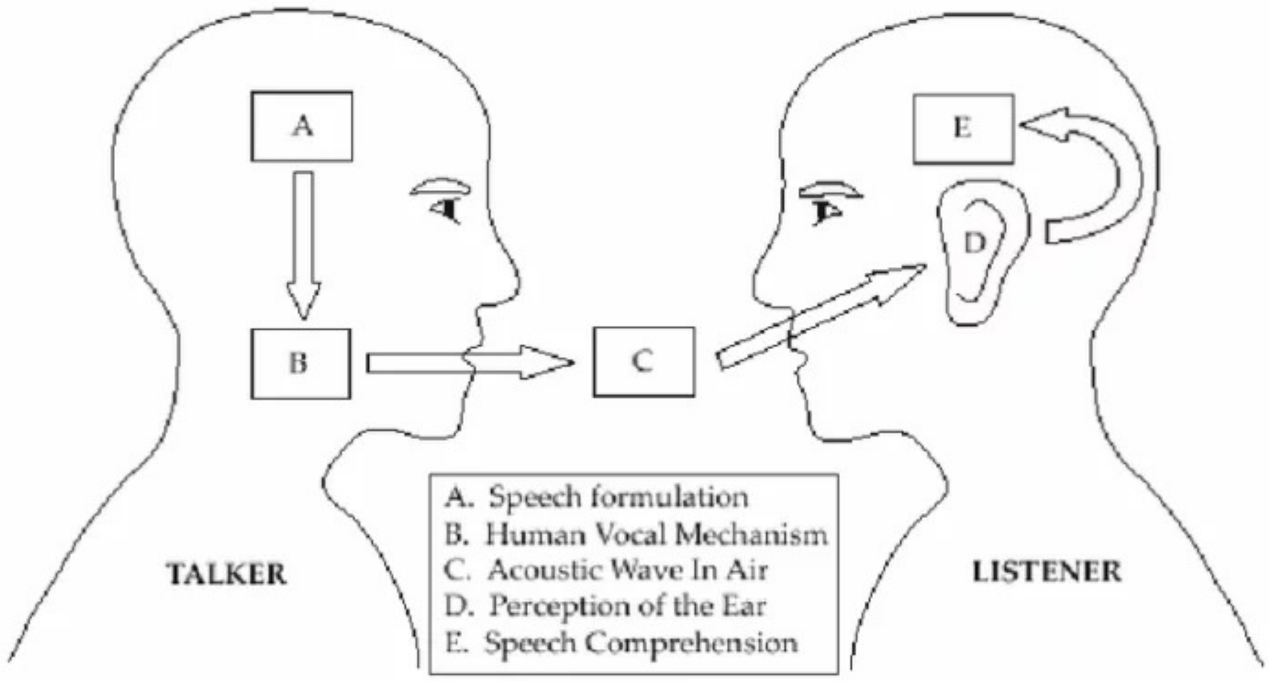
Week 4: Speech perception

Speech perception

- Speech perception is the process by which the sounds of language are heard, interpreted, and understood.
- The study of speech perception is closely linked to the fields of phonology and phonetics.



Speech perception is complicated

How do we perceive speech?

Processes involved in mapping speech signals onto language forms.

→ No problem?

Problems in speech perception

- **Sequentiality (序列性)**: humans produce 25-30 phonemes (or phonetic segment) per second. Compared to its visual counterpart is that the distribution of the auditory information is time bound, transient, and solely under the speaker’ s control.
- In particular, given that relatively little information is conveyed per unit of time, the extraction of meaning can only be done within a window of time that far exceeds the amount of information that can be held in memory.

Problems in speech perception

- **Continuity**: The inspection of a speech waveform does not reveal clear acoustic correlates of what the human ear perceives as phoneme and word boundaries. The lack of boundaries is due to **coarticulation** between phonemes.
- The lack of clear and reliable gaps between words, along with the sequential nature of speech delivery, makes speech continuity one of the most challenging obstacles.

Football is a dangerous sport.

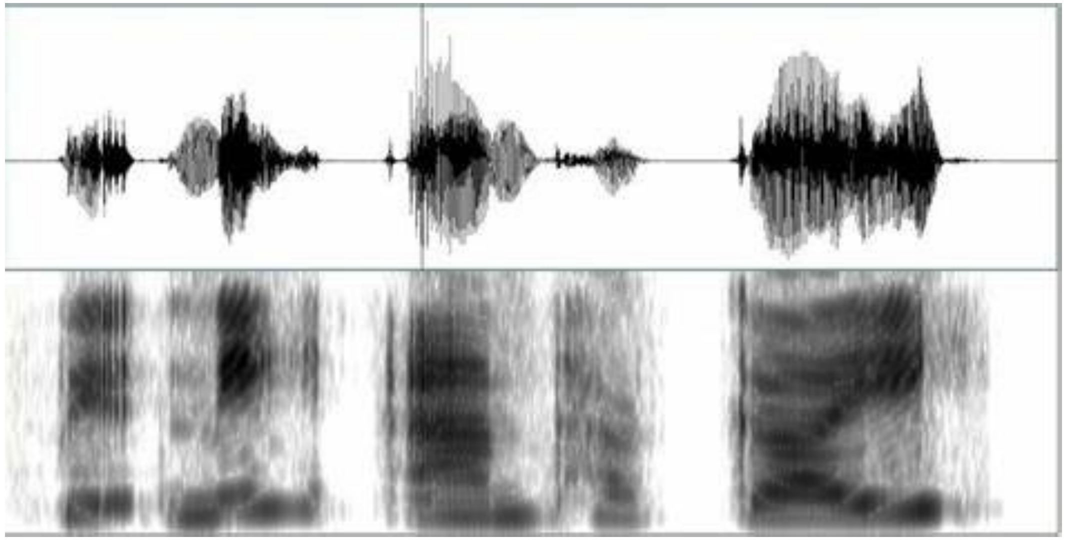
- Co-articulation
  - Coarticulation refers to the phenomenon in which the pronunciation of one speech sound is influenced by the preceding or following sounds.

    produce cat or the phonemes of cat individually

  - It is a fundamental aspect of human speech production and occurs due to the natural coordination and efficiency of articulatory movements.
  - It helps in producing speech quickly and efficiently, allowing for smooth and continuous communication.

- How do we transform continuously-chaging speech into discrete segments with co-articulation?

- Can you read the word “dog” for me?



- Co-articulation
  - Anticipatory coarticulation: If the sound becomes more like the following sound, as in the case of lamb, it is known as anticipatory coarticulation or right-to-left coarticulation (L < R).
  - Perseverative coarticulation: If the sound displays the influence of the preceding sound, it is perseverative coarticulation or left-toright coarticulation (L > R).

Problems in speech perception

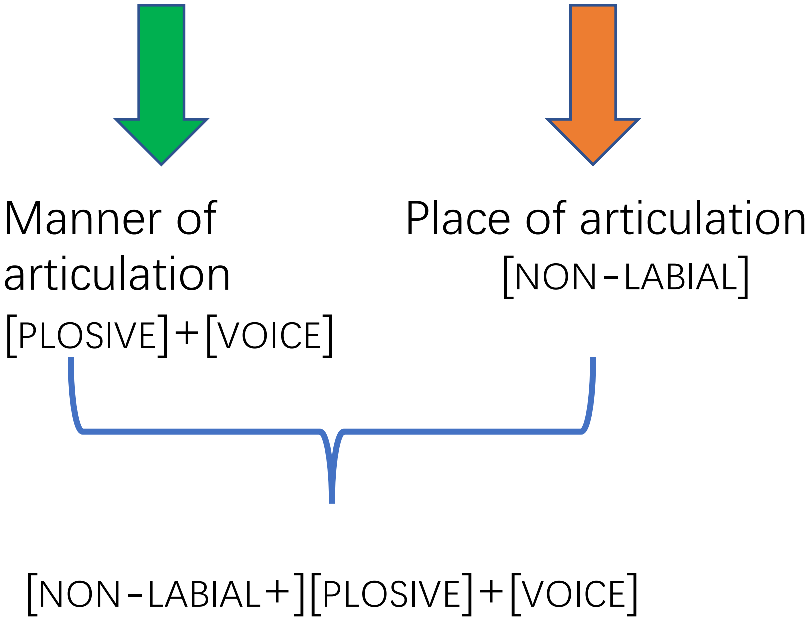
- Variability
  - There is no simple mapping of linguistic units (phonemes) to corresponding acoustic units (waveforms).
  - Everyone speaks differently, and the environment or even media alters the acoustic structure of speech signals.
  - Speech variability happens at a lower phoneme level. Unlike visual letters whose realizations have at least some commonality from one instance to another, phonemes can vary widely in their acoustic manifestation—even within the same speaker.

Pronounce the syllable by yourself

Speech perception is more complicated in real world!

