24.964

Phonetic Realization

Does Language-Specific Detail Affect Phonological Distribution?

Stop voicing & bursts

Readings for next week:

- Flemming (2001).
- Cohn (1993).

Michael's question

- So far we have concluded that the evidence for language-specific phonetic detail affecting phonological distribution is inconclusive.
- What would constitute conclusive evidence?
- 'Knock-down' arguments are rare in phonology if they exist at all - are we setting the bar too high?
- Sensible procedure: develop a model and try to test it (in comparison to alternatives).
- We have been trying to test typological predictions of models according to which lg-specific phonetic detail should affect phonological typology.

Steriade (1997) - obstruent voicing

- Markedness of obstruent voicing contrast in context C depends on strength of cues to voicing in C.

Environment		Cues
*αVoice/ [-son] _ [-son], [-son]_#, #_[-son]	>>	clo voi, clo dur
*αVoice/ V_ [-son]	>>	clo voi, clo dur, V1 dur, F0, F1 in V1
*αVoice/ V_ #	>>	clo voi, clo dur, V1 dur, F0, F1 in V1, burst dur & amp
*αVoice/ V_ [+son]	>>	clo voi, clo dur, V1 dur, F0, F1 in V1, burst dur & amp, F0, F1 in V2

Image by MIT OpenCourseWare. Adapted from Steriade, Donca. "Phonetics in Phonology: The Case of Laryngeal Neutralization." Manuscript, UCLA, 1997. (PDF)

Steriade (1997) - obstruent voicing

- Implicational universals (cf. Lombardi 1991, Wetzels and Mascaro 2001)

	#_O, O_# e.g. bsa vs. psa	R_O e.g. absa vs. apsa	R_# e.g. ab vs. ap	_R e.g. ba vs. pa	R_R e.g. aba vs. apa
Totontepec Mixe	no	no	no	no	yes
Lithuanian	no	no	no	yes	yes
French	no	no	yes	yes	yes
Shilha	no	yes	yes	yes	yes
Khasi	yes	n/a	yes	yes	yes

Key: O = obstruent, R = sonorant, inc. vowel

Image by MIT OpenCourseWare. Adapted from Steriade, Donca. "Phonetics in Phonology: The Case of Laryngeal Eutralization." Manuscript, UCLA, 1997. (PDF)

- Problem for theories that predict that a language could neutralize word-finally without neutralizing before obstruents
  - E.g. Wetzels and Mascaro 2001: independent word-final devoicing, syllable final devoicing, assimilation.

Zhang - contour tones

- If a contour tone T is permitted on a syllable with $C_{CONTOUR} = d$ then it should be permitted on syllables with $C_{CONTOUR} \geq d$.
  - Accounts for many typological patterns, e.g. if a contour tone can appear on CV it can also appear on CVV.
    - based on assumptions (mostly safe) about relative C_{CONT} values of different syllable types.
  - $C_{CONTOUR}$ cannot be determined from standard phonological representations - lg. specific phonetic detail.
    - e.g. Cantonese: C_C(CVR) > C_C(CVVO)
    - Navajo: C_C(CVVO) > C_C(CVR).
  - the data are consistent with the more specific generalization based on actual $C_{CONTOUR}$.

Zhang - contour tones

- Is there a problem for a model that allows independent restrictions *Contour/ShortV *Contour/Closed Syll, regardless of detailed duration patterns?
- predicts languages with the distributional patterns of Cantonese and Navajo but with duration patterns exchanged.
  - unattested in a sample of 3 languages.
- On the other hand, what would such a theory look like, and how would it account for the other (coarser-grained) generalizations?
  - if the distribution of contour tones is sensitive to duration, then why/how would it be sensitive to approximate, 'language-independent' durations rather than actual, language-specific durations?

Zhang - contour tones

- The existing alternative to the $C_{CONTOUR}$ analysis appeals to the mora as Tone Bearing Unit:
  - languages may require each tone to be associated to its own mora.
  - contour tones require two moras.
- Zhang's (2002, 2004) arguments against mora-based analyses:
  - Maximum of 2 or 3 moras motivated by analysis of syllable weight, but contour tone restrictions can distinguish a hierarchy of as many as four syllable types - arguably 4 levels of C_{CONT}.

Mende contour tone restrictions

Vowel length	No. of sylls in word	Syll position in word	L͡HL ok?	L͡H ok?	H͡L ok?
VV	1	final	yes	yes	yes
VV	>1	any	no	yes	yes
V	1	final	no	yes	yes
V	>1	final	no	no	yes
V	>1	non-final	no	no	no

Image by MIT OpenCourseWare. Adapted from Zhang, Jie. *The Effects of Duration and Sonority on Contour Tone Distribution: A Typological Survey and Formal Analysis*. New York, NY: Routledge, 2002.

