Chapter 14

Perceiving Speech

Icebreaker: Muffled Speech Simulation

1. The class will be broken up into groups of students.
2. Each group will select a designated speaker.
3. The speaker will say a sentence aloud, repeating it twice.
4. The speaker will cover their mouth with their hand or sleeve and repeat the same sentence twice.
5. Afterwards students will discuss the auditory differences between the clear speech and the muffled speech and how the muffled speech affected their understanding of what was said.

Chapter Overview (1 of 2)

14.01 Describe how the acoustic signal is created by the action of articulation and is represented by phonemes.

14.02 Understand the processes responsible for variability in the acoustic signal.

14.03 Describe the motor theory of speech perception and evidence for and against the theory.

14.04 Describe the multiple sources of information for speech perception.

Chapter Overview (2 of 2)

14.05 Understand how people perceive degraded speech.

14.06 Describe how research involving brain damage and neural recording has contributed to our understanding of how speech is processed by the brain.

14.07 Understand how cochlear implants work and how they have been used in children.

14.08 Describe infant-directed speech and how it affects infants.

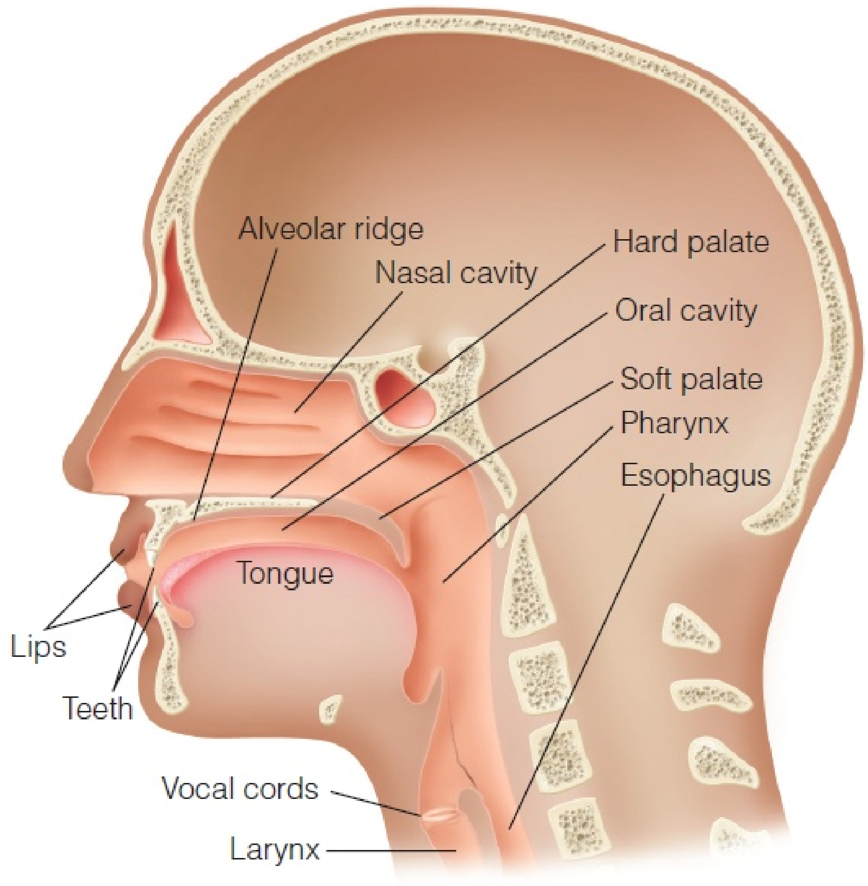
The Acoustic Signal (1 of 4)

Produced by air that is pushed up from the lungs through the vocal cords and into the vocal tract

Vowels are produced by vibration of the vocal cords and changes in the shape of the vocal tract by moving the articulators.

- These changes in shape cause changes in the resonant frequency and produce peaks in pressure at a number of frequencies, called formants.

The Acoustic Signal (2 of 4)



The Acoustic Signal (3 of 4)

The first formant has the lowest frequency, the second has the next highest, etc.

Sound spectrograms show the changes in frequency and intensity for speech.

