

MIT OpenCourseWare
<http://ocw.mit.edu>

24.910 Topics in Linguistic Theory: Laboratory Phonology
Spring 2007

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

24.910

Laboratory Phonology

Licensing by Cue

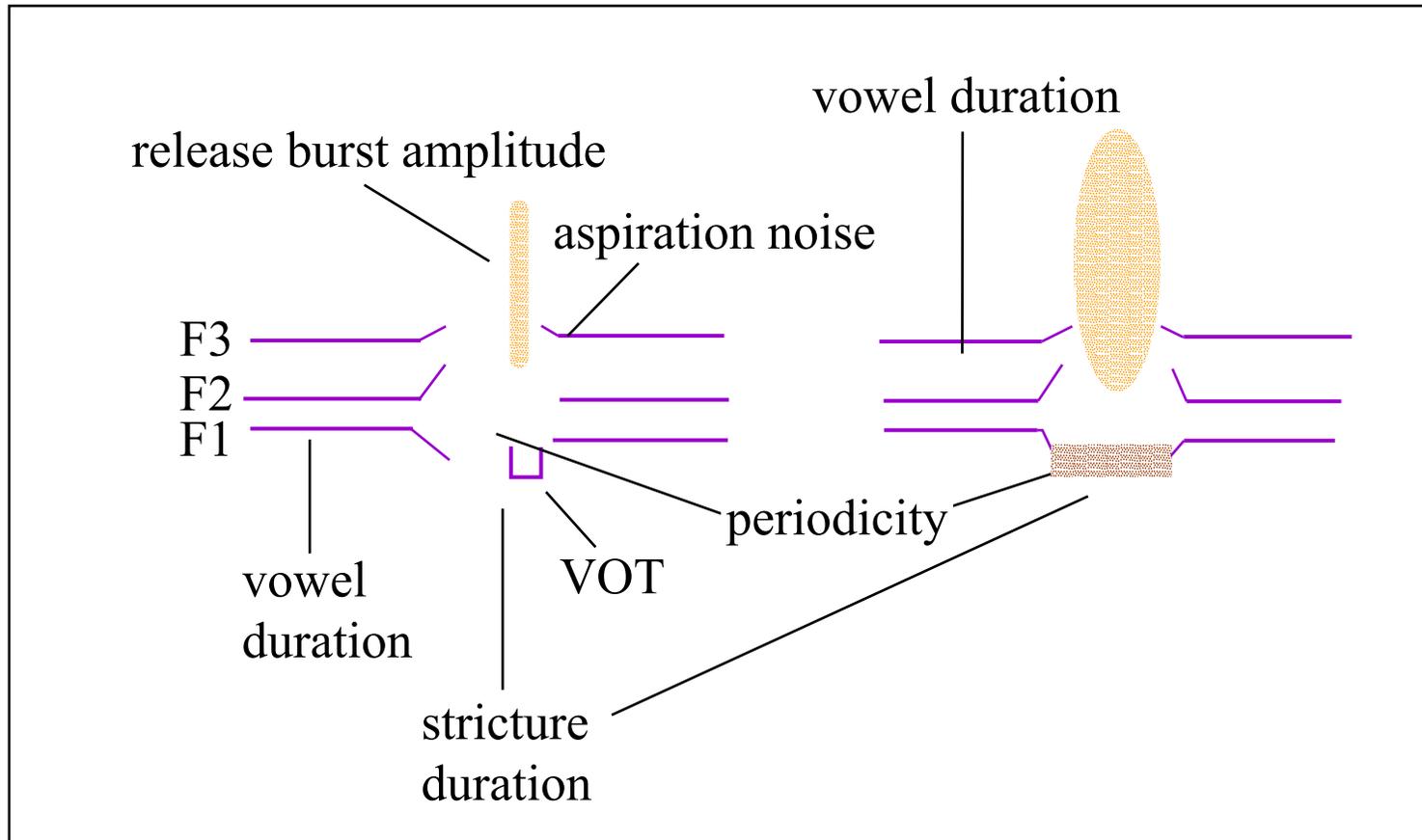


Figure by MIT OpenCourseWare.

Readings:

- ToBI tutorial, 1.0, 2.0, 2.1
- Johnson pp.28-31 (on pitch tracking).
- Ladd chapter 3

Assignment:

- Make stimuli for voicing perception experiment.