Zhang - contour tones

Arguments against mora-based analyses:

- Rising tones are more restricted than falling tones although both consist of two tones.
  - e.g. Mende. Konni: HL can appear on final CV but LH cannot.
- Domain-final position is licensing position for contour tones (attributed to final lengthening).
  - Final lengthening does not add moras for purposes of stress (Hayes 1995).
  - Final lengthening often preserves length contrasts - i.e. appears to be a sub-moraic effect.
- Moraic inconsistency: if CVR is metrically, all CVCs are metrically heavy (usually). Contour tones are licensed by sonorous codas only.

Ranking of faithfulness constraints based on langage-specific phonetic realization

- Patterns of obstruent voicing in Japanese loan words motivate a distinction between Ident(+voice)$_{singleton}$ and Ident(+voice)$_{geminate}$.
- Ident(+voice)$_{singleton}$ >> Ident(+voice)$_{geminate}$
- This ranking follows from Steriade's P-map hypothesis: correspondence constraints against perceptually larger changes are ranked higher.
- Lower distinctiveness of voicing contrasts in geminates compared to singletons seems to be a result of a language-specific devoicing process applying to voiced geminates.
  - ranking of phonological constraints depends on language-specific phonetic properties.

Voicing in Japanese loanwords

- Consonants after lax vowels are often borrowed as geminates - also orthographic geminates.
- Voiced geminates optionally devoice iff preceded by a voiced obstruent.

		Possibility of Devoicing	Examples
TVDDV words	One voiced geminate	Impossible	[eggu] * [ekku] [webbu] * [weppu]
DVDV words	Two voiced singletons	Impossible	[dagu] * [daku]. * [tagu] [giga] * [kiga]. * [gika]
DVDDV words	One voiced singleton and one geminate	Possible	[doggu] ~ [dokku] [beddo] ~ [betto]

Image by MIT OpenCourseWare. Adapted from Kawahara, S. "A Faithfulness Ranking Projected from a Perceptibility Scale: The Case of [+voice] in Japanese." *Lanugage* 82, no. 3 (2006).

Analysis

- Devoicing motivated by OCP(+voice), a general restriction in native Japanese vocabulary

  OCP(+voice): Two voiced obstruents cannot cooccur within a single stem.

- Singleton voiced obstruents can cooccur in a stem: *bagii*

  $\text{Ident(+voi)}_{sing}$ >> OCP(+voi)
- OCP can motivate devoicing of geminates: *baggu* ~ *bakku*

  OCP(+voi) >> Ident(+voi)_{gem}
- Singleton and geminate voiced stops are possible:

  $\text{Ident(+voi)}_{sing}$ >> OCP(+voi) >> Ident(+voi)_{gem} >> *VoiObs

Markedness vs. faithfulness

- Motivation for distinguishing Ident(+voi)$_{sing,}$ Ident(+voi)$_{gem}$
- There is plausibly a constraint against *VoiObsGem
  - Voiced geminates present aerodynamic difficulties (Ohala etc)
  - A number of languages have voiced voiceless singleton stops but only voiceless geminate stops (t, d, tt, *dd), e.g. Yakut, Finnish, Selayarese (Podesva 2002).
- But OCP(+voi) and *VoiObsGem cannot account for the observed pattern assuming undifferentiated Ident(voi).
  - Ident(voi) must rank above both since voiced geminates (*eggu*) and pairs of voiced singletons (*bagii*) are possible.
- Only the combination of an OCP violation *and* a voiced geminate is problematic.
  - a markedness-based analysis has to posit a conjoined constraint:
    OCP(voi)&*VoiObsGem >> Ident(voi) >> OCP(voi), *VoiObsGem
  - conjoining constraints within the domain of the stem can derive unattested patterns.
  - faithfulness analysis is simpler, motivated by P-map (might also explain the repair - devoicing as opposed to degemination).

Markedness vs. faithfulness

- Excursus: Kawahara tentatively suggests that *VoiObsGem is unnecessary given Ident(+voi)$_{sing}$, Ident(+voi)$_{gem}$. Probably both constraints are required.
- Evidence: Buginese gemination is blocked where it would create a voiced geminate stop (Podesva 1998).
  - can only be explained by markedness, not faithfulness.

Voiceless Obstruents	ma**pp**asaʔ	'to go to the market'
	ma**tt**apa	'to smoke (meat)'
	ma**tʃtʃ**iruʔ	'to sharpen'
	ma**kk**ota	'to go to the city'
	massaraŋ	'to build a nest'
Voiced Obstruents	maʔ**b**usəʔ	'to eat (rude)'
	maʔ**d**eceŋ	'to ask for forgiveness'
	maʔ**d**ʒama	'to work'
	maʔ**g**oriʔ	'to scratch'
Sonorants	mam**m**usuʔ	'to go to war'
	mannasu	'to cook'
	ma**ɲɲ**awa	'to breathe'
	ma**ll**ebo	'to form into a ball'

Geminates at the prefix-root (*maC*+ROOT) boundary

Geminate markedness vs. faithfulness

/maC-lebo/ 'form into ball'		MAX	AGREE
(a)	maʔlebo		*!
(b)	→ mallebo		
(c)	malebo	*!	