➢ **McGurk effect**

- illusion perception of audiovisual speech when auditory and visual cues are incongruent
- Visual information could be supplementary not only in noisy environment, but also when auditory signal is clear
- Examples: auditory /pa/ + visual /ka/ ⟶ perception /ta/



➢ Previous studies

- Linguistic experience indeed affect auditory-visual speech processing

  - English speakers v = ıpanese speakers (children aged betwee

  - English speakers > . Japanese speakers (adults)

  - Chinese speakers < Japanese speakers

Cultural & linguistic explanation

Tonal cues

Chinese speakers < Japanese speaker < English speaker

Implications for speech perception

- Simultaneous transmission of information
  - Each unit in the speech signal contains more than one segment
- Lack of invariance
  - The acoustic pattern that corresponds to any given segment varies considerably (between speakers, within speakers, depending on context)
    - No simple acoustic info-to-phoneme template
    - Theories must deal with these problems

Basic phenomena in speech perception

- Categorical perception
- Cue weighting
- Context
- Adaption

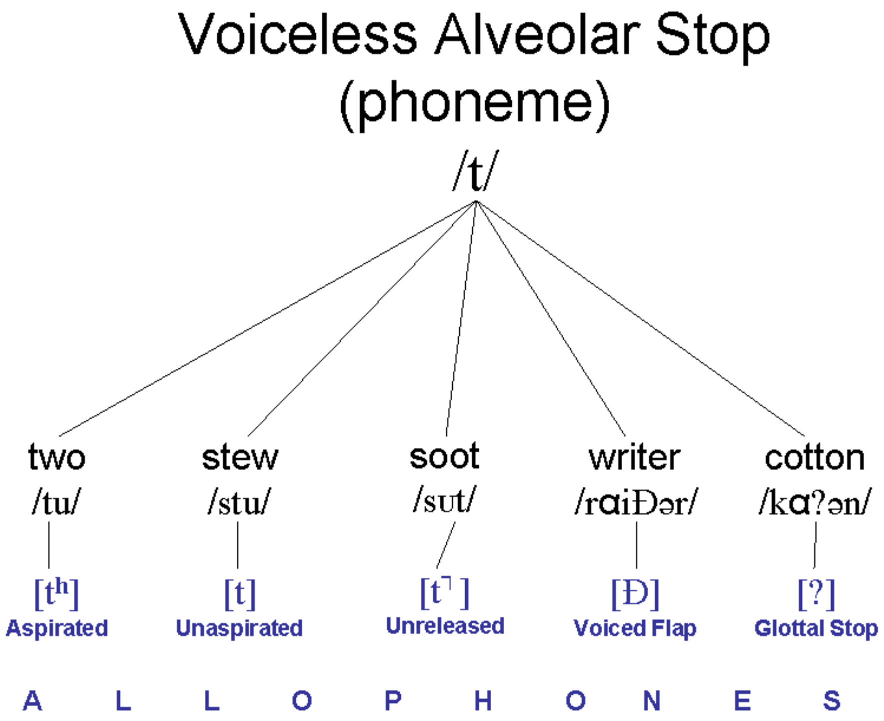
Categorical perception

- When we perceive speech, we’ re doing more than just responding to actual physical sounds out in the world. Our minds impose a lot of structure on speech sounds—structure that is the result of learning: we mentally represent sounds in terms of **abstract categories.**
- We mentally group clusters of similar sounds that perform the same function into categories called **phonemes.**

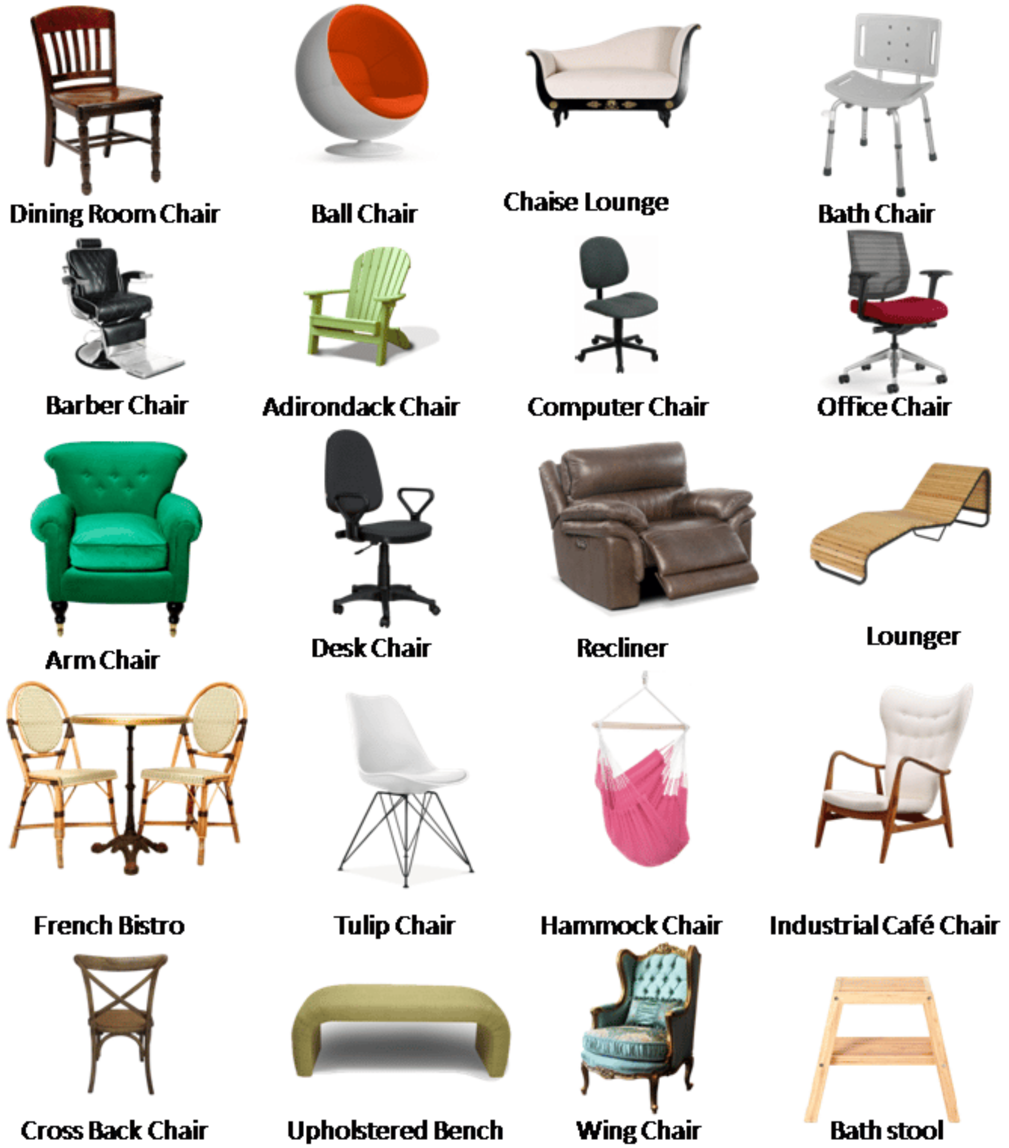
Phoneme

- The smallest unit of sound that changes the meaning of a word; often identified by forward slashes; e.g., /t/ is a phoneme because replacing it (e.g., by /p/) in the word tan makes a different word.
- This unit is an abstraction that corresponds to a variety of possible pronunciations.