II. Perceptual cues and the distribution of phonological contrasts

- Phonological contrasts generally have restricted distributions.
- E.g. Lithuanian voicing contrasts
 - a. obstruent voicing is distinctive before vocoids and consonantal sonorants:

áuk**k**le nuk**n**iaũti auglingas dregna
sil**p**nas ryt**t**metỹs skob**n**is bã**d**metys

- b. obstruent voicing is neutralized (to voiceless) word-finally:

[dau**k**] [kat**t**]

- c. obstruent voicing is neutralized before any obstruent (assimilating in voicing to following obstruent):

a[**d-g**]al mè[**z-d**]avau d̄ir[**p-t**]i d̄è[**k-t**]i

II. Perceptual cues and the distribution of phonological contrasts

Different contrasts have different characteristic patterns of distribution (Steriade 1999):

- (i) Obstruent voicing contrasts are permitted only before sonorants
(e.g. German, Lithuanian, Russian, Sanskrit).
- (ii) Major place contrasts (labial vs. coronal vs. dorsal) are permitted only before vowels
(e.g. Japanese, Luganda, Selayarese).
- (iii) Retroflexion contrasts (retroflex vs. apical alveolar) are permitted only after vowels
(e.g. Gooniyandi, Miriwung, Walmatjari).

II. Perceptual cues and the distribution of phonological contrasts

Hypothesized explanation: ‘The likelihood that distinctive values of the feature F will occur in a given context is a function of the relative perceptibility of the F-contrast in that context’ (Steriade 1999).

- Contrasts differ in their distribution of cues so they are subject to different patterns of neutralization.
- Obstruent voicing is best cued by Voice Onset Time - only realized with a following sonorant.

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

O = obstruent, R = sonorant, inc. vowel

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



more cues to voicing available

The phonological analysis

- Russian voicing (like Lithuanian)

a.	vra[k]	‘enemy’ nom. sg.	vra[g]a	gen. sg.
b.	bra[t]	‘brother’ nom. sg.	bra[t]a	gen. sg.
c.	le[d]ok	‘ice’ nom. sg.	le[tk]a	gen. sg.
d.	pro[s’]it’	‘to beg’	pro[z’b]a	‘request’

- Rule-based analysis:

- /vrag/, /brat/, /led/, /pros^j/ /-a/

- final devoicing $[-\text{son}] \rightarrow [-\text{voice}] / _ \#$

- voicing assimilation $[-\text{son}] \rightarrow [\alpha\text{voice}] / _ \left[\begin{array}{l} -\text{son} \\ \alpha\text{voice} \end{array} \right]$

Optimality Theory

- Optimality Theory separates ‘problems’ from ‘solutions’
 - E.g. $*[+voice, -son]\#$ (provisional formulation)
 - $*[\alpha voice, son][-\alpha voice, -son]$
- These are Markedness constraints - they ban dispreferred sound sequences and configurations.
- The other basic kind of constraints are Correspondence constraints - they require output forms to be as similar to the input underlying form as possible.
- If a underlying form like /sled/ is realized as [sled], it would violate $*[+voi, -son]\#$.
- If it is changed to [slet] to satisfy the markedness constraint, that violates the correspondence constraint Ident(voice).

Optimality Theory

- Constraint conflict: Any output violates one constraint or the other:

/sled/		*[+voi,-son]#	Ident(voice)
	 sled	*	
	slet		*

- Conflict is resolved by reference to a ranking of the constraints: the higher ranked constraint prevails:

/sled/		*[+voi,-son]#	Ident(voice)
	sled	*!	
	 slet		*

Optimality Theory

Components of Optimality Theory:

- A set of constraints
 - Markedness constraints - define dispreferred configurations.
 - Correspondence constraints - penalize deviations from identity between input and output.
- Generation function: takes an input (e.g. /sled/) and generates all possible output candidates (sled, slet, sle, slen, led, let, etc).
- Evaluation function: given a ranked set of constraints, identifies the candidate that best satisfies the constraint ranking.
 - That's the actual output.

Optimality Theory

- Different rankings of the same constraints (different grammars) yield different outputs:

/sled/	Ident(voice)	*[+voi,-son]#
 sled		*
slet	*!	

/sled/	*[+voi,-son]#	Ident(voice)
sled	*!	
 slet		*

Optimality Theory and Typology

- It is hypothesized that most constraints are universal - i.e. the same in every language.
- Languages differ in the ranking of those constraints.
- So cross-linguistic similarities in phonological systems result from shared constraints.
- Where do these universal constraints come from?
- Hypothesis: Most constraints arise from:
 - basic requirements for rapid, robust communication
 - limitations of speech production and perception apparatus
 - cognitive limitations
- Common to all languages.