/maC-goriʔ/ 'scratch'		MAX	*DD	AGREE
(a)	→ maʔgoriʔ			*
(b)	maggoriʔ		*!	
(c)	magoriʔ	*!		

- *DD = *VoiObsGem
- Low ranked $\text{Ident}(+\text{voi})_{\text{gem}}$ cannot explain the resistance of voiced stops to gemination.

P-map

- Kawahara argues that the ranking Ident(+voi)$_{sing,}$>> Ident(+voi)$_{gem}$ follows from a difference in perceptibility of voicing in singletons vs. geminates.
- NB there is no evidence for this ranking from native Japanese phonology.
- Based on the P-map hypothesis (Steriade 2001).

P-map

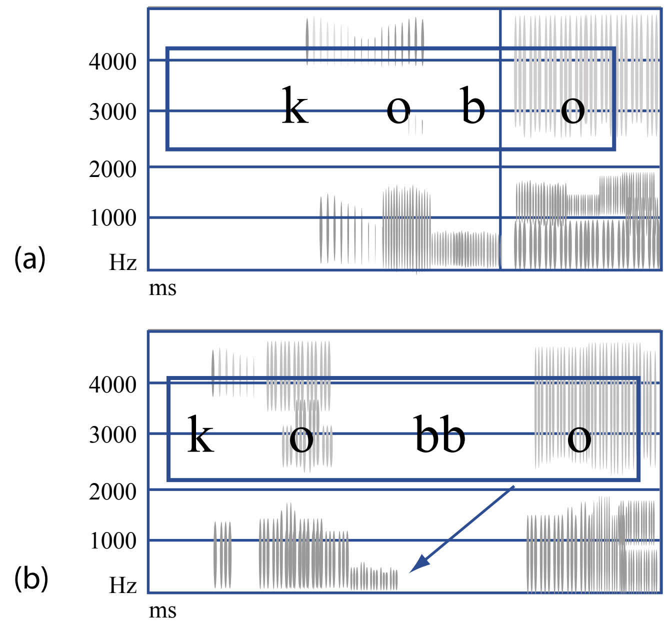
- ‘The aim, in any departure from the UR, is to change it *minimally* to achieve compliance with the phonotactics’
  - ‘minimal’ = perceptually minimal
- Necessary knowledge about perceptual similarity is encoded in the P-map:
  - ‘The P-map is a set of statements about absolute and relative perceptibility of different contrasts, across the different contexts where they might occur.’
  - e.g. [p]-[b] contrast is more distinct before [+son] than before [-son].
- The P-map ‘project[s] correspondence constraints and determine[s] their ranking’.

P-map

- Projecting correspondence constraints
  - For any two P-map cells, x-y/$\_K_i$ and w-z/$\_K_j$, associated with different confusability indices, there are distinct sets of correspondence conditions, Corresp. (x-y/$\_K_i$) and Corresp (w-z/$\_K_j$).
  - e.g. Ident(+voi)$_{sing}$, Ident(+voi)$_{gem}$
- Ranking correspondence constraints by relative distinctiveness
  - For any two P-map cells, x - y/ $\_K_i$ and w - z/ $\_K_j$, if x-y/$\_K_i$ ➢ w - z/ $\_K_j$ then any correspondence constraint referring to x - y/ $\_K_i$ outranks any parallel constraint referring to w - z/ $\_K_j$
  - E.g. Ident(+voi)$_{sing}$ >> Ident(+voi)$_{gem}$ if t-d ➢ tt-dd

Distinctiveness of voicing contrasts in Japanese

- Singleton stops are fully voiced, geminates are only partially voiced:



Spectrograms of a singleton [b] (a) and geminate [bb] (b).
Time scales are the same (350 ms).

Image by MIT OpenCourseWare. Adapted from Kawahara, S. "A Faithfulness Ranking Projected from a Perceptibility Scale: The Case of [+voice] in Japanese." *Lanugage* 82, no. 3 (2006).

