Sound Category Learning as a Key Link

Patel ch. 2.4

Sound category learning as a key link

- Pitch and timbre are organized quite differently in musical and linguistic sound systems.
- But the fact that the mind has found two entirely different ways of building organized sound systems suggests that sound category learning is a fundamental aspect of human cognition.

Two Hypotheses

- Hypothesis I
  - The mechanisms that create and maintain learned sound categories in music and language are separate.
- Hypothesis II
  - Shared sound category learning mechanism hypothesis (SSCLMH)
  - The end products of development are domain specific.
  - But developmental processes are domain general

From Tan, S., Pfordresher, P. and R. Harré. Psychology of Music. 2010. Psychology Press.

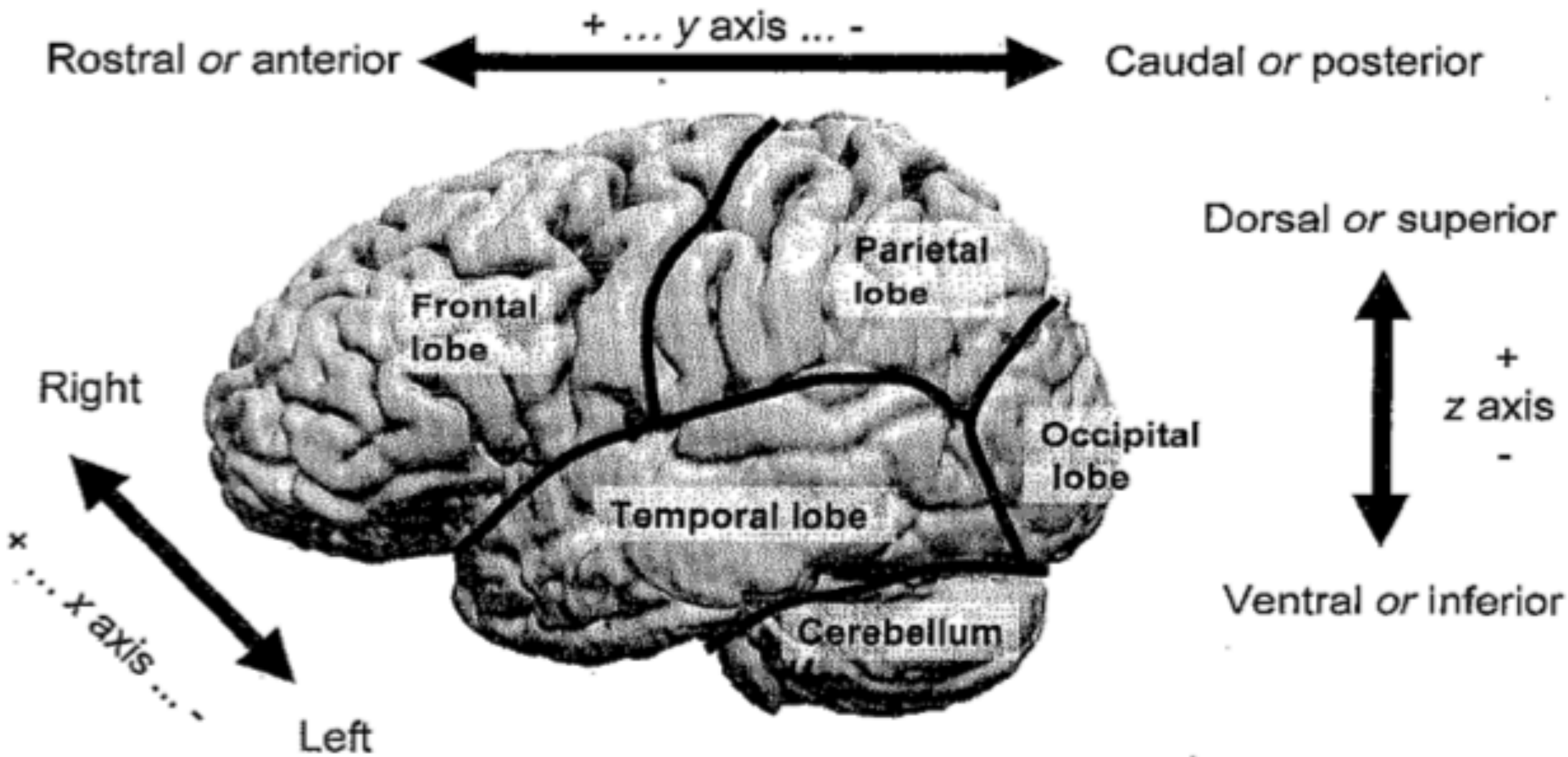


Figure 4.1 Lobes of the neocortex and the cerebellum, along with coordinate labels. The left hemisphere of author PQP's brain is shown.

Source: Brain image scans produced in collaboration with the Buffalo Neuroimaging Analysis Center (Robert Zivadinov, M.D., Ph.D., Jennifer L. Cox, Ph.D.) www.bnac.net. Used by permission from BNAC.

From Tan, S., Pfordresher, P. and R. Harré. Psychology of Music. 2010 . Psychology Press.