Consonants are produced by a constriction of the vocal tract.

Formant transitions: rapid changes in frequency preceding or following consonants

The Acoustic Signal (4 of 4)

The first formant has the lowest frequency, the second has the next highest, etc.

Sound spectrograms show the changes in frequency and intensity for speech.

Consonants are produced by a constriction of the vocal tract.

Formant transitions: rapid changes in frequency preceding or following consonants

Basic Units of Speech

Phoneme: the smallest unit of speech that changes meaning of a word.

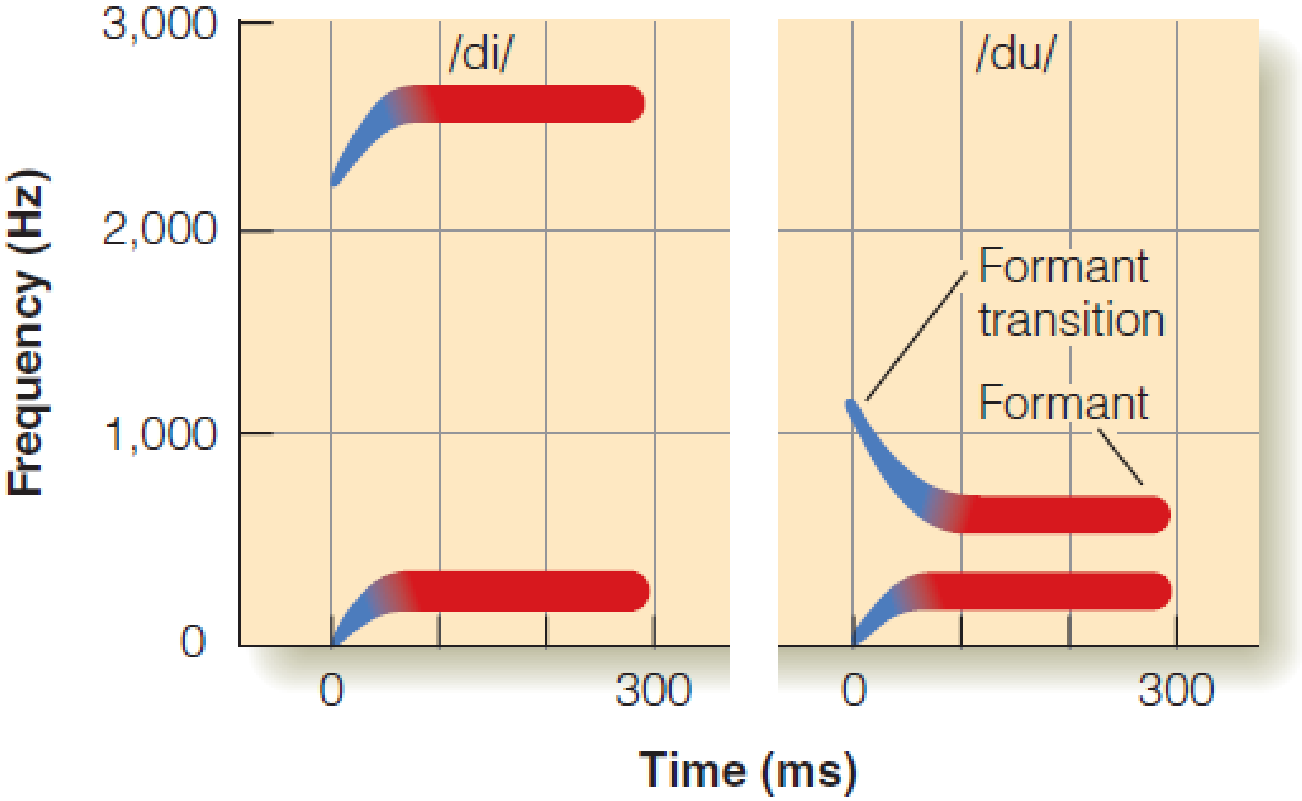
- In English, there are 47 phonemes:
  - 13 major vowel sounds
  - 24 major consonant sounds
- Number of phonemes in other languages varies—11 in Hawaiian and 60 in some African dialects.