Typology of Voicing Neutralization

- Basic requirement for rapid, robust communication: perceptually distinct contrasts.
 - cf. analysis of vowel inventories.
- Steriade (1997, 1999) argues that perceptual considerations shape the typology of voicing neutralization:
 - Less distinct contrasts are dispreferred.
 - Specifically: contrasts are neutralized first in environments where they would be less distinct.

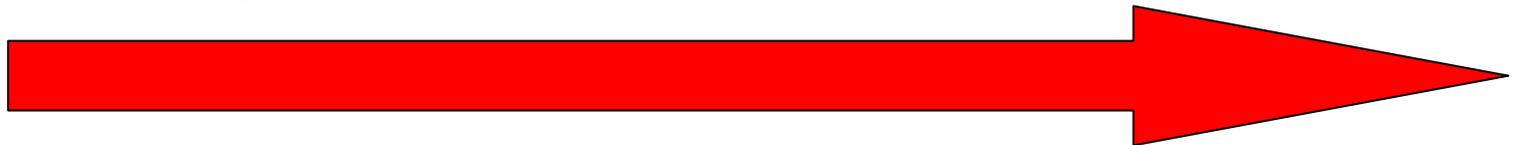
Steriade (1997) - obstruent voicing

- 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

O = obstruent, R = sonorant, inc. vowel

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



more cues to voicing available

Steriade (1997) - obstruent voicing

- Implementation: constraints against obstruent voicing contrasts in different contexts.
- Ranked according to the strength of the cues available in that context (fixed ranking).

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

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

Analyses

- Russian:

*[αvoice]/_[-son] >> * [αvoice]/_# >> **Ident(voice)** >> *[αvoice]/_[+son]

- Hungarian:

*[αvoice]/_[-son] >> **Ident(voice)** >> * [αvoice]/_# >> *[αvoice]/_[+son]

	*[αvoice]/_#	Ident(voice)	*[αvoi]/_[+son]	*[+voi, -son]
sled-slet	*!			
 slet		*		
sled		*		*!

	*[αvoice]/_#	Ident(voice)	*[αvoi]/_[+son]	*[+voi, -son]
 sleda-sleta			*	*
slet		*!		
sled		*!		*

Hungarian

- Contrast before sonorants

ved**m**eg ‘buy it!’ a:tmeɟ ‘to cross’

- Contrast word-finally

rɔ**b** ‘prisoner’ kɔlɔ**p** ‘hat’

ha:**z** ‘house’ kɛrt ‘garden’

- Neutralization before obstruents

ha:s-to:l ‘from the house’ kɛr**d**-bɛ ‘in the garden’

*-z**t**- *-t**b**-

Hungarian

- Russian:

*[αvoice]/_[-son] >> * [αvoice]/_# >> **Ident(voice)** >> *[αvoice]/_[+son]

- Hungarian:

*[αvoice]/_[-son] >> **Ident(voice)** >> * [αvoice]/_# >> *[αvoice]/_[+son]

	Ident(voice)	*[αvoice]/_#	*[αvoi]/_[+son]	*[+voi, -son]
☞ ha:z-ha:s		*		*
ha:s	*!			
ha:z	*!			*

	Ident(voice)	*[αvoice]/_#	*[αvoi]/_[+son]	*[+voi, -son]
☞ edme-etme			*	*
etme	*!			
edme	*!			*

Neutralization with assimilation

- *TD: *[-voice][-son, +voice]
 - It is more difficult to initiate voicing during an obstruent than to maintain voicing during an obstruent.
 - Universal ranking: *TD >> *[-son, +voice]

	*[αvoi]/ _[-son]	Ident(voi)	*[αvoice]/ _#	*TD	*[+voi, -son]
kεrtbe - kεrdbε	*!			*	***
kεrtbe		*		*!	*
 kεrdbε		*			**

Speech perception and phonology

- Steriade's analysis of the typology of obstruent voicing depends on the ranking of *[voice] constraints.
- This is supposed to follow from the relative strength of cues to voicing available in each context.
- What are these cues? How do we know how strong they are?

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

Speech perception

- The problem faced by the listener: To extract meaning from the acoustic signal.
- This involves the recognition of words, which in turn involves discriminating the segmental contrasts of a language.
- Much phonetic research in speech perception has been directed toward identifying the perceptual cues that listeners use.

Speech perception

- Production studies can reveal many differences between minimal contrasting words, e.g. contrasting vowels of English differ in formant frequencies and duration.
 - Are listeners sensitive to these differences in speech perception?
 - What is the nature of the perceptual representations of speech?
- These questions are addressed through perceptual experiments (cf. Johnson p.70).
- Most direct test of perceptual significance of an acoustic property: manipulate the acoustic property synthetically and see if perceptual response is affected. E.g. vary formant frequencies in synthetic vowels, and have subjects categorize the vowels.