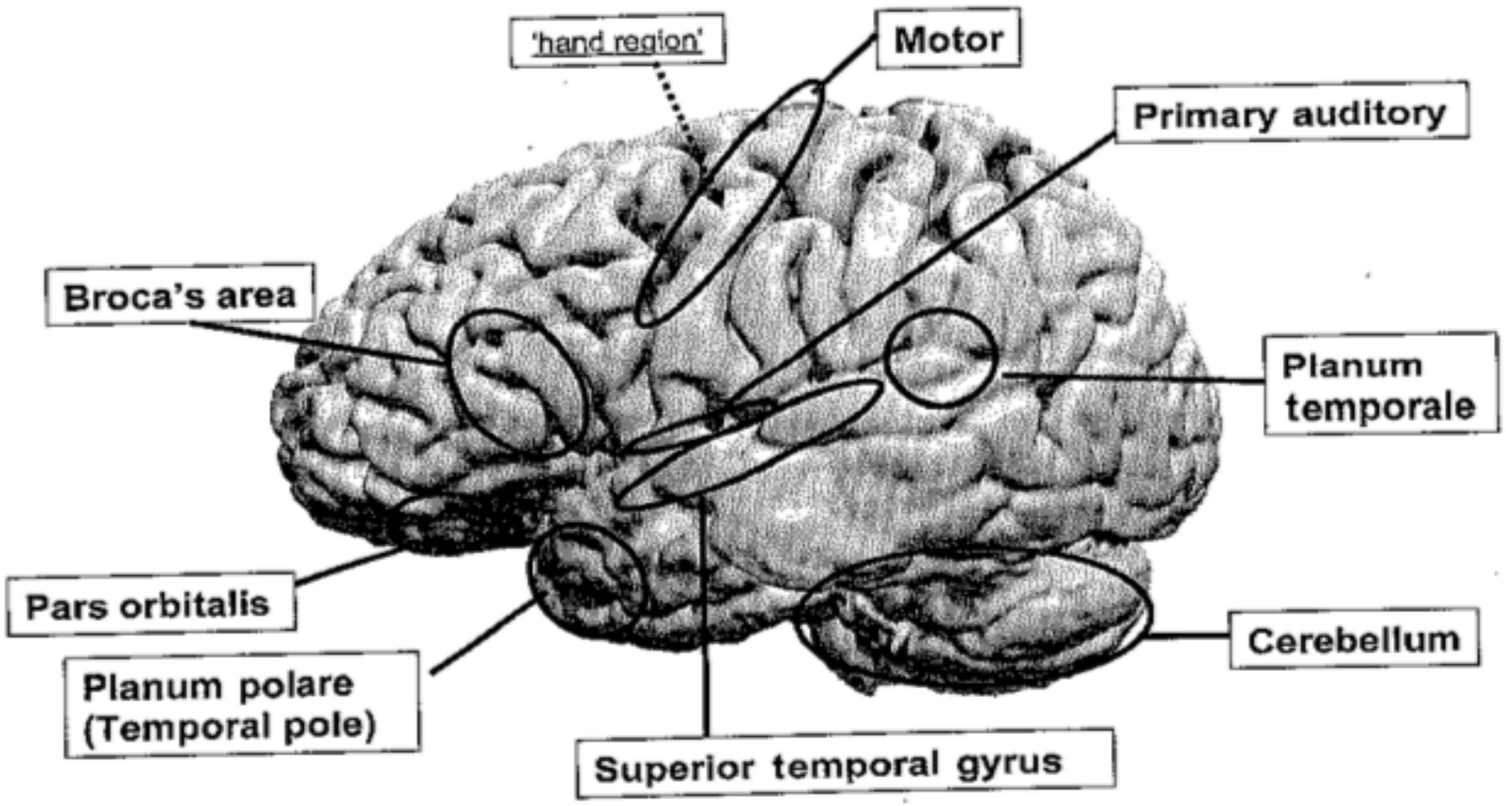


Figure 4.2 Surface regions of the brain associated with music. The left hemisphere of author PQP's brain is shown. Arrows are used to indicate regions on a more localized scale.

Source: Brain image scans produced in collaboration with the Buffalo Neuroimaging Analysis Center (Robert Zivadinov, M.D., Ph.D., Jennifer L. Cox, Ph.D.) www.bnac.net. Used by permission from BNAC.

Aphasia

- Broca's aphasia (agrammaticism):
  - Ah ... Monday ... ah, Dad and Paul Haney [himself] and Dad ... hospital. Two .. .ah, doctors ... and ah ... thirty minutes .. .and yes ... ah ... hospital. And, er, Wednesday ... nine o'clock. And er Thursday, ten o'clock .. .doctors. Two doctors ... and ah ... teeth. Yeah, ... fine.

- Wernicke's aphasia (anomia):
  - [*Trying to describe a picture of a child taking a cookie*.]

    Uh, well this is the ... the /dodu/ of this. This and this and this and this. These things going in there like that. This is /sen/ things here. This one here, these two things here. And the other one here, back in this one, this one /gesh/ look at this one.

[Source: Brain and Language Course, Mark Liberman: http: www.ling.upenn.edu/courses/Fall_2001/ling001/neurology.html]

Background: Dissociations

- Evidence from brain damage shows that music and language are can be differentially affected.
- Acquired amusia:
  - Can no longer recognize melodies (presented without words) that were highly familiar to them before the onset of brain damage, but are normal at recognizing spoken lyrics and spoken words in general.
- Congenital amusia:
  - Can recognize the lyrics of familiar songs but are unable to recognize the tune that usually accompanies them.

(Peretz & Coltheart 2003)

Dissociations cont.

- Tervaniemi et al 2006 had participants listen to a repeating speech sound (a nonsense syllable), or a repeating musical sound (a saxophone tone) that were roughly matched for duration, intensity and spectral content.
- fMRI scan showed both produced activation in bilateral auditory cortex in the superior temporal gyrus (STG), but in different places.