The Variability of the Acoustic Signal (1 of 3)

The variability problem: there is no simple correspondence between the acoustic signal and individual phonemes.

- Variability comes from a phoneme’s context.
- Coarticulation: overlap between articulation of neighboring phonemes also causes variation.

The Variability of the Acoustic Signal (2 of 3)



The Variability of the Acoustic Signal (3 of 2)

Variability from different speakers

- Speakers differ in pitch, accent, speed in speaking, and pronunciation.
- This acoustic signal must be transformed into familiar words.

People perceive speech easily despite the variability problems due to perceptual constancy.

Categorical Perception (1 of 3)

Occurs when a wide range of acoustic cues results in the perception of a limited number of sound categories.

- An example of this comes from experiments on voice onset time (VOT): time delay between when a sound starts and when voicing begins
  - Stimuli are /da/ (VOT of 17ms) and /ta/ (VOT of 91ms)

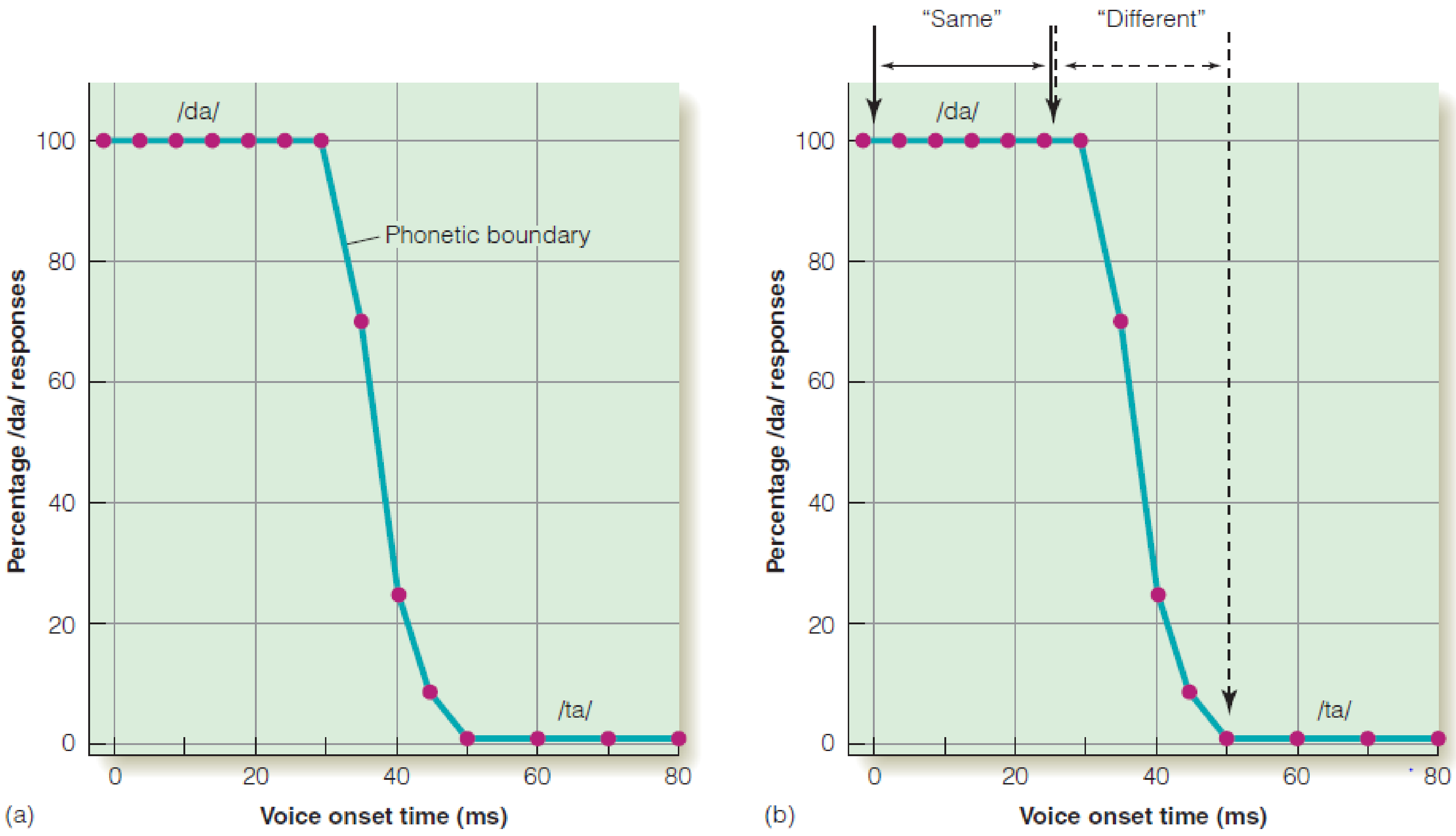
Categorical Perception (2 of 3)