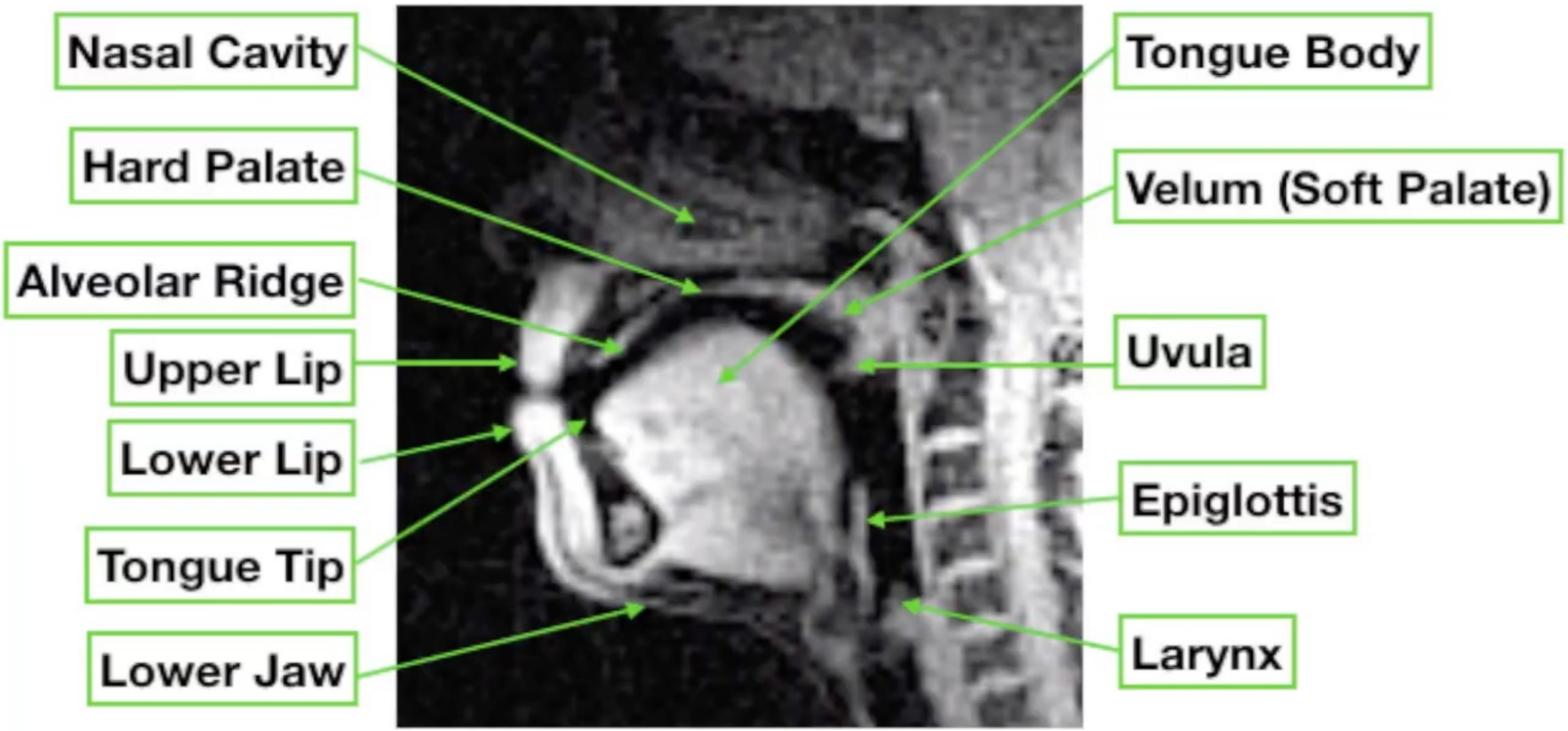
- The category of phonemes is broken down into variants—**allophones (音位变体)**—that are understood to be part of the same abstract category.
- Two or more similar sounds that are variants of the same phoneme



Types of Chair



- **Organs of articulation**: tongue, teeth, lips, nasal cavity, alveolar ridge, palate, etc.

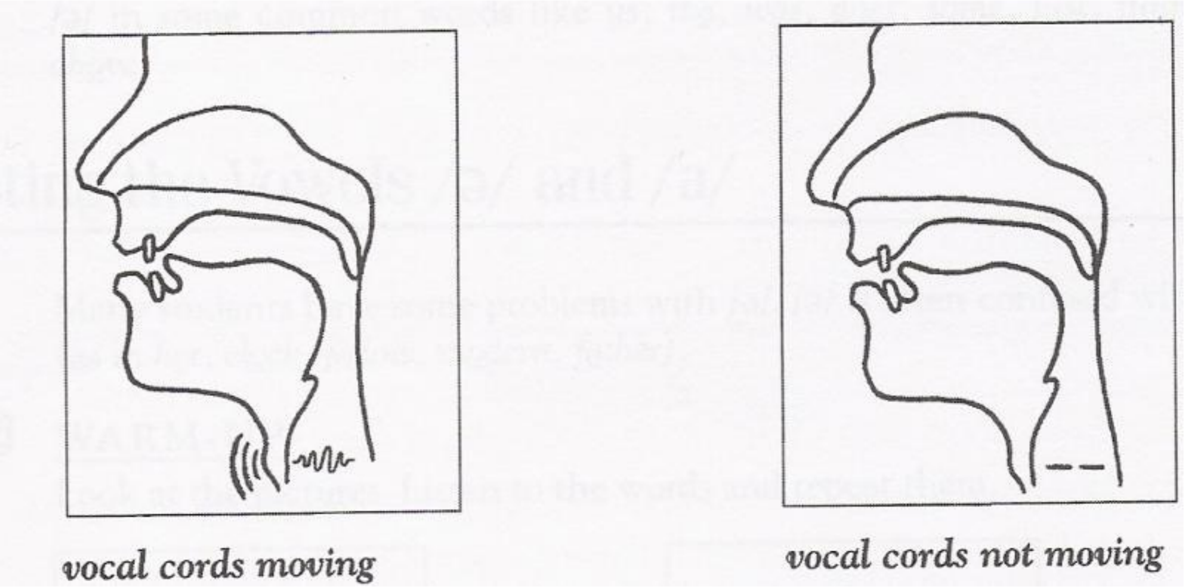


Consonant

- A class of speech sounds
- It includes sounds produced with a certain obstruction/block of the air stream using speech organs
- We focus on three dimensions: the **voiced/voiceless distinction**, the **place of articulation**, and **the manner of articulation**

Voicing can refer to the articulatory process in which the vocal folds vibrate:

- voiced: when the vocal folds are drawn together, the air from the lungs repeatdedly pushes it apart as it passes through, with a vibration effect
- voicelss: when the vocal folds are spread apart, the air from the lungs passes between them without obstruction, producing voiceless sounds



van – fan

Place of articulation

- the place of articulation (also point of articulation) of a consonant is a location along the vocal tract to create consonant sounds where its production occurs.
    - Air builds up in the lungs;
    - Air then moves toward the trachea, larynx;
    - The chest muscles control airflow;
    - The vocal cords in the larynx, start a vibration cycle which builds up air pressure and generates acoustic waves;
    - The airflow can go either through the oral cavity or the nasal cavity, depending on the sound.
    - Finally, the air is modified by the articulators (lips, tongue, teeth and palate).

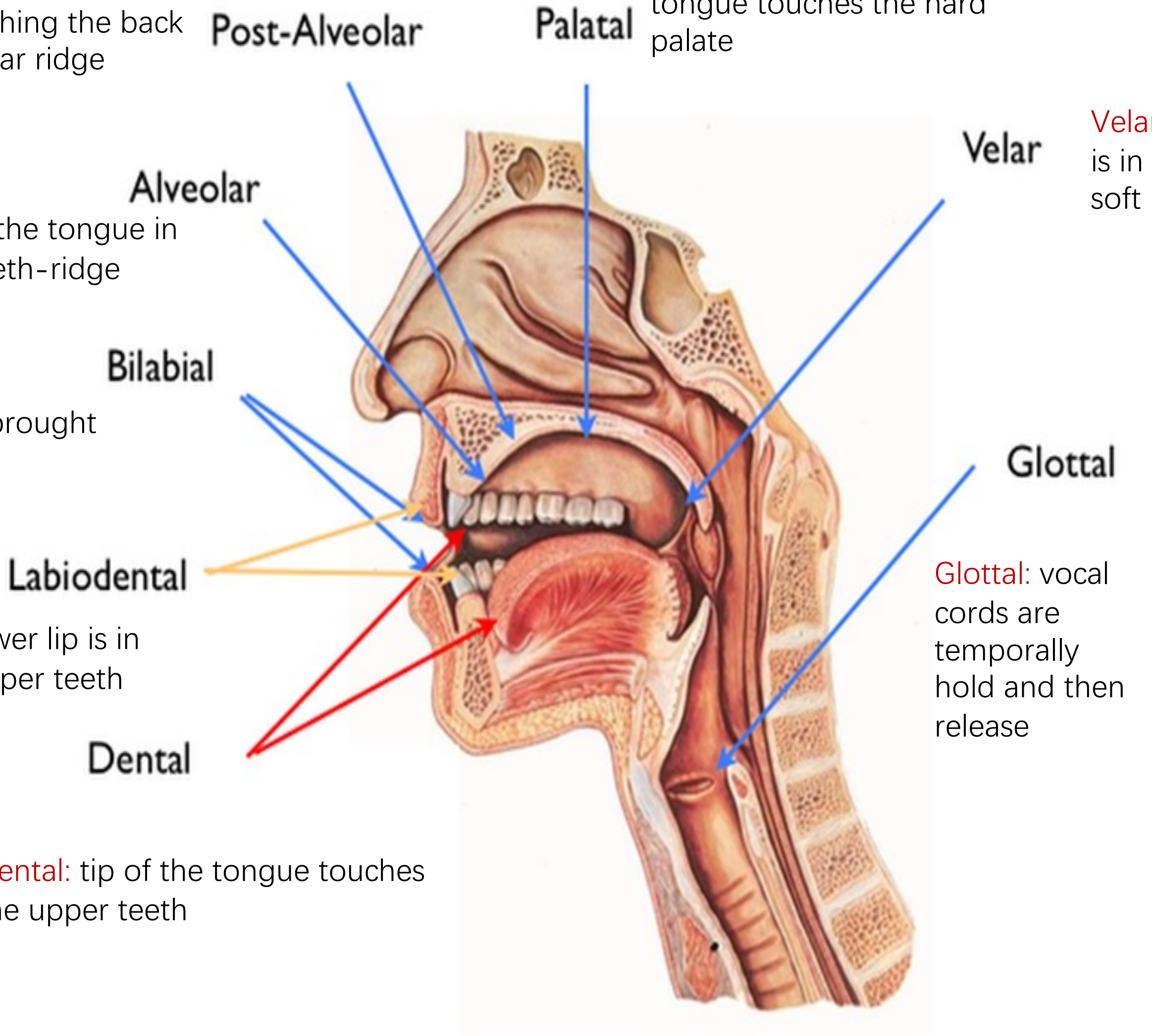
Post-alveolar: the tongue near or touching the back of the alveolar ridge

