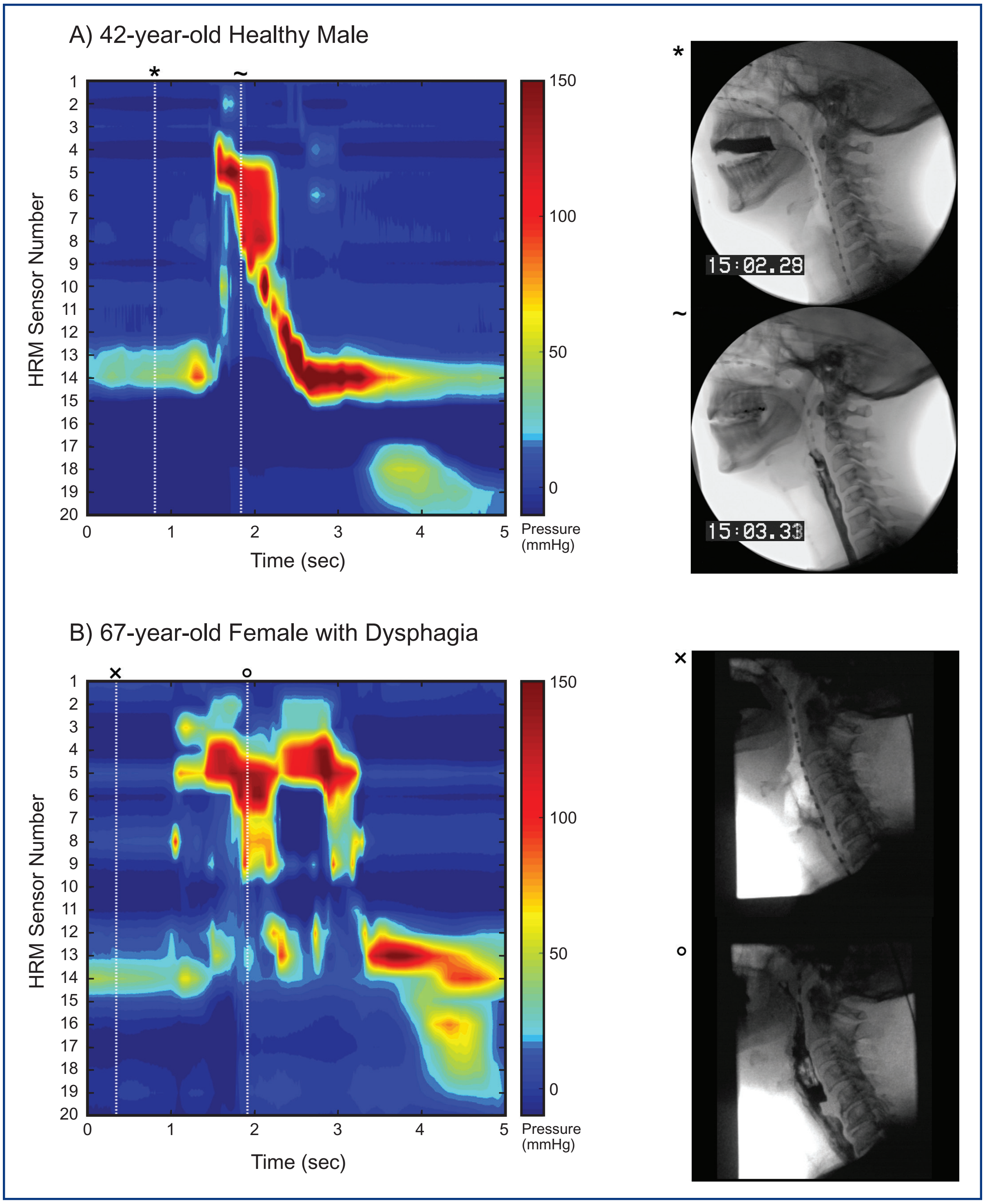


Figure 16–12. Simultaneous videofluoroscopy and pharyngeal high-resolution manometry of a 10 cc thin barium swallow from a 42-year-old healthy man (**A**) and a 67-year-old woman with dysphagia (**B**). High-resolution manometry sensors appear as black rectangles on the videofluoroscopy stills. Videofluoroscopy still images correspond to the time indicated by the vertical lines on the manometry plot with the same symbol at the top. In the data from the healthy man, pressures in the pharynx are low at rest (sensors 4–12: *dark blue*), whereas pressure is higher in the upper esophageal sphincter (sensors 13–14: light *blue/green*). During swallowing, the pharynx constricts, creating high pressures (*orange/red*) at the same time the upper esophageal sphincter relaxes (*dark blue*). The data from the woman with dysphagia reveals that she swallowed twice to clear the bolus, as indicated by the gap in the pressure wave (sensors 10–11: dark blue). Also note the area of elevated pressure in the upper esophageal sphincter (sensor 13: *light blue/green*) during opening. Courtesy of Timothy McCulloch, MD, and Corinne Jones, PhD, CCC-SLP.

Preclinical Speech Science Workbook

Preclinical Speech Science Workbook

THIRD EDITION

Jeannette D. Hoit
Gary Weismer

Plural Publishing Inc.

PLURAL PUBLISHING
INC.

5521 Ruffin Road
San Diego, CA 92123

e-mail: info@pluralpublishing.com