Computers were used to create stimuli with a range of VOTs from long to short.

Listeners do not hear the incremental changes, instead they hear a sudden change from /da/ to /ta/ at the phonetic boundary.

Thus, we experience perceptual constancy for the phonemes within a given range of VOT.

Categorical Perception (3 of 3)

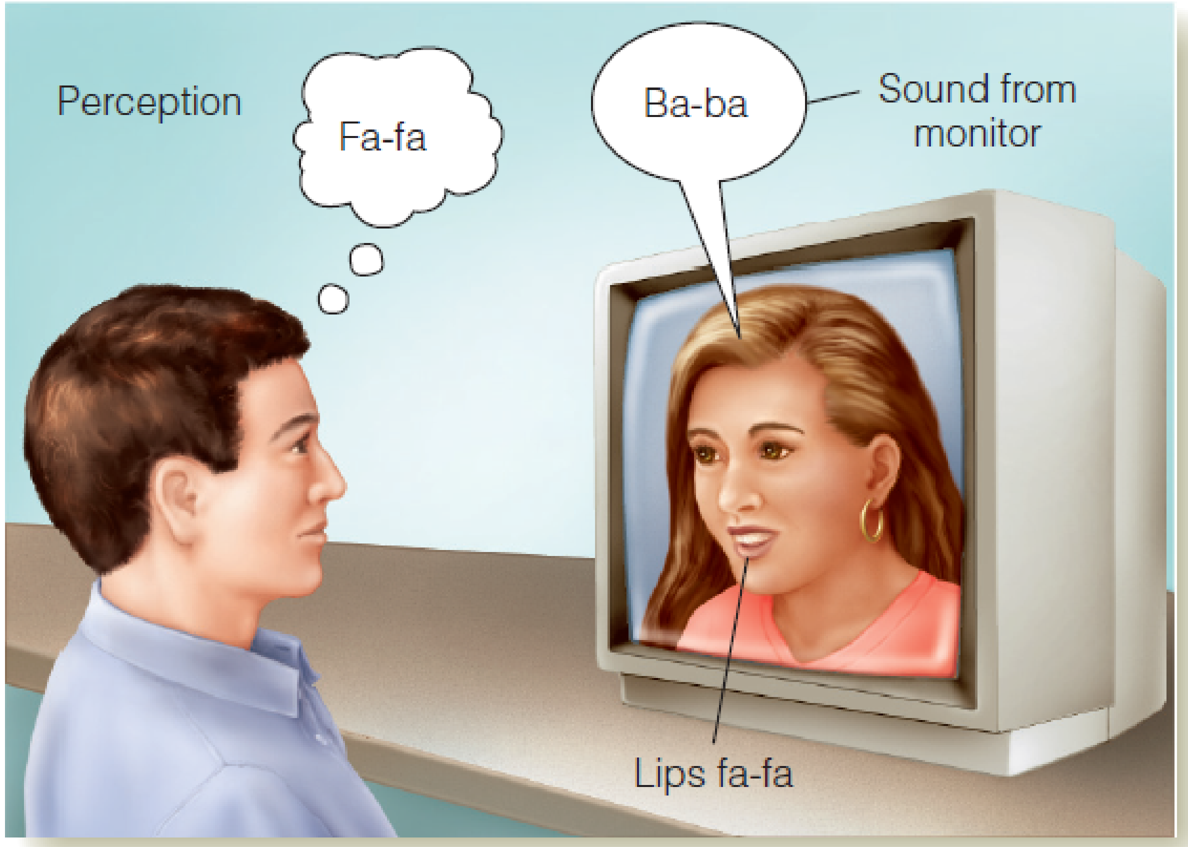


Information Provided by the Face (1 of 3)