Distinctiveness of voicing contrasts in Japanese

- Singleton stops are fully voiced, geminates are only partially voiced:

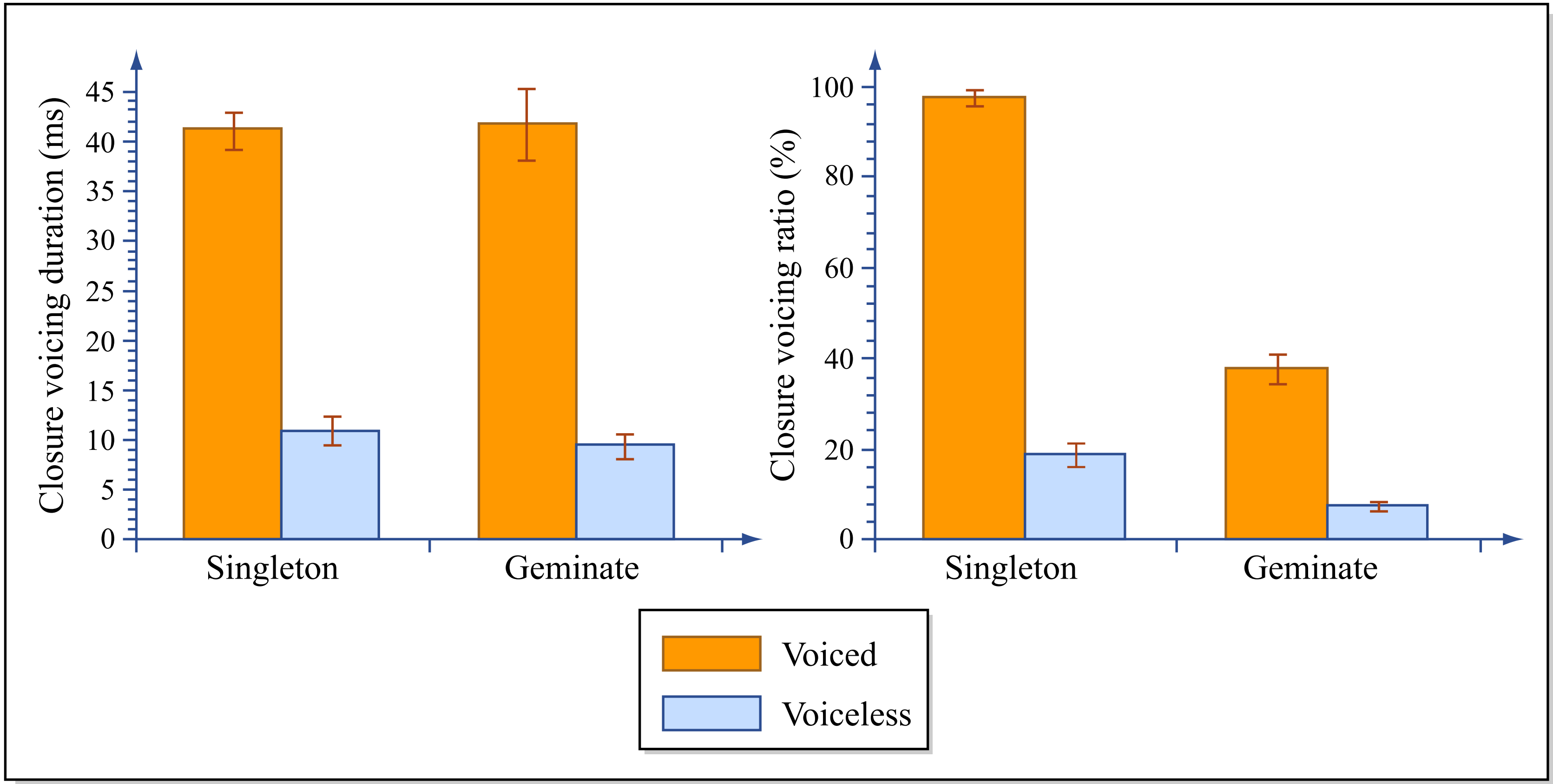


Image by MIT OpenCourseWare. Adapted from Kawahara, S. "A Faithfulness Ranking Projected From a Perceptibility Scale: The Case of [+voice] in Japanese." *Lanugage* 82, no. 3 (2006).

Distinctiveness of voicing contrasts in Japanese

- Perception test: kVC(C)V words in sentence frame
  - V = e, a, o, C = p, t, k, b, d, g, pp, tt, kk, bb, dd, gg
- Presented in 'cocktail party' noise
- 17 subjects, 2 alternative forced choice voiced vs. voiceless, presented in orthography (Katakana).
- measured d' for each subject

Useful background: detection theory

- d' is a detection-theoretic measure of perceptual distance between stimuli.
- measures perceptual sensitivity, independent of bias (predisposition to respond in a particular way).

References:

Green, D. M. and Swets, J. A. (1966) *Signal detection theory and psychophysics.*

MacMillan and Creelman (2004). *Detection Theory: A User's Guide*. 2nd Edition.