Palatal: the back of the tongue touches the hard palate

Velar: the back of tongue is in contact with velum, soft palate

Alveolar: tip of the tongue in contact with teeth-ridge

Bilabial: lips brought together



Glottal: vocal cords are temporally hold and then release

Labiodental: lower lip is in contact with upper teeth

Dental: tip of the tongue touches the upper teeth

Consonants	Voiceless	Voiced	Place of articulation
Bilabials	[p]	[b], [m], [w]	both (=bi) lips (=labia)
	pet, tape	***bet, met, wet***	
Labiodentals	[f]	[v]	upper teeth with lower lip
	fat, safe	***vat, save***	
Dentals	[θ]	[ð]	tongue tip behind upper teeth
	thin, bath	***then, bathe***	
Alveolars	[t], [s]	[d], [z], [n], [l], [r]	tongue tip to alveolar ridge
	top, sit	***dog, zoo, nut, lap, rap***	
Palatals	[ʃ], [ʧ]	[ʒ], [ʤ], [j]	tongue and palate
	ship, chip	***casual, gem, yet***	
Velars	[k]	[g], [ŋ]	back of tongue and velum
	cat, back	***gun, bang***	
Glottals	[h]		space between vocal folds
	hat, who		

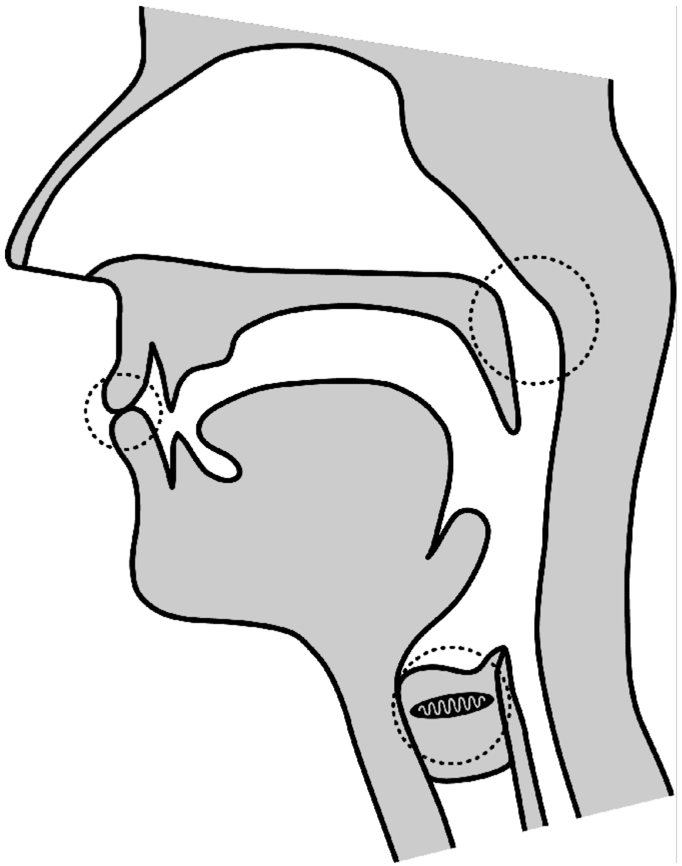
Manner of articulation

- the manner of articulation refers to how air flows through the vocal tract, based on the size and shape of the constriction between the articulators.

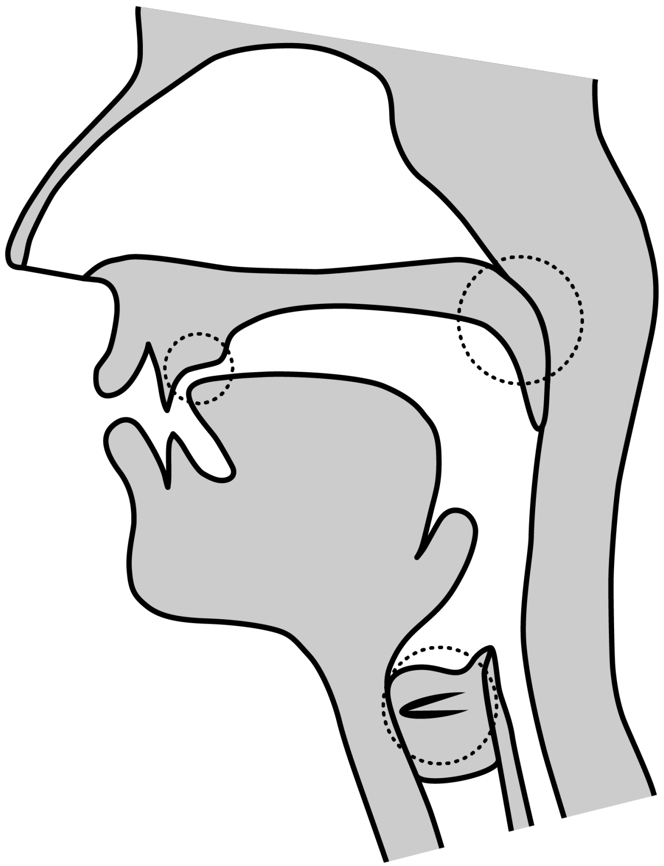
Consonants	Voiceless	Voiced	Manner of articulation
Stops	[p], [t], [k]	[b], [d], [g]	block airflow, let it go abruptly
	pet, talk	*bed, dog*	
Fricatives	[f], [θ], [s], [ʃ], [h]	[v], [ð], [z], [ʒ]	almost block airflow, let it escape through a narrow gap
	faith, house, she,	*vase, the, rouge*	
Affricates	[ʧ]	[ʤ]	combine a brief stop with a fricative
	cheap, rich	*jeep, rage*	
Nasals		[m], [n], [ŋ]	lower the velum, let air flow out through nose
		morning, name	
Liquids		[l], [r]	raise and curl tongue, let airflow escape round the sides
		load, light, road, write	
Glides		[w], [j]	move tongue to or from a vowel
		we, want, yes, you	

PVM Chart: English

			PLACE							
			LABIAL		CORONAL				DORSAL	
	MANNER	VOICING	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Palatal	Velar	Glottal
OBSTRUENTS	Stop	Voiceless	p			t			k	ʔ
		Voiced	b			d			g	
	Fricative	Voiceless		f	θ	s	ʃ			h
		Voiced		v	ð	z	ʒ			
	Affricate	Voiceless					tʃ			
		Voiced					dʒ			
SONORANTS	Nasal	Voiced	m			n			ŋ	
	LIQUID Lateral	Voiced				l				
	LIQUID Rhotic	Voiced					ɹ			
	Glide	Voiced	w					j	w	



bilabial nasal voiced



alveolar fricative voiceless

Vowel

- A class of speech sounds
- It includes sounds that their production does not involve any block of the air as it comes out from the lungs
- A vowel sound is a sound that is produced by an open configuration of the vocal tract with a vibration of the vocal cords

	Front		Central	Back	
High	i				u
	ɪ			ʊ	
Mid	e		ə		o
	ɛ		ʌ	ɔ	
Low	æ				
			a	ɑ	

Categorical perception

- Each phoneme can be thought of as a cluster of features—for example, /p/ can be characterized as a voiceless labial stop, whereas /z/ is an alveolar voiced fricative
- Establishing a phonological representation for the input involves perceiving the features of phonemes
- The power of these mental representations in guiding the perception of speech is indicated by categorical perception

Categories

Are these continuous changes in a stimulus being perceived gradually?

Figure 7.2 Is it a cup or a bowl? The category boundary isn't clear, as evident in these images inspired by an experiment of linguist Bill Labov (1973). In contrast, the boundary between different phonemic categories is quite clear for many consonants, as measured by some phoneme-sorting tasks. (Photograph by David McIntyre.)

Categorical perception

- Categorical perception is a phenomenon of perception of distinct categories when there is gradual change in a variable along a continuum. A pattern of perception where continuous changes in a stimulus are perceived not as gradual, but as having a sharp break between discrete categories.
- Mental categories impose sharp boundaries, so that you perceive all sounds that fall within a single phoneme category as the same, even if they differ in various ways, whereas sounds that straddle phoneme category boundaries clearly sound different.
- This means that you’re rarely in a state of uncertainty about whether some has said “bear” or “pear,” even if the sound that’s produced falls quite near the voiced/unvoiced boundary.

How to prove it?