Auditory-visual speech perception

- The McGurk effect
  - Visual stimulus shows a speaker saying “ga-ga”.
  - Auditory stimulus has a speaker saying “ba-ba”.
  - Observer watching and listening hears “da-da”, which is the midpoint between “ga” and “ba”.
  - Observer with eyes closed will hear “ba”.

Information Provided by the Face (2 of 3)



Information Provided by the Face (3 of 3)

The link between vision and speech has a physiological basis.

- Calvert et al. showed that the same brain areas are activated for lip reading and speech perception.

Information From Our Knowledge of Language (1 of 2)

Experiment by Rubin et al.

- Short words (sin, bat, and leg) and short non-words (jum, baf, and teg) were presented to listeners.
- The task was to press a button as quickly as possible when they heard a target phoneme.
- On average, listeners were faster with words (580 ms) than non-words (631 ms).

Information From Our Knowledge of Language (2 of 2)

Experiment by Warren

- Listeners heard a sentence that had a phoneme covered by a cough.
- The task was to state where in the sentence the cough occurred.
- Listeners could not correctly identify the position and they also did not notice that a phoneme was missing:- called the phonemic restoration effect.

Perceiving Words in Sentences (1 of 2)

Experiment by Miller and Isard

- Stimuli were three types of sentences:
  - Normal grammatical sentences
  - Anomalous sentences that were grammatical
  - Ungrammatical strings of words
- Listeners were to shadow (repeat aloud) the sentences as they heard them through headphones.

Perceiving Words in Sentences (2 of 2)

Results:

- 89% accurate with normal sentences
- 79% accurate for anomalous sentences
- 56% accurate for ungrammatical word strings
- Differences were even larger if background noise was present.

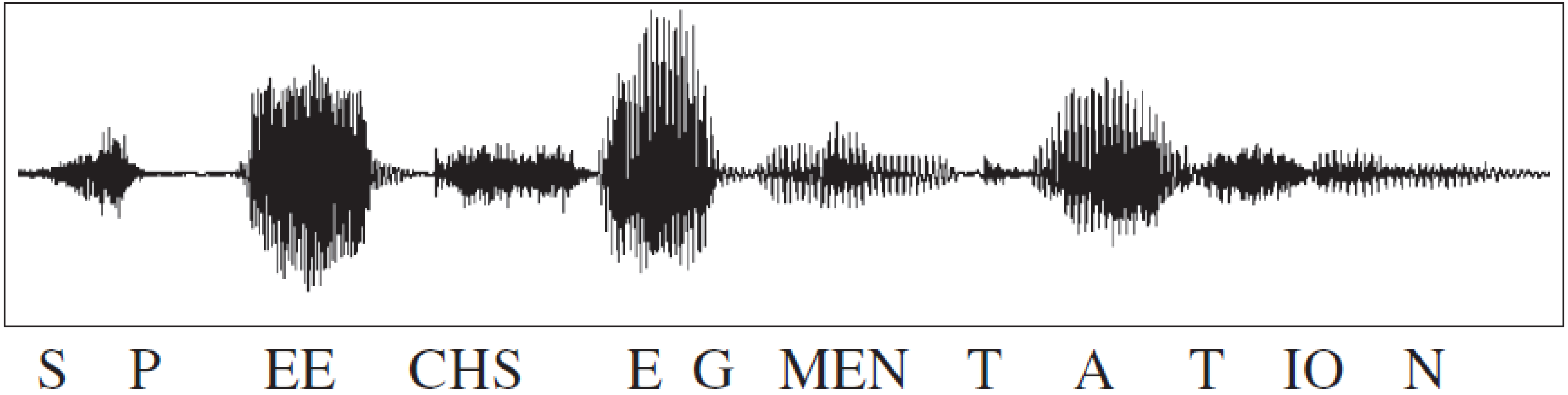
Perceiving Breaks between Sequences of Words (1 of 8)

The segmentation problem: there are no physical breaks in the continuous acoustic signal.