Cues to vowel quality

- The main cues to vowel quality are related to the frequencies of the first two or three formants.

Cues to consonant contrasts

- Place cues (Wright, Frisch and Pisoni 1999)

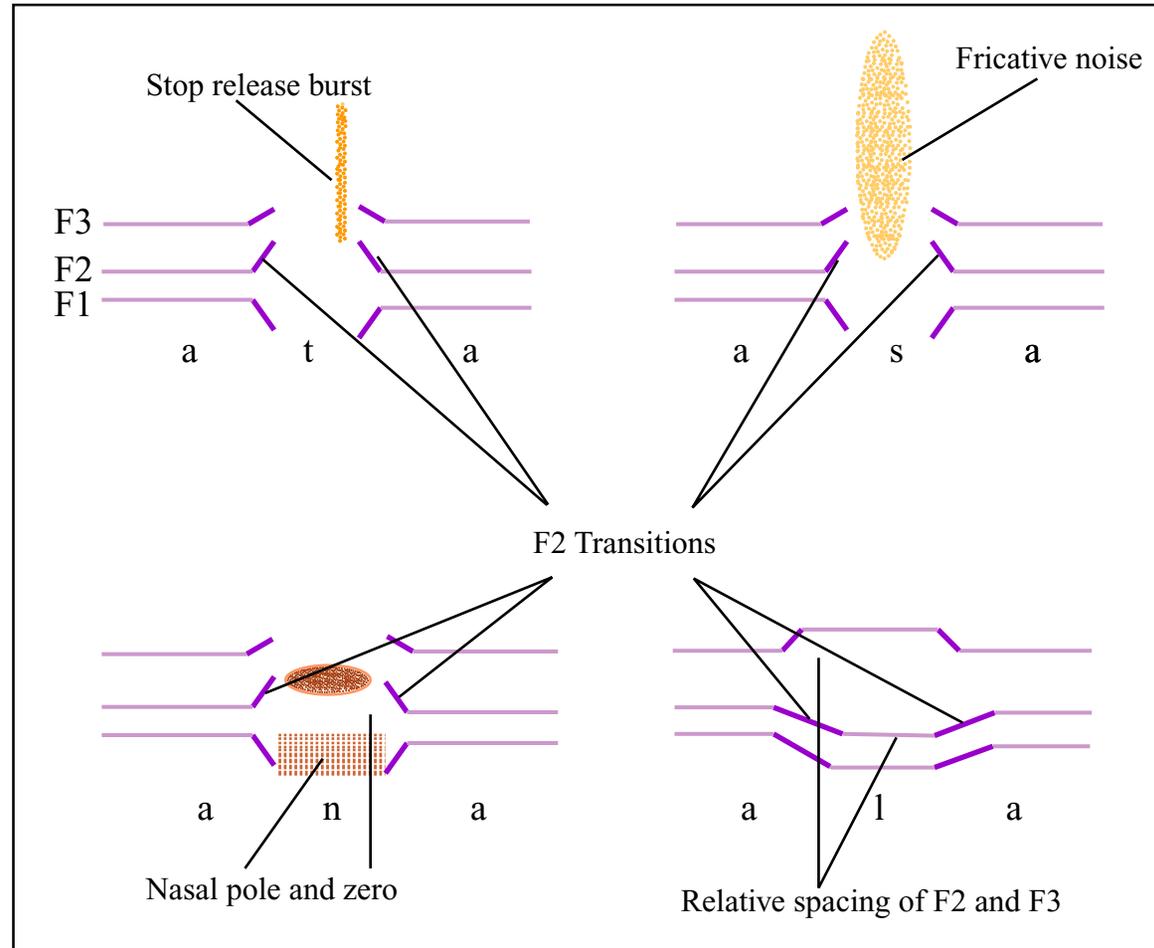


Figure by MIT OpenCourseWare. Adapted from Wright, R., S. Frisch, and D. B. Pisoni. "Speech Perception." In *Wiley Encyclopedia of Electrical and Electronics Engineering*. Vol. 20. New York, NY: John Wiley and Sons, 1999, pp. 175-195.