website: http://www.pluralpublishing.com

Typeset in 12/14 Palatino by Flanagan's Publishing Services, Inc.
Printed in South Korea through Four Colour Print Group

Library of Congress Cataloging-in-Publication Data:

ISBN-13: 978-1-63550-063-9
ISBN-10: 1-63550-063-X

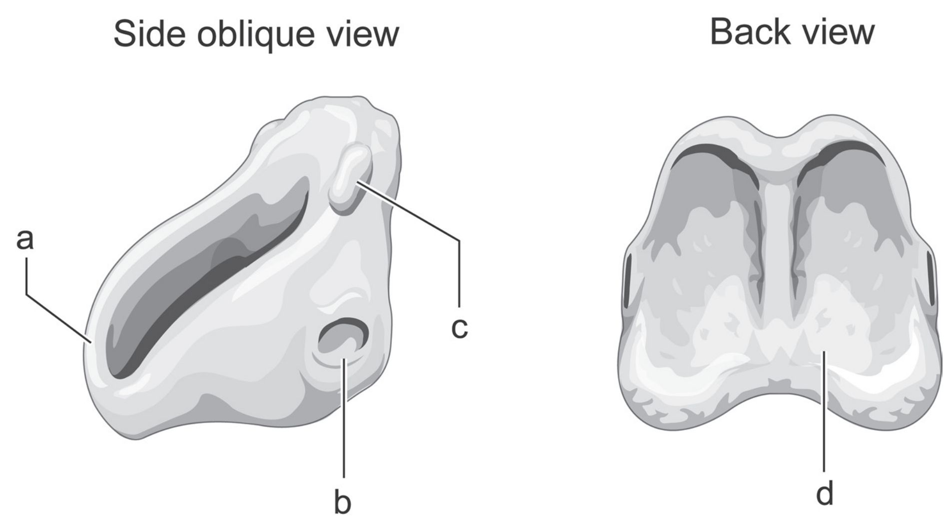
Contents

Preface

The *Preclinical Speech Science Workbook, Third Edition* is a natural companion to the *Preclinical Speech Science, Third Edition* textbook. It has been carefully designed to help students reinforce, integrate, apply, and go beyond the material presented in the textbook.

The workbook contains a wide variety of activities. These include anatomic labeling, measuring physiologic and acoustic data, interpreting graphs, calculating quantitative problems, answering thought questions about material presented in the textbook, and conducting simple experiments (without the use of special equipment). The solutions to all these activities are provided at the back of the workbook; however, we strongly encourage students to work through each activity independently and refer to the solutions only when completely satisfied with their answers. This will provide the best learning experience and will help students make the transition from passive learners to active participants in their development toward becoming speech-language pathologists, audiologists, and clinical scientists.