Top-down processing, including knowledge a listener has about a language, affects perception of the incoming speech stimulus.

Segmentation is affected by context, meaning, and our knowledge of word structure.

Perceiving Breaks between Sequences of Words (2 of 8)



Perceiving Breaks between Sequences of Words (3 of 8)

Knowledge of word structure

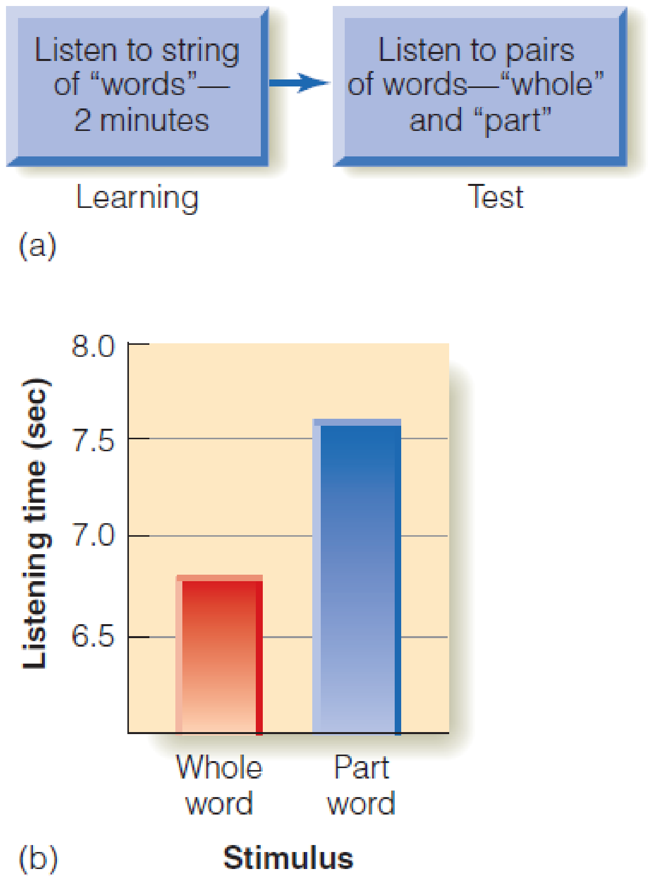
- Transitional probabilities: the chance that one sound will follow another in a language.
- Statistical learning: the process of learning transitional probabilities and other language characteristics.
  - Infants as young as eight months show statistical learning.

Perceiving Breaks between Sequences of Words (4 of 8)

Experiment by Saffran et al.

- Learning phase: infants heard nonsense words in two-minute strings of continuous sound that contained transitional probabilities.
- Nonsense words were in random order within the string.
- If infants use transitional probabilities, they should recognize the words as units even though the string of words had no breaks.

Perceiving Breaks between Sequences of Words (5 of 8)



Perceiving Breaks between Sequences of Words (6 of 8)

Examples of transitional probabilities

- Syllables within a word: bidaku: syllable “da” always followed “bi”, which is a transitional probability of 1.0
- Syllables between words: “ku” from “bidaku” was not always followed by “pa” from “padoti” or “tu” from “tupiro”
  - The transitional probability of either of these combinations occurring was .33.

Perceiving Breaks between Sequences of Words (7 of 8)

Testing phase: infants presented with two types of three syllable stimuli from the strings.

- Whole-words: stimuli (bidaku, tupiro, padoti) that had transitional probabilities of 1.0 between the syllables.
- Part-words: stimuli created from the beginning and ends of two words (tibida from the end of padoti and the beginning of bidaku).

Perceiving Breaks between Sequences of Words (8 of 8)

During the testing phase, each stimulus was preceded by a flashing light near the speaker that would present the sound.

- Once the infant looked at the light, the sound would play until the infant looked away.

Infants listened longer to the part-words, which were new stimuli, than to the whole-words.

Perceiving Degraded Speech

How well can we understand speech heard under adverse conditions?