Cues to consonant contrasts

- Manner cues (Wright, Frisch and Pisoni 1999)

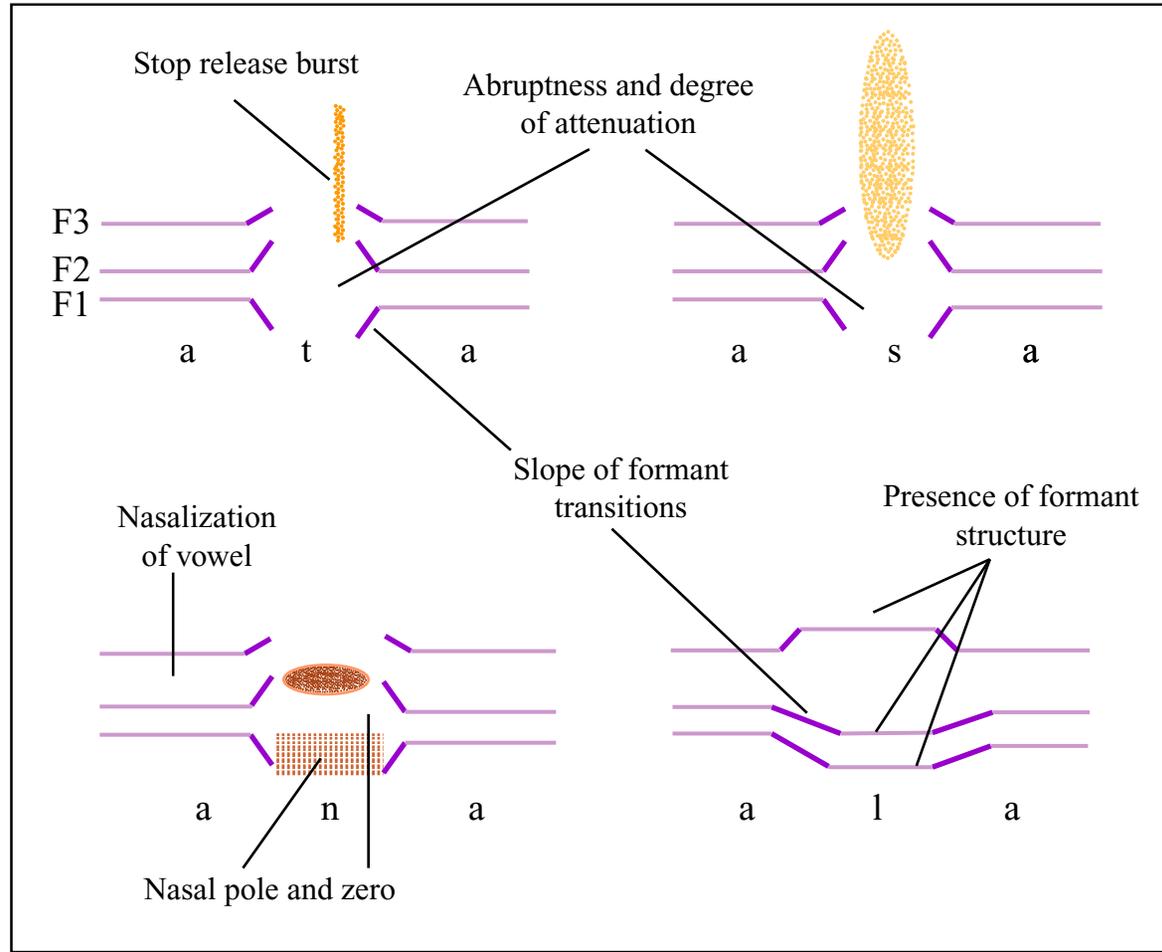


Figure by MIT OpenCourseWare. Adapted from Wright, R., S. Frisch, and D. B. Pisoni. "Speech Perception." In *Wiley Encyclopedia of Electrical and Electronics Engineering*. Vol. 20. New York, NY: John Wiley and Sons, 1999, pp. 175-195.

Cues to consonant contrasts

- Obstruent voicing cues (Wright, Frisch and Pisoni 1999)