3–3. Label the parts of the cricoid cartilage indicated in the figures.



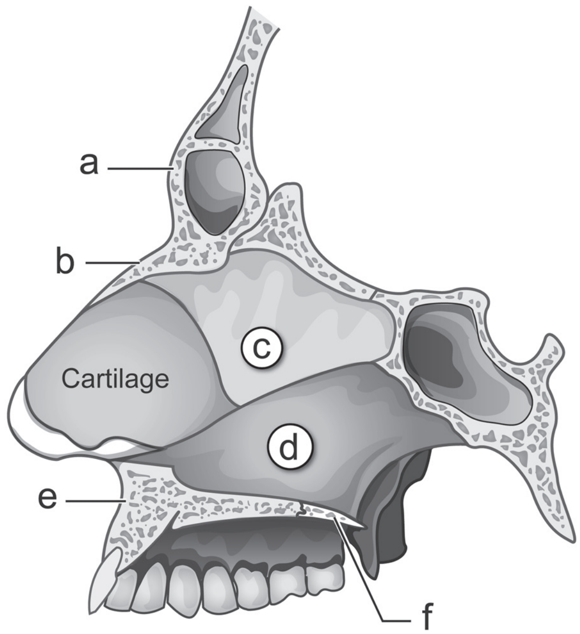
a ______________________________

b ______________________________

c ______________________________

d ______________________________

4–6. Label the bones indicated in the figure.



a ______________________ d ______________________

b ______________________ e ______________________

c ______________________ f ______________________

(a) The two nasal cavities are separated from one another by the nasal ______________, which is made up of [Check one]

______ tendons and ligaments.

______ a matrix of soft tissue.

______ muscle.

______ cartilage and bone.

(b) The hard palate is made up of the ______________ bone and the ______________ bone.

6–21. The velopharyngeal orifice area can be estimated by a method developed by Warren and DuBois (1964; see Figure 6–15 in your textbook). Estimate the velopharyngeal orifice area using the formula and the values given below for oral pressure (P_1, in dynes/cm^2), nasal pressure (P_2, in dynes/cm^2), and nasal flow (in cubic centimeters per second, cc/s). The formula is:

$$\text{Velopharyngeal port area} = \frac{\text{Nasal airflow}}{k\sqrt{\dfrac{2\,(\text{Air pressure differential})}{\text{Density of air}}}}$$

Note that dynes/cm^2 is a unit of measure for pressure that is much smaller than cmH_2O (specifically, 1 $cmH_2O \approx 980$ dynes/cm^2, so 1 dyne/cm$^2 \approx 0.001$ cmH_2O). Also, note that k is a constant that adjusts for the fact that airflow is often turbulent during speech production, rather than laminar (smooth). The suggested value for k is 0.65, density of air $\approx$ .001 (g/cm^3), and the air pressure differential = $P_1 - P_2$.

Velopharyngeal orifice area is expressed in square centimeters (cm^2). Calculate the velopharyngeal orifice area from the oral pressure, nasal pressure, and nasal flow values given below.

Oral Pressure (P_1; dynes/cm^2)	**Nasal pressure (P_2; dynes/cm^2)**	**Nasal Flow (cc/s)**	**Velopharyngeal Orifice Area (cm^2)**
100	80	200	__________
100	0	0	__________
100	20	30	__________

Indicate which of the calculated values above best describes the velopharyngeal orifice area for:

Sustained vowel with normal voice quality __________

Sustained vowel with hypernasal voice quality __________

Sustained /m/ __________

13–30. When is the electrical potential of the hair cells of the organ of Corti and of the crista ampullaris changed?

____

____

____

13–31. The parts of the vestibular system that sense position of the head in the front-to-back and side-to-side dimensions are the ________________ and ________________, respectively. These two structures are part of the organ called the ________________.

13–32. The core of the cochlea is called the ________________, which contains ________________ originating at the base of the hair cells as well as the group of cell bodies called the ________________.

13–33. The membranes that separate the three cochlear ducts are ________________ and ________________.

13–34. In three sentences or less, describe the organ of Corti.

____

____

____

____