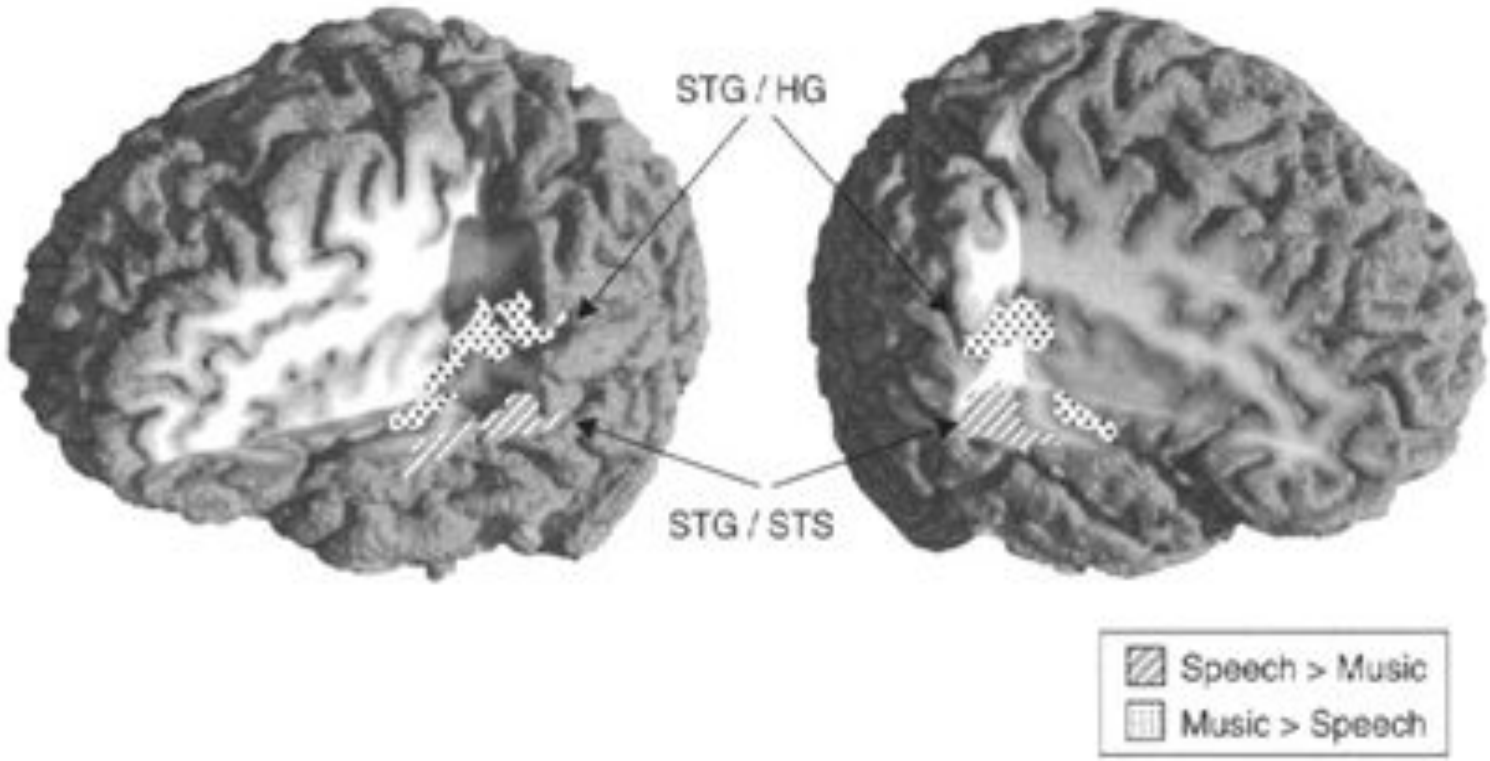


Figure 2.24 Brain regions showing significantly more activation during perception of speech versus musical sounds (diagonal lines) or vice versa (stippled dots). STG = superior temporal gyrus, HG = Heschl's gyrus (which contains primary auditory cortex), STS = superior temporal sulcus. From Tervaniemi et al., 2006.

Background: hemispheric asymmetries

- Traditionally, speech is assumed to be processed in the left hemisphere, and music in the right.
- But both hemispheres are involved in processing both.
- Hemispheric asymmetries are more subtle than generally appreciated.
- Patel: Nothing in this literature contradicts the idea of shared learning mechanisms for sound categories in the two domains.