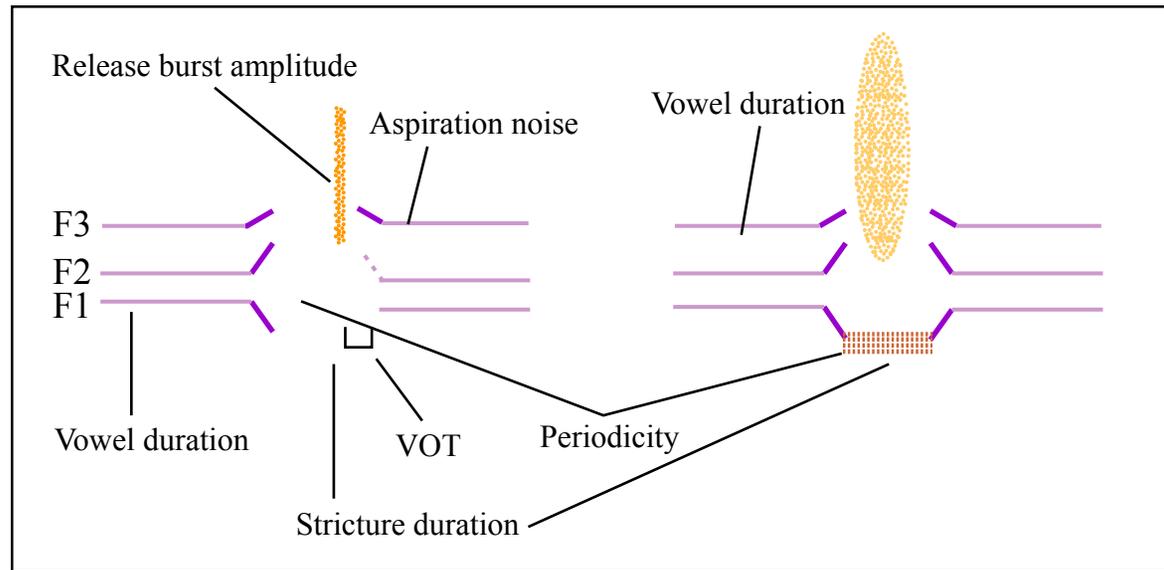


Figure by MIT OpenCourseWare. Adapted from Wright, R., S. Frisch, and D. B. Pisoni. "Speech Perception." In *Wiley Encyclopedia of Electrical and Electronics Engineering*. Vol. 20. New York, NY: John Wiley and Sons, 1999, pp. 175-195.

The nature of acoustic cues

- There are multiple cues to every contrast – the speech signal is highly redundant.
- E.g. stop voicing in English
 1. Low-frequency spectral energy, periodicity (Stevens and Blumstein 1981:29)
 2. Voice onset time (Lisker 1975)
 3. Amplitude of aspiration (Repp 1979)
 4. Amplitude of release burst (Repp 1979)
 5. Closure duration (Lisker 1957)
 6. Duration of the preceding vowel (Massaro and Cohen 1983)
 7. F1 adjacent to closure (Lisker 1975, Raphael 1972)
 8. f0 adjacent to the closure (Haggard, Ambler and Callow 1970)
 9. Amplitude of F1 at release (Lisker 1986).