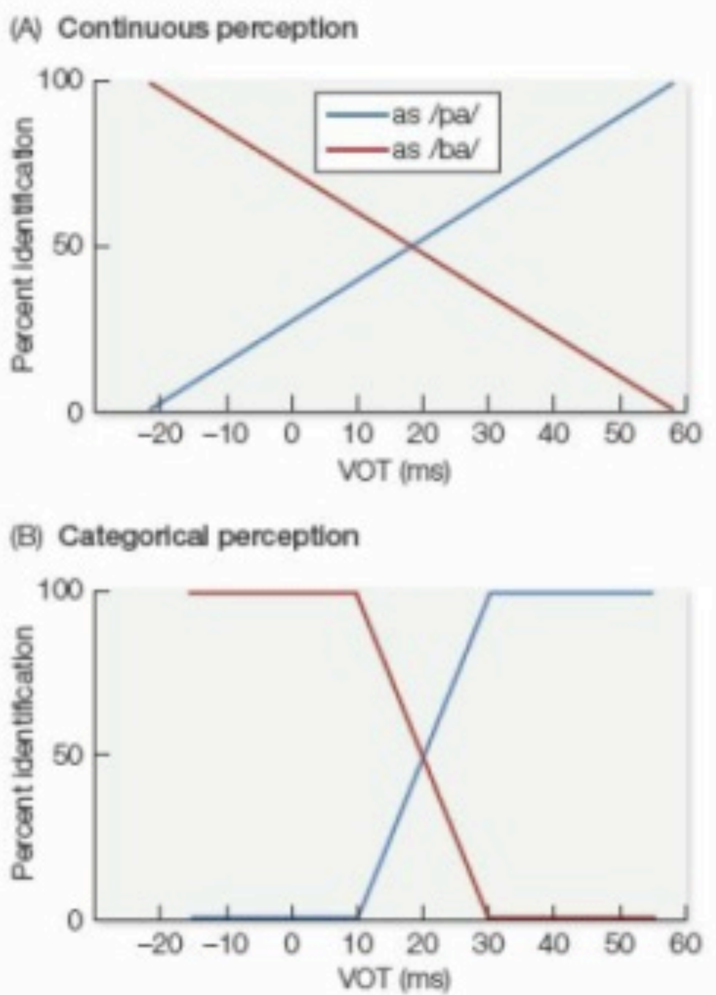


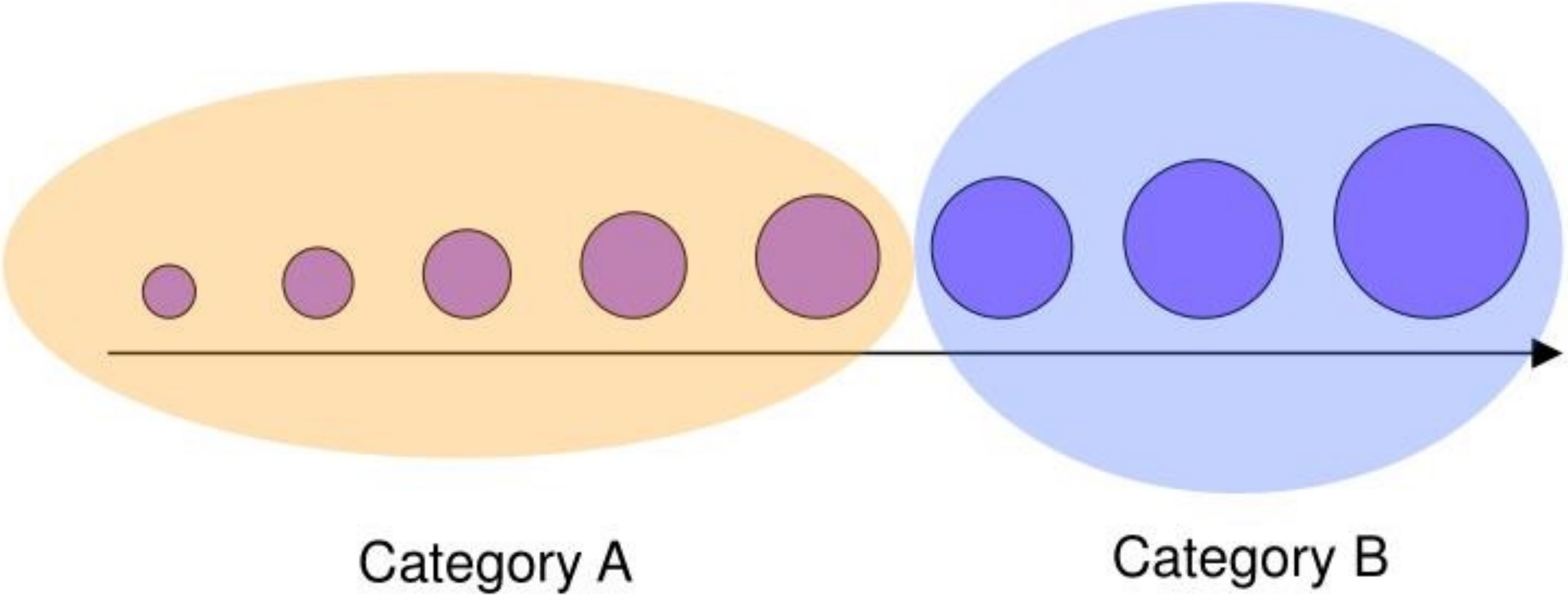
Figure 7.4 Idealized graphs representing two distinct hypothetical results from a phoneme forced-choice identification task. (A) Hypothetical data for a perfectly continuous type of perception, in which judgments about the identity of a syllable gradually slide from /ba/ to /pa/ as VOT values increase incrementally. (B) Hypothetical data for a sharply categorical type of perception, in which judgments about the syllable's identity remain absolute until the phoneme boundary, where the abruptly shift. Many consonants that represent distinct phonemes yield results that look more like (B) than (A); however, there is variability depending on the specific task and specific sounds.

Let’ s look at an example

- The study was done by Liberman and his colleagues (1957)
- **A forced-choice identification task** was used to test categorical perception.
- The strategy is to have people listen to many examples of speech sounds and indicate which one of two categories each sound represents (for example, /pa/ versus /ba/).
- The speech sounds are created in a way that varies the voice onset time (VOT) in small increments—for example, participants might hear examples of each of the two sounds at 10-ms increments

What happened?

- When the VOT was shorter than 30ms, the perception shifted from "pa" to "ba"
- The shift did not gradually as VOT decreased
- This is the example of categorical perception



Categorical perception isn’ t limited to VOT

- Compare da and ba
- Compared sha and zha

- How much of speech perception is specifically about speech?
- Does our tendency to sharply categorize speech sounds stem from our capacity for speech, or is this about how our auditory system deals with sound more general?

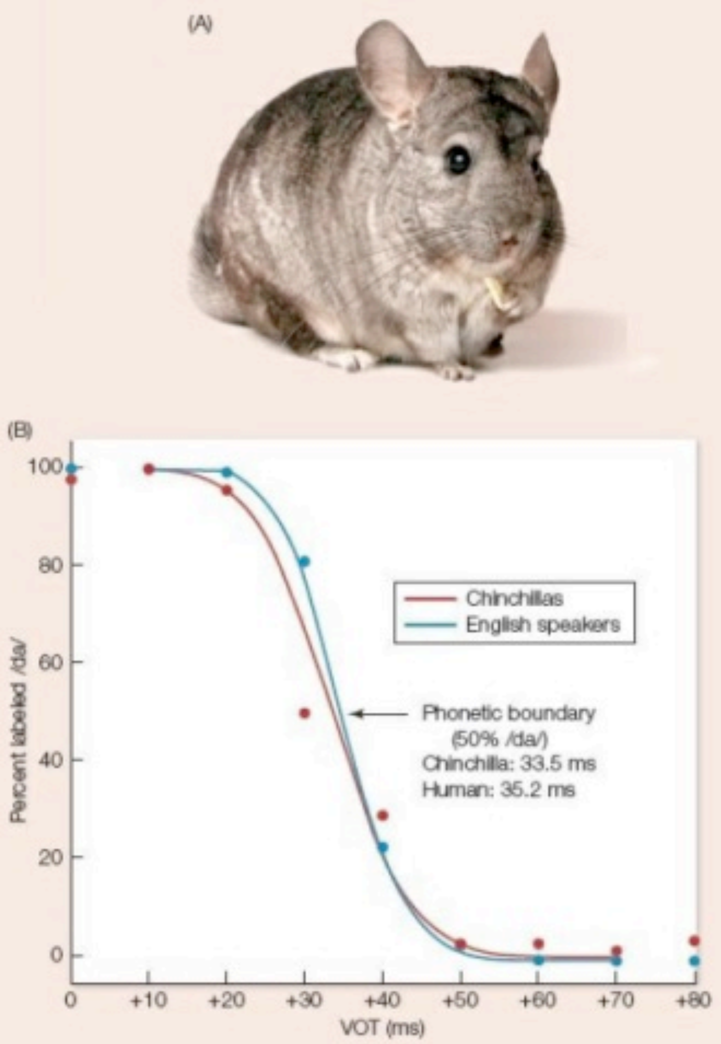


Figure 7.5 (A) Chinchillas are rodents about the size of a squirrel. They are a good choice for auditory studies because their range of hearing (20–30 kHz) is close to that of humans. (B) Results from Kuhl and Miller's categorical perception experiment, comparing results from the animals and human adults. The graph shows the mean percentage of trials in which the stimulus was treated as an instance of the syllable /da/. For humans, this involved asking the subjects whether they'd heard a /da/ or /ta/ sound; for chinchillas, it involved seeing whether the animals fled to the other side of the cage or stayed to drink water. (A photo by David McIntyre; B after Kuhl & Miller, 1975, *Science* 190, 69.)