Heeger, D. (2003). http://www.cns.nyu.edu/~david/sdt/sdt.html

Some basic terminology

- Identification (e.g. voiced vs. voiceless)
- terminology only makes sense here if identification of [+voice] is taken to be the task.

		Response		
		+voice	-voice	Total
stimulus class	+voi	hits (20)	misses (5)	(25)
	-voi	false alarms (7)	correct rejections (18)	(25)

Some basic terminology

- Given two of the numbers from the contingency table the other can be calculated.
- $H = P($"+voi"$|$ +*voi*$)$ (hit rate)
- $F = P($"+voi"$|$ -*voi*$)$ (false alarm rate)

		Response +voice	-voice	Total
stimulus class	+voi	hits (20)	misses (5)	(25)
	-voi	false alarms (7)	correct rejections (18)	(25)

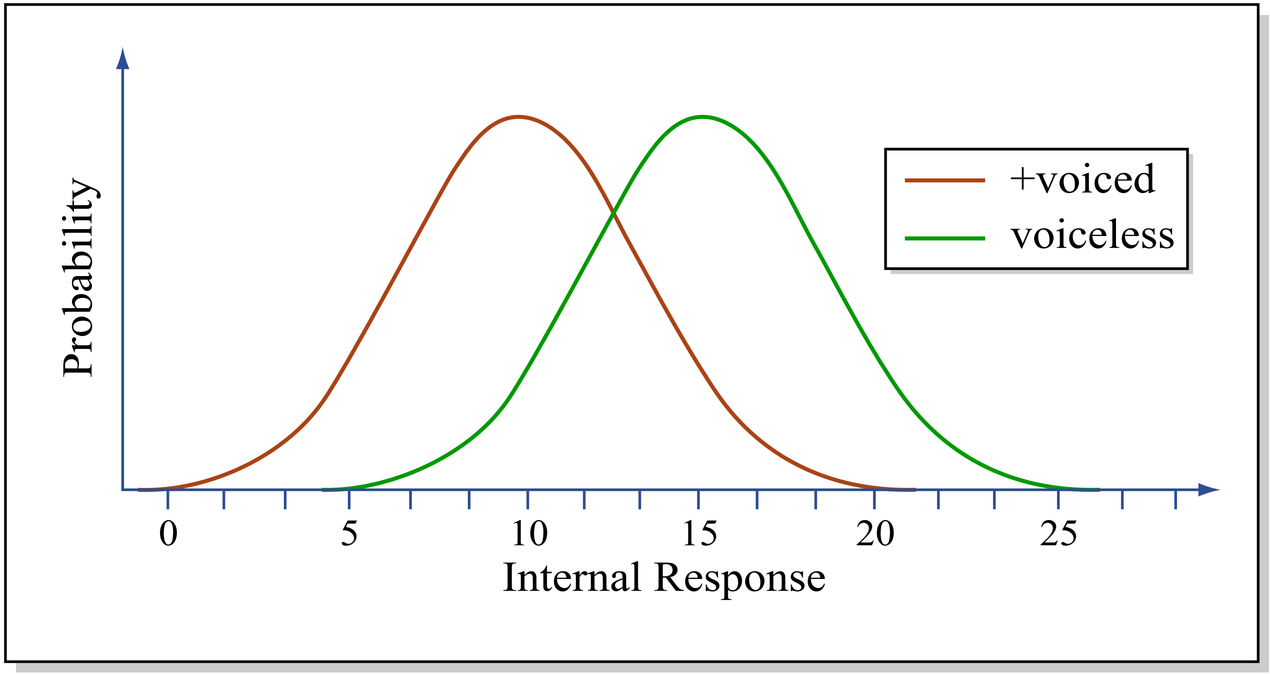
The Detection Theory model of perception tasks

Detection Theory distinguishes two components of tasks like identification and discrimination:

- Information acquisition/sensory operation: Stimulus is mapped onto a value on some internal variable (e.g. a value on a perceptual dimension).
- Decision: Select a response by comparing this internal response to a ‘criterion’ value.

Noise

- Sensory process is affected by internal and/or external noise, so multiple presentations of a single stimulus yield a distribution of perceptual values.



Bias

- Decision is made by establishing a criterion - a threshold value on the internal perceptual dimension.