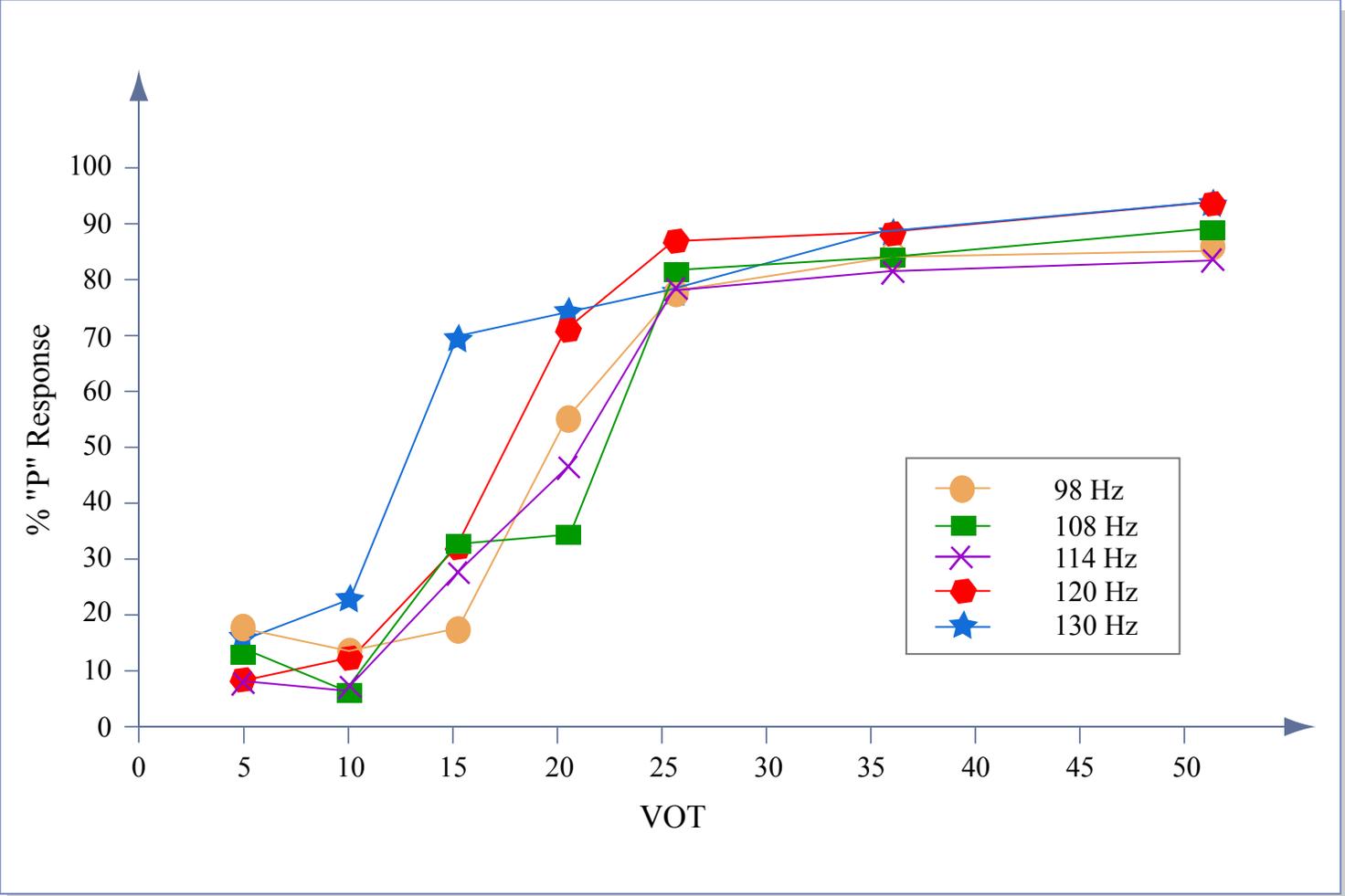


Figure by MIT OpenCourseWare.

Whalen, Abramson, Lisker & Mody (1993).

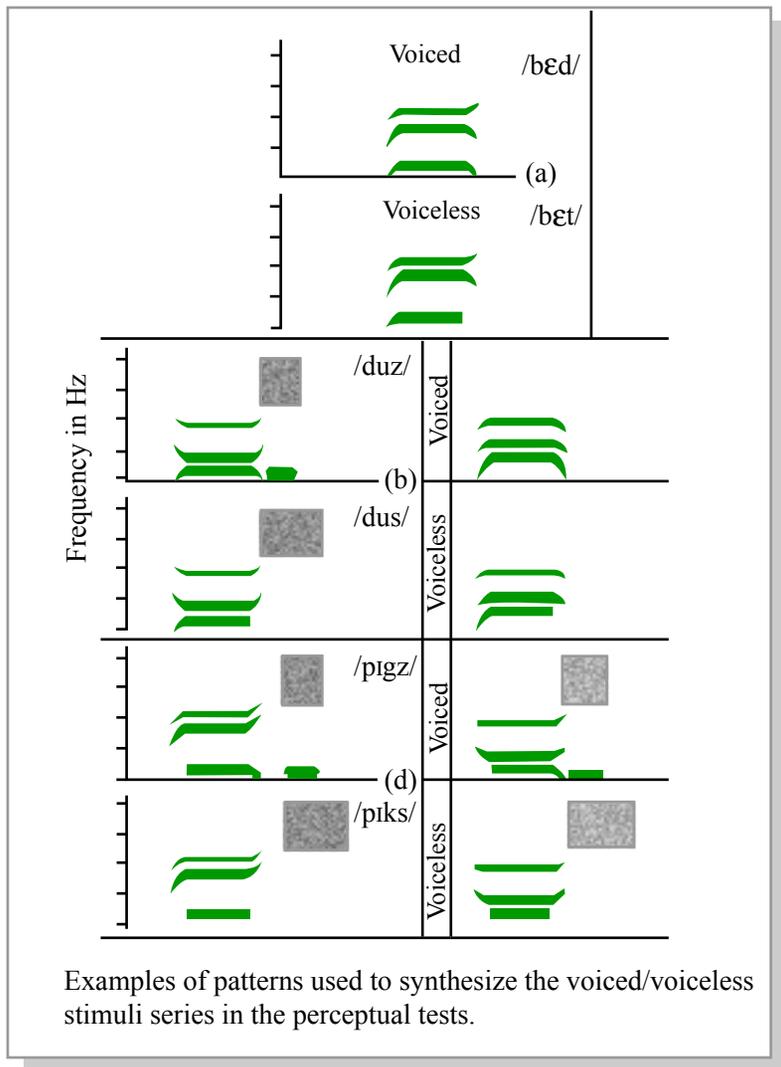


Figure by MIT OpenCourseWare.