Recent development

- McMurray (2008):
  - **Exp. A:** classic force-choice identification task (/ba/ or /pa/?): participants heard nine examples of the syllables /ba/ and /pa/ evenly spaced along the VOT continuum, and pressed buttons labeled "b" or "p" to show which sound they thought they'd heard.
  - **Exp. B:** hear actual words (e.g., "beach" or "peach"), choose image: participants heard actual words, such as "beach" or "peach," while looking at a computer screen that showed images of both possibilities, and then had to click on the image of the word they thought they'd heard.

McMurray and his colleagues argued that true perception is fairly continuous, rather than categorical.

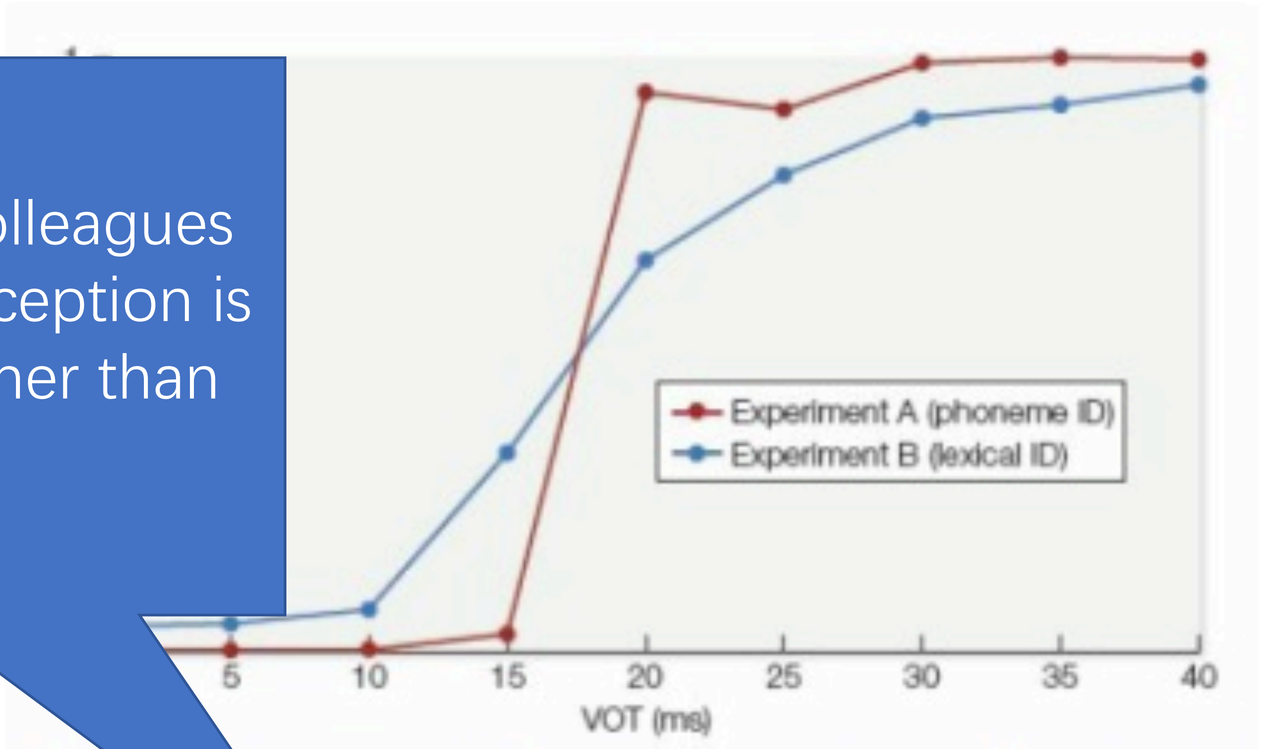


Figure 7.6 Proportion of trials identified as /p… …unction of VOT for Experiment A, which involved a forced-choice identification task where subjects had to indicate whether they heard /ba/ or /pa… …nd Experiment B, a word recognition task in which subjects had to click on the image that corresponded to the word they heard (e.g., *beach* or *peach*). (After McMurray et al., 2008, *J. Exp. Psych.: Hum. Percep. Perform.* 34, 1609.)

1. Typical categorical perception curve in Exp. A, with a very steep slope right around the VOT boundary;
2. The slope around the VOT boundary is much more graded, looking much more like the graph depicting continuous perception.

Alternatively?

- An fMRI study led by Emily Myers (2009) suggests a phenomenon known as **repetition suppression (重复抑制)**, a phenomenon in which the repetition of the same stimulus over time leads to a decrease in neural activity in the regions of the brain that process that type of stimulus.
- Some regions were sensitive to changes in VOT even when the two sounds fell within the same phonemic category. But other regions were mainly sensitive to VOT differences across phonemic boundaries, and less so to differences within categories.

Task difference

- Both of these tasks require subjects to consciously direct their attention to the speech sounds they’ re hearing, and in this sense, they may both differ from the word choice task set up by McMurray and colleagues—here, the hearer’ s task might be described as mapping the incoming sounds onto a word representation and then indicating the degree of confidence in one choice or another.
- This could arguably be done without the hearer consciously attending to the first sound in the word and deciding whether it’ s a /p/ or a /b/ sound.

Importance of categorical perception

- We have phonemic inventory in our native languge, and we use this inventory to categorize speech sounds as phonemes
- Highly variable speech sounds fall into discrete phonemic categories
- In this way, our categorical perception abilities allow us to make sense of speech signal rapidly and efficiently

Cue weighting

- Why most studies of categorical perception have focused on VOT?

  --- Well because it’ s easy to manipulate in the lab!!

  --- out in the real, noisy, messy world, it is far more complex!!

- 16 possible different acoustic cues between *rapid* and *rabid* (Lisker, 1986)

  --- the duration of the vowel

  --- the pitch contour

  --- how long the lips stay closed during the “p” or “b” sound; the pitch contour before and after the lip closure

  ---Etc..

- None of these cues is absolutely reliable on its own. On any given occasion, they won’ t all point in the same direction— some cues may suggest a voiced consonant, whereas others might signal a voiceless one. The problem for the listener, then, is how to prioritize and integrate the various cues

- The process of prioritizing the acoustic cues that signal a sound distinction, such that some cues will have greater weight than others

- VOT is a more reliable cue to voicing for English talkers

  --- than it is for Korean talkers

  --- Korean talkers rely heavily on, say, pitch at vowel onset, and the length of closure of the lips (Schertz et al., 2015)

- Different experiences with speech might result in different perceptual organization

Segmental vs. suprasegmental cues

- Listeners use suprasegmental cues for locating word boundaries. For example, languages whose words have a predominant rhythmic pattern (e.g., word-initial stress is predominant in English; word-final lengthening is predominant in French) provide a relatively straightforward—though probabilistic—segmentation strategy to their listeners.

Context effects in speech perception

- This is a basic cognitive function: to appreciate the color on an object, **the illumination source** has to be taken into account, which the brain does continuously.
  - Shadows are blue, so we mentally subtract the blue light in order to view the image, which then appears in bright colors—gold and white
  - However, artificial light tends to be yellowish, so if we see it brightened in this fashion, we factor out this color, leaving us with a dress that we see as black and blue.

Context

- Another way to achieve stable sound representations in the face of variable acoustic cues might be to use contextual cues to infer sound categories.