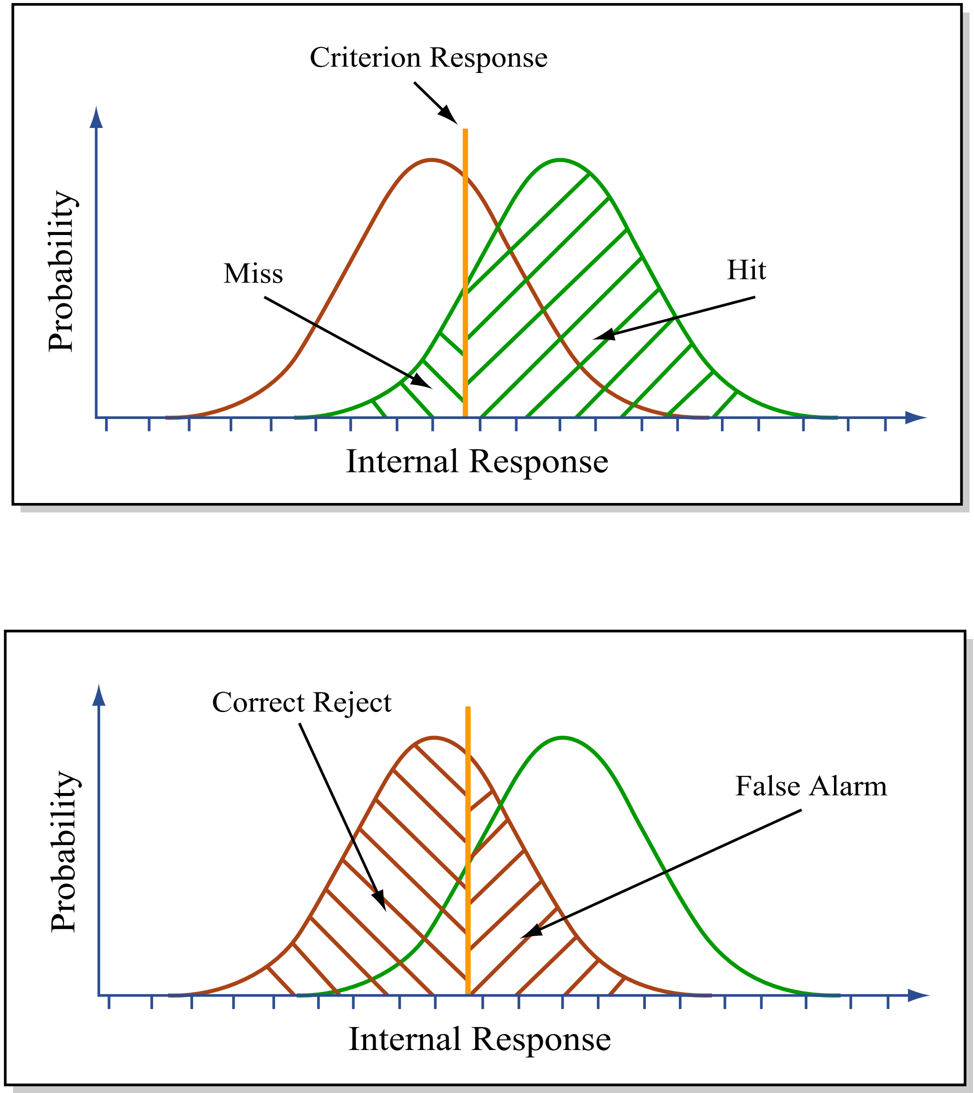


Image by MIT OpenCourseWare. Adapted from Heeger, David. "Signal Detection Theory." 1997.

Please see: "Signal detection (advanced)" handout on David Heeger's Web page.

Bias

- Decision is made by establishing a criterion - a threshold value on the internal perceptual dimension.
- Decision criterion can vary depending on *bias*, i.e. a predisposition to respond "voiced" or " voiceless".
  - There is a trade-off between *H* and *F*. The relative importance of each could influence choice of criterion (see applet).
  - Prior probability might influence choice of criterion (e.g. relative frequency of [p] and [b]).

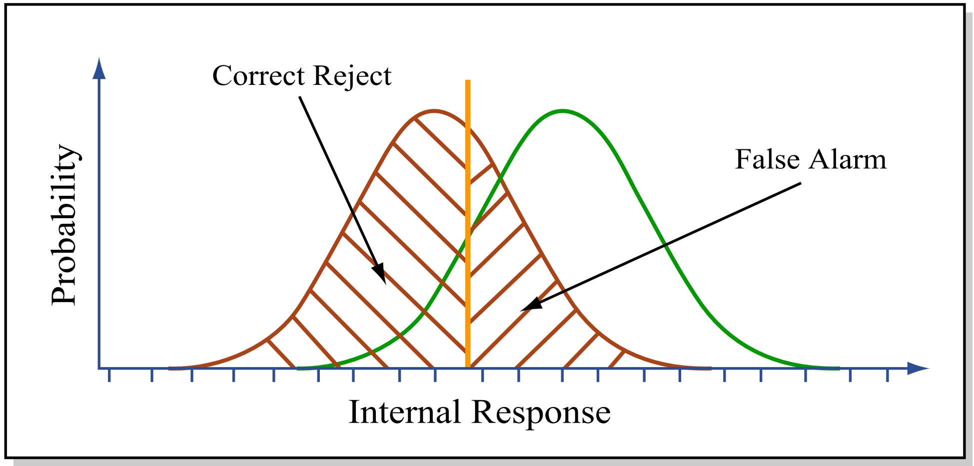


Image by MIT OpenCourseWare. Adapted from Heeger, David. "Signal Detection Theory." 1997. Also available as (PDF)

Sensitivity

- Crucially, *sensitivity* to the difference between stimuli classes depends on the two distributions, not on the criterion value. It is a measure of perceptual difference, independent of bias.