Davis and colleagues (2005): noise vocoding

- Pop-out effect

Speech Perception and the Brain (1 of 4)

Broca's aphasia: individuals have damage in Broca's area in frontal lobe

- Labored and stilted speech and short sentences but they understand others.

Wernicke's aphasia: individuals have damage in Wernicke's area in temporal lobe.

- Speak fluently, but the content is disorganized and not meaningful
- Also have difficulty understanding others, and word deafness may occur in extreme cases.

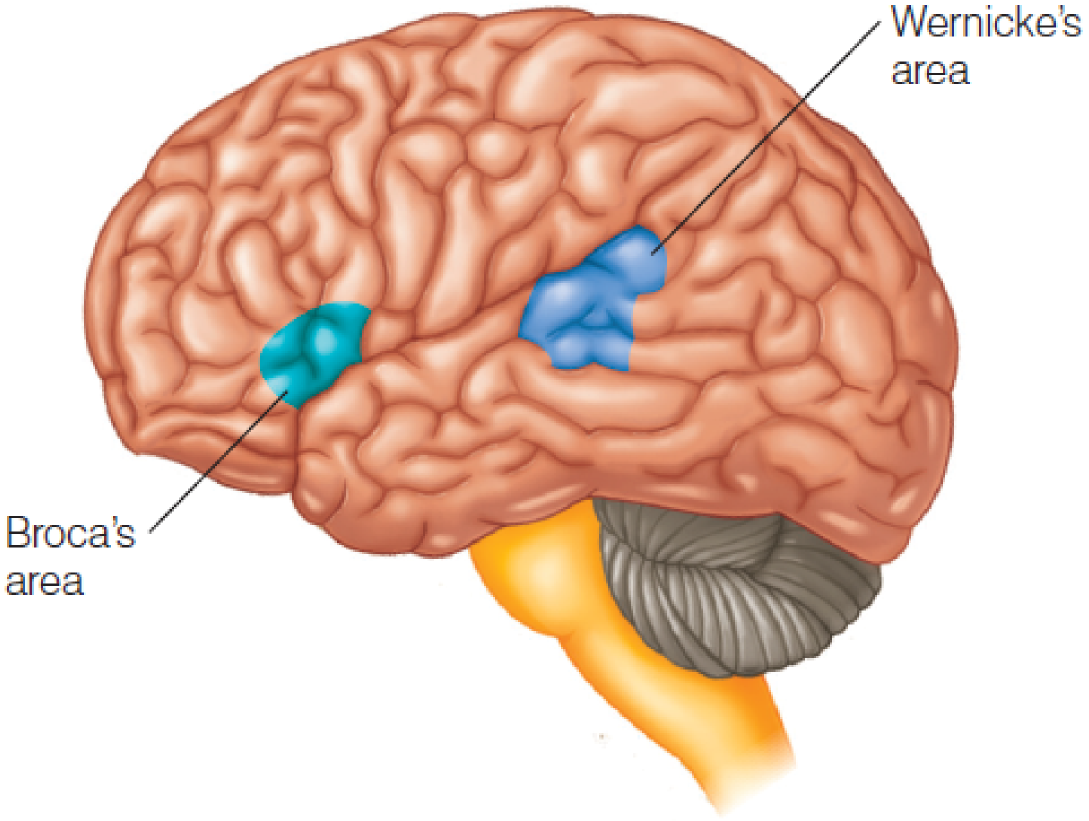
Speech Perception and the Brain (2 of 4)

Brain images show that some patients with brain damage can discriminate syllables but are able to understand words.

Brain scans have also shown that there is:

- A “voice area” in the STS that is activated more by voices than other sounds.
- A ventral stream for recognizing speech and a dorsal stream that links the acoustic signal to movements for producing speech, called the dual stream model of speech perception.

Speech Perception and the Brain (3 of 4)



Speech Perception and the Brain (4 of 4)

Berlin and colleagues (2000) used fMRI to locate a “voice area” in the superior temporal sulcus.

Perrodin and coworkers (2011): voice cells

Dual-stream model of speech perception

Mesgarani and coworkes (2014) recorded from electrodes directly placed on the temporal lobe of patients undergoing brain surgery.