- **Ganong effect:** an effect in which listeners perceive the same ambiguous sound differently depending on which word it is embedded within
- e.g., __ ask   vs.   ___ash   for a sound between /t/ and /d/

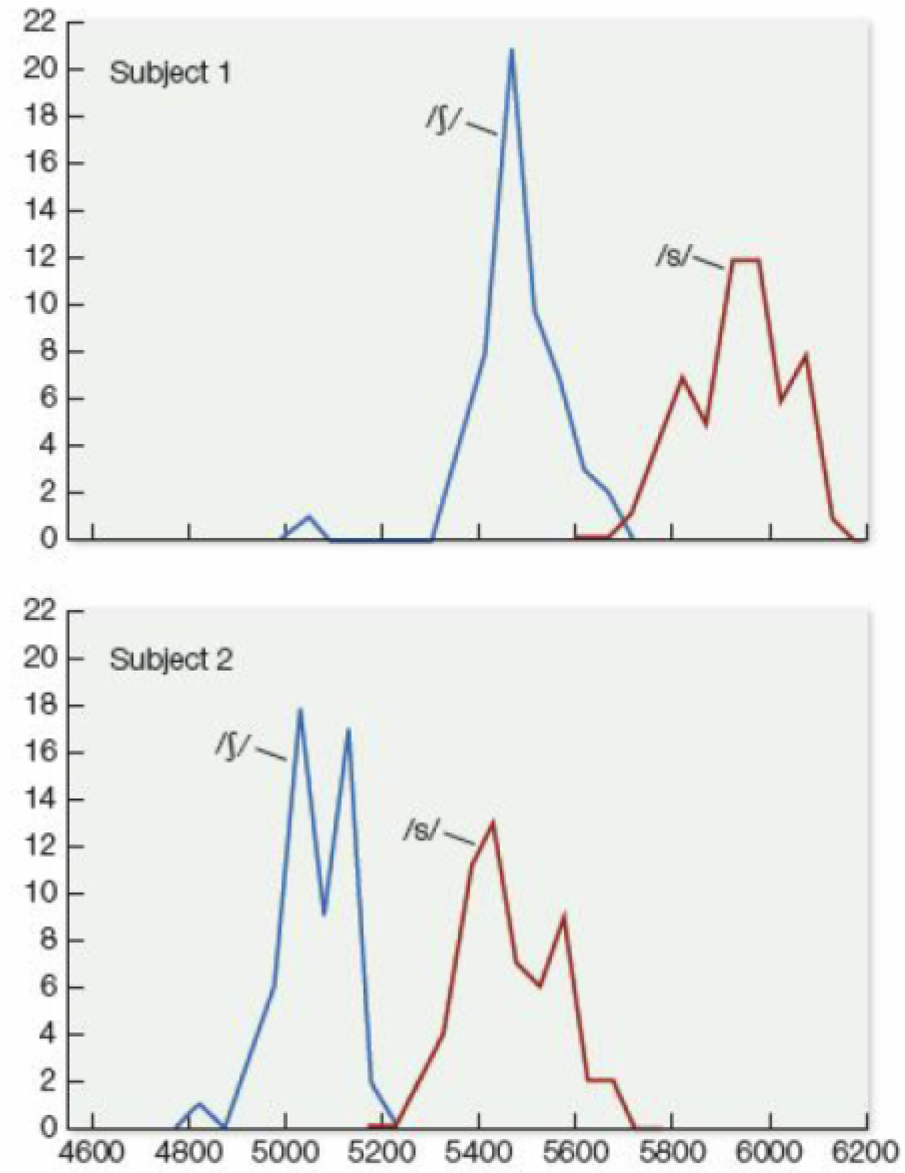
- **Phoneme restoration effect:** is a perceptual phenomenon where under certain conditions, sounds actually missing from a speech signal can be restored by the brain and may appear to be heard
- e.g., *legislature*

Adaption to a variety of talkers

- There' s a whole other layer of complexity we need to add: systematic variation across speakers.
  - unique voice
  - speaking style
  - age
  - gender
  - native language
  - etc.

Individual differences

- One talker’ s /ʃ/ sounds overlapping heavily with the other person’ s /s/ sounds. How do listeners cope with such variation?
- Do they base their sound categories on the average distributions of all the talkers they’ ve ever heard?
- If so, talkers who are outliers might be very hard to understand. Or do listeners learn the various speech styles of individual talkers and somehow adjust for them?



Evidence for adaption

- Lynne Nygaard and David Pisoni (1998), who trained listeners on sentences spoken by 10 different talkers over a period of 3 days, encouraging listeners to learn to label the individual voices by their assigned names.
- new sentences mixed with noise by familiar talkers vs. unfamiliar talkers after three days
- Those who transcribed the speech of the familiar talkers performed better than those tested on the new, unfamiliar voices
- the more noise was mixed into the signal, the more listeners benefitted from being familiar with the voices.

More evidence

- In a study by Eva, Reinisch and Lori Holt (2004), listeners heard training words in which sounds that were ambiguous between /s/ and /f/ were spliced into the speech of a talker who had a noticeable Dutch accent.
- Listeners adjusted their categories based on the speech sample they' d heard, and, moreover, carried these adjustments over into their perception of speech produced by a different talker who also had a Dutch accent.
- So, despite the fact that the Dutch speakers differed from the average English speaker on many dimensions, listeners were able to zero in on this systematic difference, and generalize across accented talkers.

Two theories of speech perception

1)Motor Theory （肌动理论）

2)Auditory Theory （听觉理论）

Motor theory

- the perception of speech sounds involves accessing representations of the articulatory gestures that are required to make those speech sounds.

PVM Chart: English			PLACE							
			LABIAL		CORONAL				DORSAL	
	MANNER	VOICING	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Palatal	Velar	Glottal
OBSTRUENTS	Stop	Voiceless	p			t			k	ʔ
		Voiced	b			d			g	
	Fricative	Voiceless		f	θ	s	ʃ			h
		Voiced		v	ð	z	ʒ			
	Affricate	Voiceless					tʃ			
		Voiced					dʒ			
SONORANTS	Nasal	Voiced	m			n			ŋ	
	LIQUID Lateral	Voiced				l				
	LIQUID Rhotic	Voiced					ɹ			
	Glide	Voiced	w					j	w	

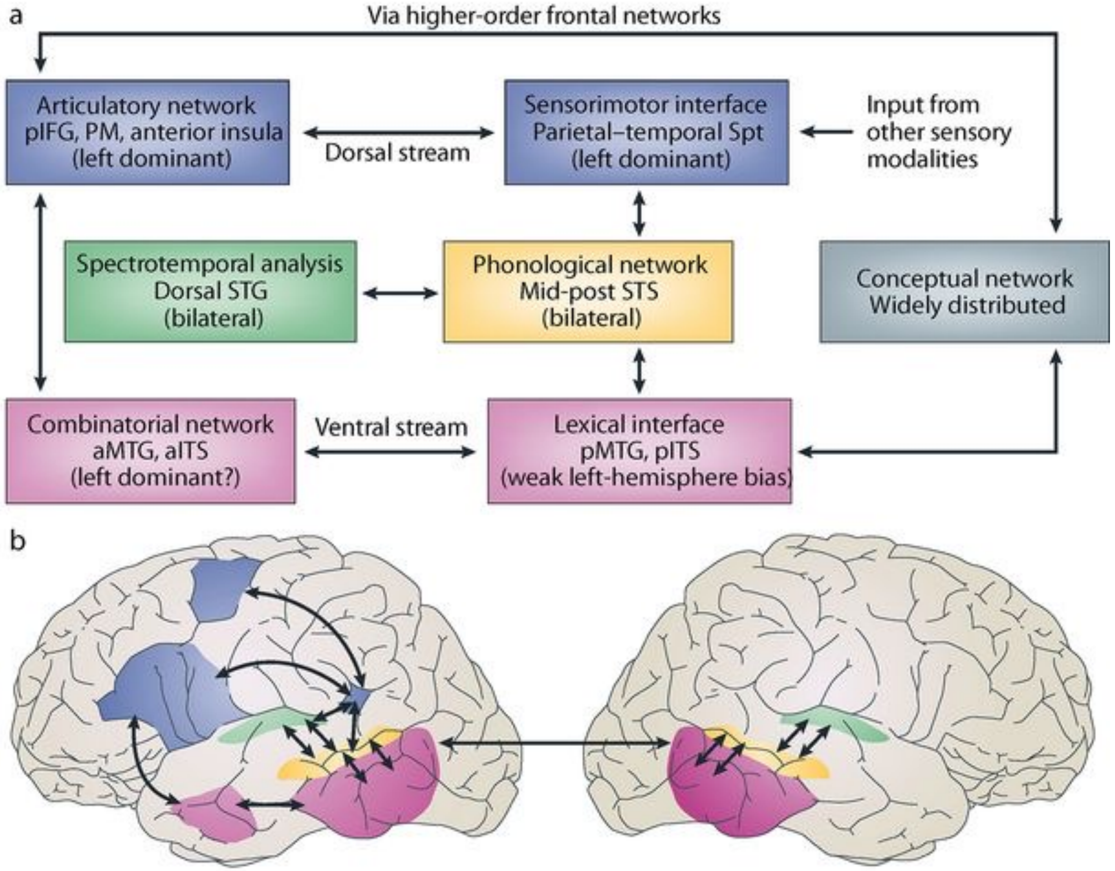
Motor theory

- Alvin Liberman (1967)
- Listeners perceive spoken words by reproducing the movements of the speakers' vocal tract rather than by identifying the sound patterns that speech generates
- motor system is involved in both production and perception of speech

Supporting evidence from behavioral studies

- McGurk effect
  - Visual information about how the sound is made (not the sound itself) affects perception
- Categorical perception
  - Phonemes can be defiend by their place and manner of articulation
  - One hear either a clear /ba/, or clear /pa/ after a certain point
  - There is agreement that our motor system produces only /pa/ or /ba/