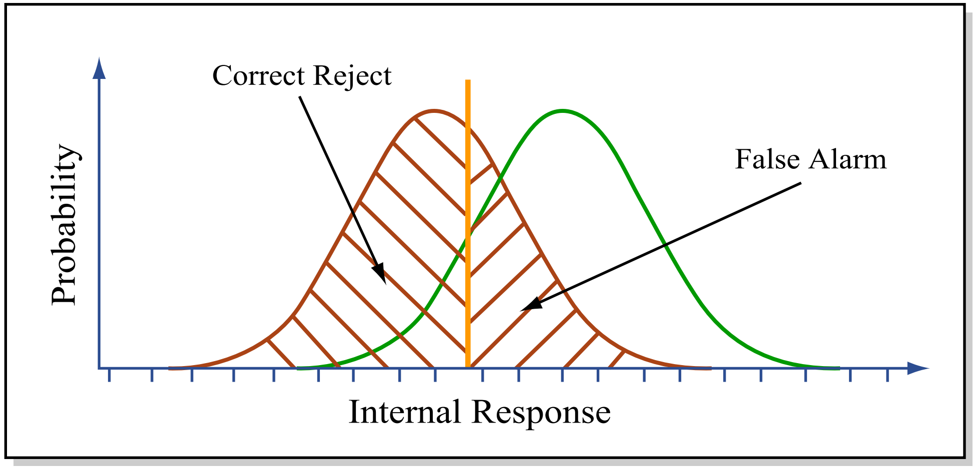


Image by MIT OpenCourseWare. Adapted from Heeger, David. "Signal Detection Theory." 1997. Also available as (PDF).

Measuring sensitivity - d'.

- d' = separation/spread

  = (mean(+*voi)* - mean (-*voi*))/s.d.

- $d' = z(H) - z(F)$
  - z is the inverse of the normal distribution
  - $z(p)$ is number of standard deviations from the mean at which probability of more extreme value is p.
  - assumes standard deviation (noise) is the same for each stimulus.

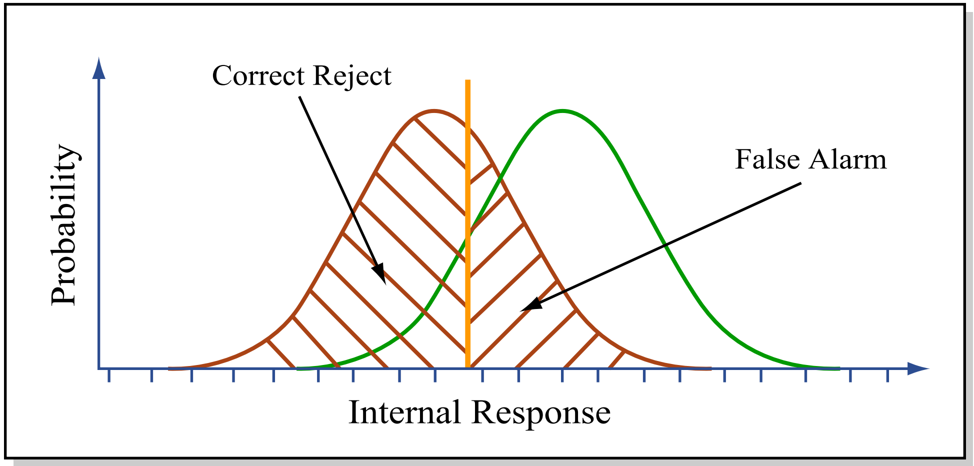


Image by MIT OpenCourseWare. Adapted from Heeger, David. "Signal Detection Theory." 1997. Also available as (PDF)

Measuring sensitivity - *d′*.

Properties of *d′*.

- Infinite when H=1, F=0 (perfect sensitivity)
  - Common to replace 0 and 1 with 1/(2N) and 1-1/(2N) respectively to avoid infinite values.
- 0 (lowest value) when H=F (NB this could be any value).