- Responses to single phonemes
- Responses to phonetic features

Speech Perception and Action (1 of 2)

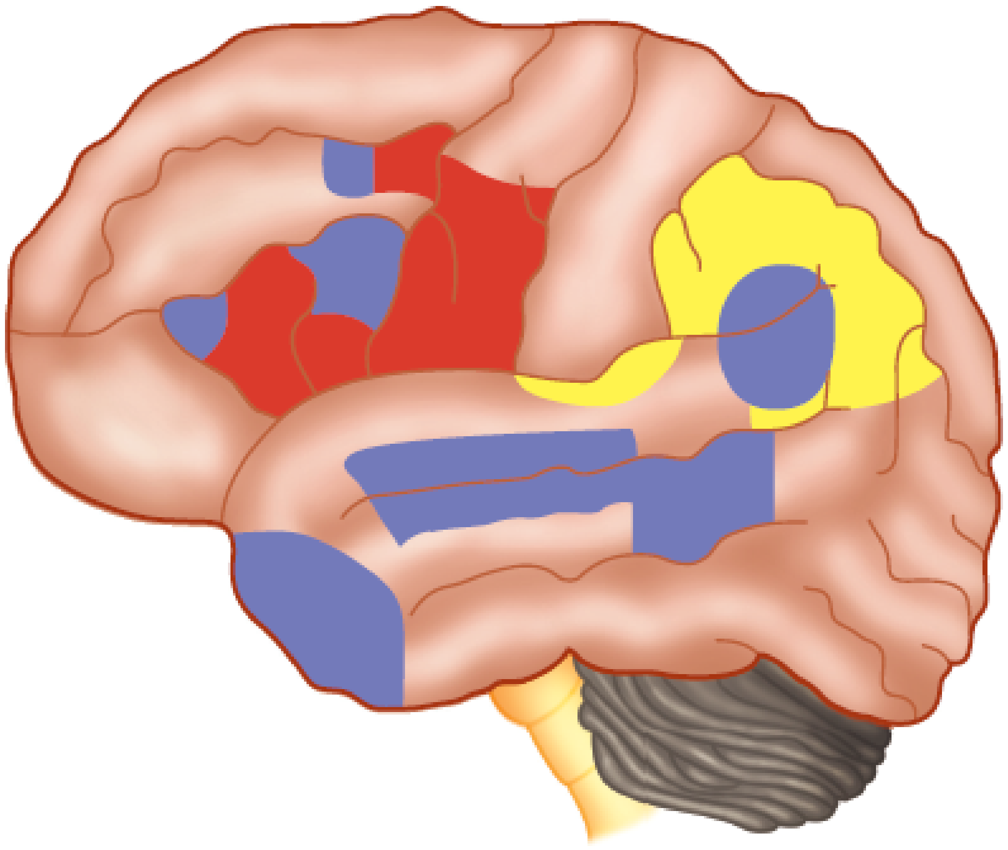
Liberman et al. proposed that motor mechanisms responsible for producing sounds activate mechanisms for perceiving sound.

Evidence from monkeys comes from the existence of audiovisual mirror neurons.

Experiment by D’Ausilio et al.

- Focal transcranial magnetic stimulation
- Demonstrated a link between production and perception

Speech Perception and Action (2 of 2)



Infant Speech Perception

Categorical perception

- Habitation procedure

Learning the sounds of language

- Very young infants: “citizens of the world”

Experience-dependent plasticity: a change in the brain’s ability to respond to specific stimuli that occurs as a result of experience.

Chapter Summary (1 of 2)

Describe how the acoustic signal is created by the action of articulation and is represented by phonemes.

Understand the processes responsible for variability in the acoustic signal.

Describe the motor theory of speech perception and evidence for and against the theory.

Describe the multiple sources of information for speech perception.

Chapter Summary (2 of 2)

Understand how people perceive degraded speech.

Describe how research involving brain damage and neural recording has contributed to our understanding of how speech is processed by the brain.

Understand how cochlear implants work and how they have been used in children.

Describe infant-directed speech and how it affects infants.