Supporting evidence from cognitive neuroscience

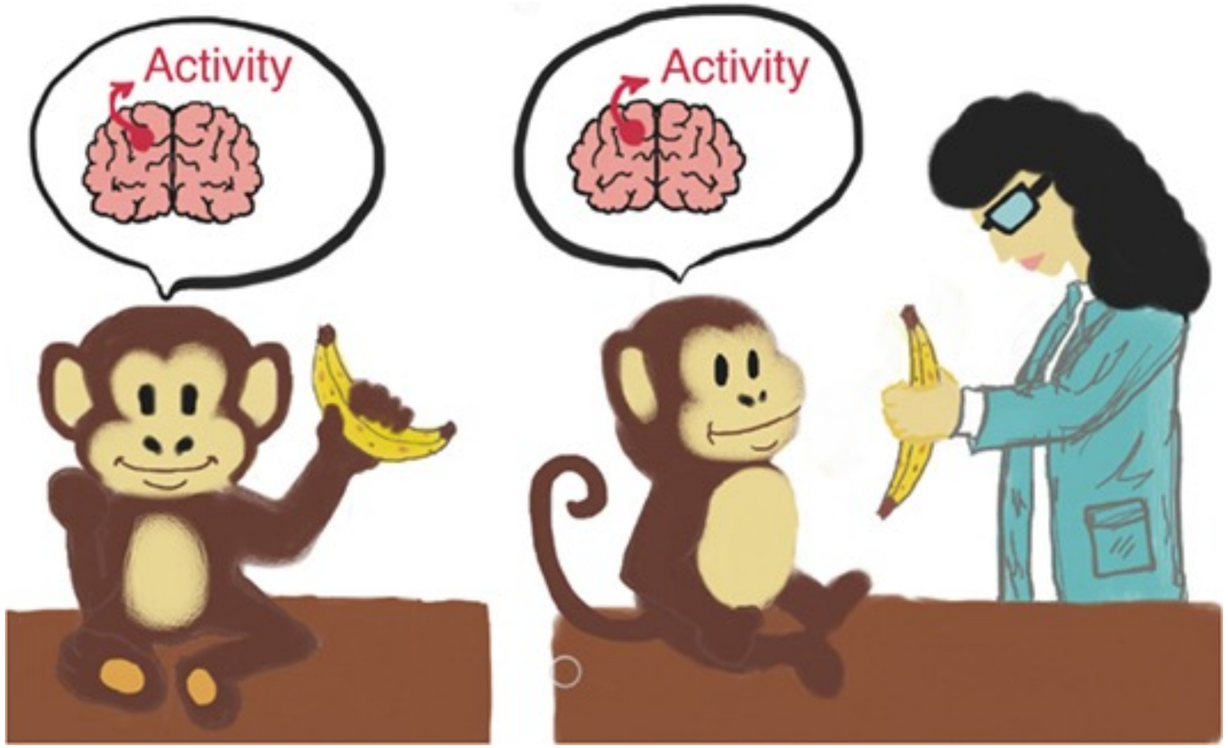


Ventral stream （腹测流）
dorsal stream （背测流）

- Brain imaging reveals that listening to speech/speech perception activates motor areas
- Disrupting the premotor cortex impairs speech perception

Supporting evidence from cognitive neuroscience

- Mirror neurons (Rizollatti et al., 1990s)
  - A mirror neuron is a neuron that fires both when an animal acts and when the animal observes the same action performed by another. Thus, the neuron "mirrors" the behavior of the other, as though the observer were itself acting



Motor theory

- Perception is based on production
  - Listeners know how speech is produced, therefore they can accurately perceive phonetic segments despite of coarticulation
- Perception is species-specific
  - Only humans know how speech is produced, so only humans can perceive phonetic structure of utterances
- Perception is innate
  - The specialized processing system is given

Problesm of the Motor theory

- Infants are good at speech perception, but not at production
  - Learning of speech must be based on different mechanisms
- Top-down influence
  - Context affects understanding, but motor cortex does not code meaning

Auditory theory

- Perception is not based on production
  - Specialized auditory system allows to perceive sounds
- Perception is not species-specific
  - As long as other animals have auditory system like humans, then these animals should perceive speech similarly to humans
- Perception could be innate
  - Auditory system is well developed in infancy, so perception is innate

Shen, G., & Froud, K. (2019). Electrophysiological correlates of categorical perception of lexical tones by English learners of Mandarin Chinese: an ERP study. *Bilingualism: Language and Cognition*, *22*(2), 253-265.

Abstract

This study examines brain responses to boundary effects with respect to Mandarin lexical *tone* continua for three groups of adult listeners: (1) native English speakers who took advanced Mandarin courses; (2) naïve English speakers; and (3) native Mandarin speakers. A cross-boundary tone pair and a within-category tone pair derived from tonal contrasts *(Mandarin Tone* 1/Tone 4; *Tone* 2/Tone 3) with equal physical/acoustical distance were used in an auditory oddball paradigm. For native Mandarin speakers, the cross-category deviant elicited a larger MMN over left hemisphere sensors and larger P300 responses over both hemispheres relative to within-category deviants, suggesting categorical perception of tones at both pre-attentive and attentional stages of processing. In contrast, native English speakers and Mandarin learners did not demonstrate categorical effects. However, learners of Mandarin showed larger P300 responses than the other two groups, suggesting heightened sensitivity to tones and possibly greater attentional resource allocation to tone identification.

Moon, H., & Magne, C. (2015). Noun/verb distinction in English stress homographs: An ERP study. *Neuroreport*, *26*(13), 753-757.

Abstract

Sensitivity to speech rhythm, especially the pattern of stressed and unstressed syllables, is an important aspect of language acquisition and comprehension from infancy through adulthood. In English, a strong correlation exists between speech rhythm and grammatical class. This property is well illustrated by a particular group of noun/verb homographs that are spelled the same but are pronounced with a lexical stress on the first syllable when used as a noun or on the second syllable when used as a verb. The purpose of this study was to further examine the neural markers of speech rhythm and its role in word recognition. To this end, event-related brain potentials were recorded while participants listened to spoken sentences containing a stress homograph either in a noun or a verb position. The rhythmic structure of the stress homographs was manipulated so that they were pronounced with a stress pattern that either matched or mismatched their grammatical class. Results of cluster-based permutation tests on the event-related brain potentials revealed larger negativities over the centrofrontal scalp regions when the stress homographs were mispronounced, in line with previous studies on lexical ambiguity resolution. In addition, differences between rhythmically unexpected nouns and verbs could be seen as early as 200 ms, suggesting that listeners are sensitive to statistical properties of their language rhythm. Together, these results support the hypothesis that information about speech rhythm is rapidly integrated during speech perception and contributes to lexical retrieval.

Zou, Y., Tsang, Y. K., & Wu, Y. (2019). Semantic radical activation in Chinese phonogram recognition: evidence from event-related potential recording. *Neuroscience*, *417*, 24-34.

Abstract

Some previous studies suggested that semantic radicals are activated during Chinese character recognition. However, many details about semantic radical processing remain unresolved. This study examines an often-overlooked factor, namely the "character status" of the semantic radicals. To be specific, some semantic radicals are themselves stand-alone characters (e.g., "口" in "咱"), while others are not (e.g., "亻" in "仿"). A masked priming character decision experiment with event-related potential (ERP) recording was conducted to compare the processing of these two types of radicals. Results showed that character semantic radicals elicited earlier onset, but less widely distributed, P200 than non-character semantic radicals (150 ms vs. 200 ms). Character radicals also elicited a statistically significant N400 earlier and with broader scalp distribution than non-character radicals (300 ms vs. 350 ms). Finally, only the character semantic radicals showed an effect on late positive complex (LPC). The differences in priming effects suggested that character and non-character semantic radicals are processed differently. The implication of the study was discussed with reference to the hierarchical model of Chinese character recognition.

Wang, Y., Li, Z., Jiang, M., Long, F., Huang, Y., & Xu, X. (2024). Time course of Chinese compound word recognition as revealed by ERP data. *Language, Cognition and Neuroscience*, *39*(1), 55-75.

ABSTRACT

Previous studies have yielded conflicting results regarding the onset of semantic processing in compound word recognition. This study examined the role of semantics in morphological processing using event-related potentials (ERP) recorded for Chinese compound targets primed by W+M+, W–M+, W–M– (W = whole-word semantics, M = morpheme meaning, + = congruent, and – = incongruent), semantically related and unrelated primes. Two experiments were conducted. In Experiment 1 of a masked priming lexical-decision task (SOA = 50 ms), EEG results demonstrated that the brain was sensitive to semantic information as early as between 100 and 250 ms. In Experiment 2 of an unmasked priming lexical-decision task (SOA = 200 ms), data confirmed early semantic access. The two EEG experiments also showed that the semantics of constituent morphemes may have little bearing on compound recognition. Overall, these results seem to converge with a form-and-meaning account of compound recognition.

and morpheme levels. Specifically, the targets (e.g. "干旱/drought") were primed by W+M+ ("干燥/arid", W = whole word semantics, M = morpheme meaning, + = congruent, and – = incongruent), W–M+ ("干洗/dry-cleaning"), W–M– ("干涉/interfere"), semantically related ("枯萎/withered", semantically but not morphologically or orthographically related) and unrelated primes ("歌曲/song"). Thus, the five priming conditions