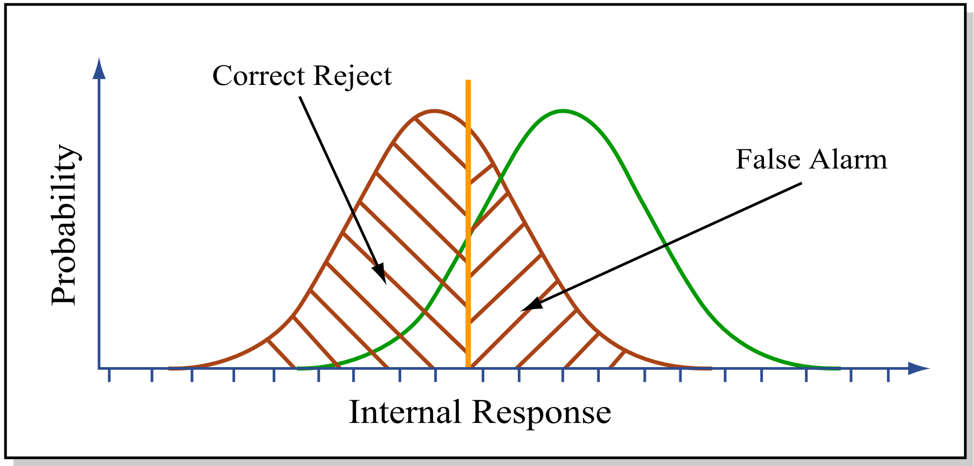


Image by MIT OpenCourseWare. Adapted from Heeger, David. "Signal Detection Theory." 1997. Also available as (PDF)

d' vs. percent correct

- Sensitivity measures such as d' are theoretically motivated measures of perceptual distance that control for bias
- *H* is obviously a bad measure of perceptual sensitivity because it could be 100%, but if *F* is 100% also, subject is obviously not sensitive to the distinction between voiced and voiceless.
- But percent correct - percentage of both voiced and voiceless stimuli that are labeled correctly – does not face the same problem. It would only be 50% if the subject labels all stimuli as voiceless (assuming equal numbers of each stimulus).
- The advantage of d' is that it is theoretically free of bias, whereas %correct can vary with bias, and d' is theoretically comparable across experimental paradigms.

Some limitations of detection theoretic analysis

- Assumptions about signal distributions.
- Detection theory is only easy to apply where stimuli can be regarded as lying on a single dimension.
  - Problematic: more than two stimuli, more than one dimension, e.g. identification of natural [ba, da, ga] stimuli.

Distinctiveness of voicing contrasts in Japanese

- Perception test: kVC(C)V words in sentence frame
  - V = e, a, o, C = p, t, k, b, d, g, pp, tt, kk, bb, dd, gg
- Presented in 'cocktail party' noise
- 17 subjects, 2 alternative forced choice voiced vs. voiceless, presented in orthography (Katakana).
- measured d' for each subject
  - T vs. D: mean $d' = 3.79$
  - TT vs. DD: mean $d' = 0.71$ (difference is significant)

Language-specificity

- Not all languages partially devoice voiced geminates, e.g. Egyptian Arabic.

Image removed due to copyright restrictions. Source: Kawahara, S. "A Faithfulness Ranking Projected from a Perceptibility Scale: The Case of [+voice] in Japanese." *Language* 82, no. 3 (2006).

- Not clear the ranking would ever be reversed, but might expect languages in which there is no difference between singletons and geminates in perceptibility of voicing.

Phonetics-phonology interaction

- If the P-map refers to phonological features like [+/-voice] (as opposed to [pː-bː], [pː-bp]), then the P-map and associated rankings must be language-specific.
- Must be able to determine the phonetic realizations of segments in context, independently of the phonology.
  - NB relevant realizations are often non-occurring. E.g. in Steriade's analysis of final devoicing, the acceptability of final devoicing depends in part on the phonetic similarity between non-occurring [ab] and [ap].
- If P-map specifies similarity between detailed representations then the relevant portion can be universal
  - [p-b], [pː-bː] > [pː-bp], *100%voi→5%voi >> *40%voi →5%voi

Phonetics-phonology interaction

- But this line of analysis faces a 'Richness of the Base' problem:
  - *100%voi→5%voi >> OCP(voi) >> *40%voi →5%voi
  - Works if input /baggu/ is actually [bagku], but [baggu] will not devoice.
  - Given that there is variation in this case, we could adopt a proposal from (2002) 'faithfulness among variants', [bagku] is the basic form, andvariant [bakku] is required to correspond to this surface form.
  - More generally it appears that some phonological constraints need to refer to the way in which an input *would* be phonetically realized.
    - /baggu/ can devoice because it would otherwise be realized as [bagku].
    - /pad/ can devoice to [pat] because the realization [pad] would be insufficiently distinct from the realization of /pat/.

References

- Kawahara, S. (2006). A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* 82.3.
- Wetzels, W.L., and J. Mascaró (2001) The typology of voicing and devoicing.*Language* 77, 207-244.

24.964 Topics in Phonology: Phonetic Realization
Fall 2006