Zatorre, Belin & Penhune 2002

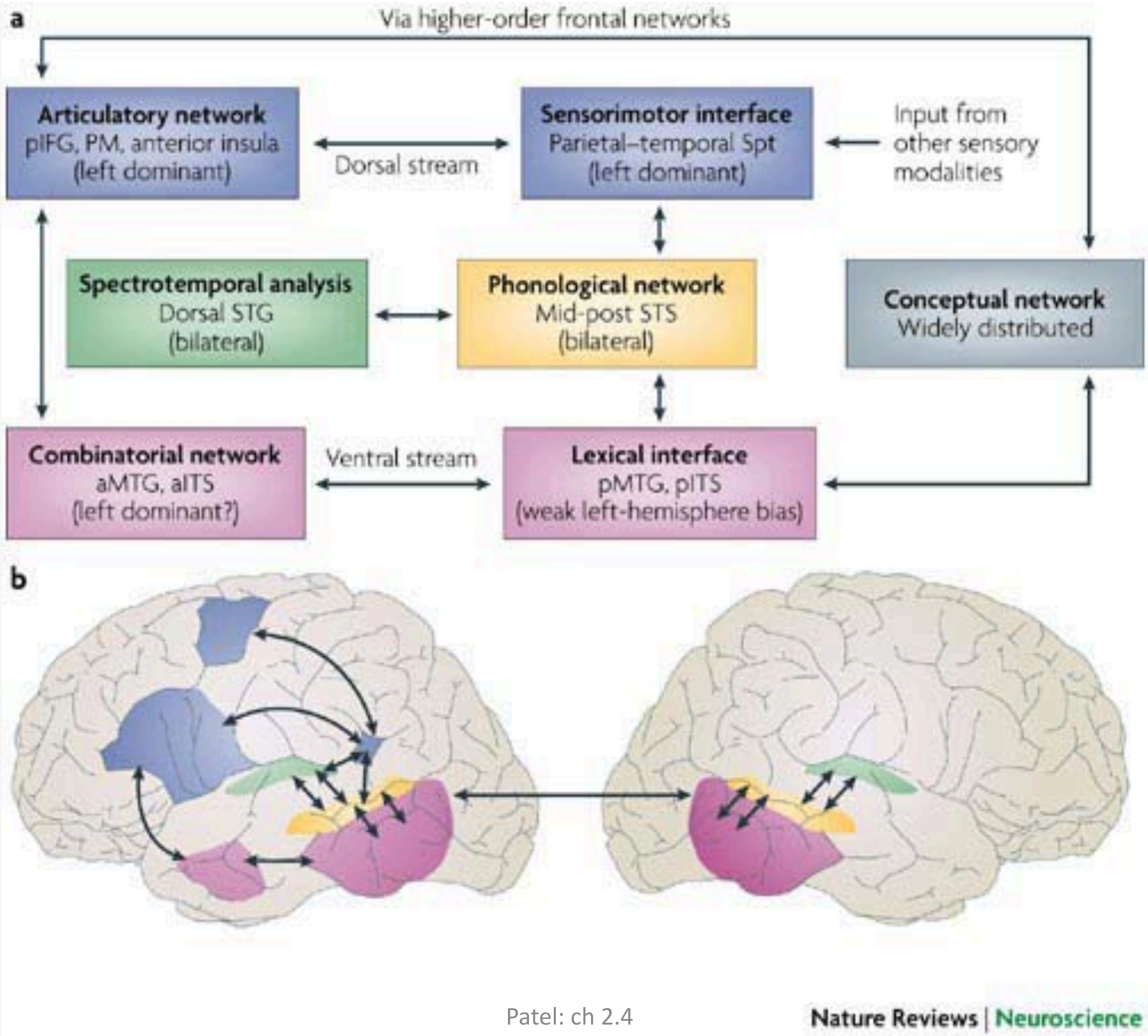
- "We examine the evidence that speech and musical sounds exploit different acoustic cues: speech is highly dependent on rapidly changing broadband sounds, whereas tonal patterns tend to be slower, although small and precise changes in frequency are important.
- We argue that the auditory cortices in the two hemispheres are relatively specialized, such that temporal resolution is better in left auditory cortical areas and spectral resolution is better in right auditory cortical areas.
- We propose that cortical asymmetries might have developed as a general solution to the need to optimize processing of the acoustic environment in both temporal and frequency domains."

Hemispheric asymmetries cont.

- Best & Avery 1999
  - Zulu but not English listeners exhibited a left-hemisphere bias in listening to click sounds.
- Gandour et al 2000
  - The perception of Thai tonal contrasts activates left frontal cortex in Thai but not in English or Mandarin listeners

Hickok & Poeppel 2007

- "Despite decades of research, the functional neuroanatomy of speech processing has been difficult to characterize. A major impediment to progress may have been the failure to consider task effects when mapping speech-related processing systems.
- We outline a dual-stream model of speech processing that remedies this situation. In this model, a ventral stream processes speech signals for comprehension, and a dorsal stream maps acoustic speech signals to frontal lobe articulatory networks.
- The model assumes that the ventral stream is largely bilaterally organized — although there are important computational differences between the left- and right-hemisphere systems — and that the dorsal stream is strongly left-hemisphere dominant."


a
Via higher-order frontal networks
Articulatory network
pIFG, PM, anterior insula
(left dominant)
Dorsal stream
Sensorimotor interface
Parietal–temporal Spt
(left dominant)
Input from
other sensory
modalities
Spectrotemporal analysis
Dorsal STG
(bilateral)
Phonological network
Mid-post STS
(bilateral)
Conceptual network
Widely distributed
Combinatorial network
aMTG, aITS
(left dominant?)
Ventral stream
Lexical interface
pMTG, pITS
(weak left-hemisphere bias)
b
Nature Reviews | Neuroscience

Hemispheric asymmetries cont.

- Liégeois-Chauvel et al 1998
  - 65 patients who had undergone unilateral temporal lobe cortical excisions for the relief of epilepsy.
  - Excisions to *both* hemispheres impaired the use of pitch interval information.
  - Only right hemisphere excisions impaired the use of pitch contour information.
  - Thus the use of learned sound categories of musical intervals appear to have bilateral representations.

Background: A "speech mode" of perception

- Ramez et al 1981.
  - Synthesize a sine wave for first three formants that exactly reproduces the formant's pattern of frequency change over time.
  - Sine-wave speech
  - Sounds meaningless at first, but then sounds like speech, seemingly using a mode of perception specific to speech.
- But there could still be shared processing mechanisms for *developing* learned sound categories in speech and music.
- And there could be an analog of sine-wave speech for music.

Evidence: Relations between musical ability and linguistic phonological abilities

- Anvari et al 2002
  - 5 year olds: performance on musical pitch tasks predicted variance in reading abilities.
- Slevc & Miyake 2006
  - 50 Japanese adult learners of English.
  - Music ability (for pitch pattern perception) predicted phonological skills in the L2 (English).

Evidence: Sound category-based distortions of auditory perception

- Perception Magnet Effect (PME), Kuhl 1991.
  - Best exemplar of a vowel such as /i/ was designated the “prototype”(P), and a poor exemplar was designated the “nonprototype” (NP).
  - Listeners heard either P or NP repeating as a background stimulus, and had to indicate when this sound changed to another version of /i/.
  - Listeners were less sensitive to sound change when the prototype served as the standard, as if the prototype was acting like a perceptual “magnet” that warped the space around it.

Sound category-based distortions of auditory perception cont.

- Acker et al. 1995 tested the PME in music, using chords instead of vowels.
  - Varied the tuning of E, G in a C major triad (C-E-G).
  - But in performing discrimination tasks, listeners performed *better* in the vicinity of the prototype.
  - They suggested that in music, category prototypes act as anchors rather than magnets.
  - But the subjects were all musically trained.
- Barrett 1997, 2000 tested musicians and nonmusicians.
  - Found that nonmusicians showed worse discrimination in the vicinity of the prototype.
  - Musicians learn to pay close attention to sounds in the vicinity of the prototype in order to tune properly.

Evidence: Decay in sensitivity for nonnative sound categories

- Lynch et al. 1990
  - Tested the ability of American infants and adults to detect mistunings in melodies based on familiar and unfamiliar scales.
    - Western major, minor scales, Javanese pelog scale.
    - Infants showed equally good discrimination on familiar and unfamiliar scales; while adult nonmusicians were better on familiar scales.
- Lynch & Eilers 1992, Lynch et al. 1995
  - Mistuning was randomly located rather than always on the same note.
    - Western infants did show an advantage for familiar scale, perhaps due to simpler frequency ratios of such scales.
- But the repeating background melody always occurred at the same pitch level instead of being transposed, so discrimination could have been based on absolute frequency instead of intervals.

Evidence: Exploring a common mechanism for sound category learning

- Statistical learning has been demonstrated for segmentation of word boundaries from sequences of syllables.
- Maye & Weiss 2003
  - Two groups of infants were exposed for 3 minutes to different distributions of the same speech tokens: either bimodal or unimodal.
  - Only the bimodal group discriminated tokens 3 and 6.
- An analogous experiment could be conducted with music intervals.
  - e.g. intervals that vary between a major third and a minor third in terms of pitch steps.

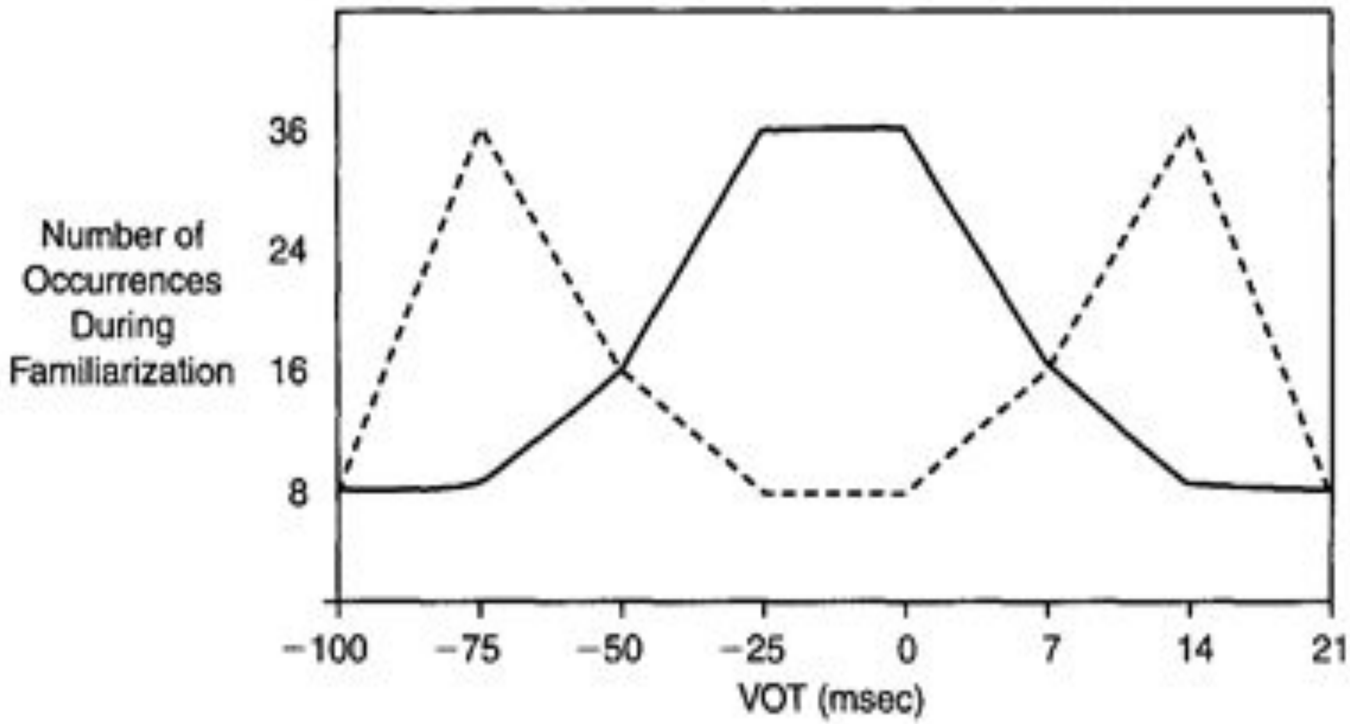


Figure 2.25 Distribution of stimuli along an acoustic continuum from /da/ to /ta/. One group of infants heard stimuli from a bimodal distribution (dashed line), whereas another group heard stimuli from a unimodal distribution (solid line). Importantly, both groups heard tokens 3 and 6 (–50 ms and 7 ms VOT, respectively) equally often.