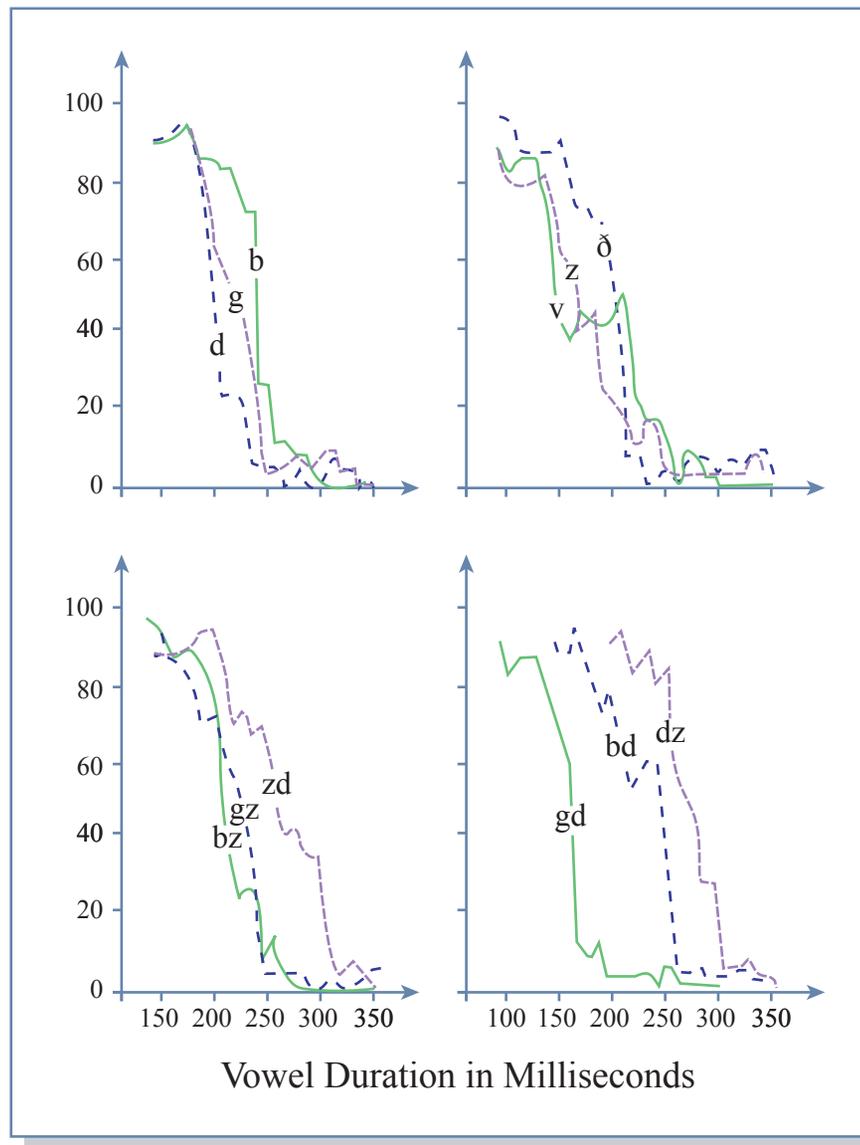


Figure by MIT OpenCourseWare.

Raphael (1972) *JASA*

The distribution of acoustic cues

- Cues to a contrast are temporally distributed and cues to more than one contrast may be present in the signal simultaneously (i.e. no strict segmentation).
- The availability and nature of the cues to a given contrast type vary systematically with context.
 - This observation is central to ‘licensing by cue’ analyses of the distribution of phonological contrasts.
- Broad distinction between internal and external cues to a contrast:
 - Internal: cues realized during the segment itself, e.g. vowel formants, fricative noise.
 - External: cues realized on adjacent segments, e.g. VOT, formant transitions.
- External cues typically depend on the presence of particular segment types in the context, e.g. VOT requires a following voiced sonorant.
 - So the presence of external cues can be highly variable across contexts.

Steriade (1997) - obstruent voicing

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

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

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

Assumptions:

- More cues are better.
- Release cues are stronger than preceding/closure cues.

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

O = obstruent, R = sonorant, inc. vowel

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

Assessing relative cue strength

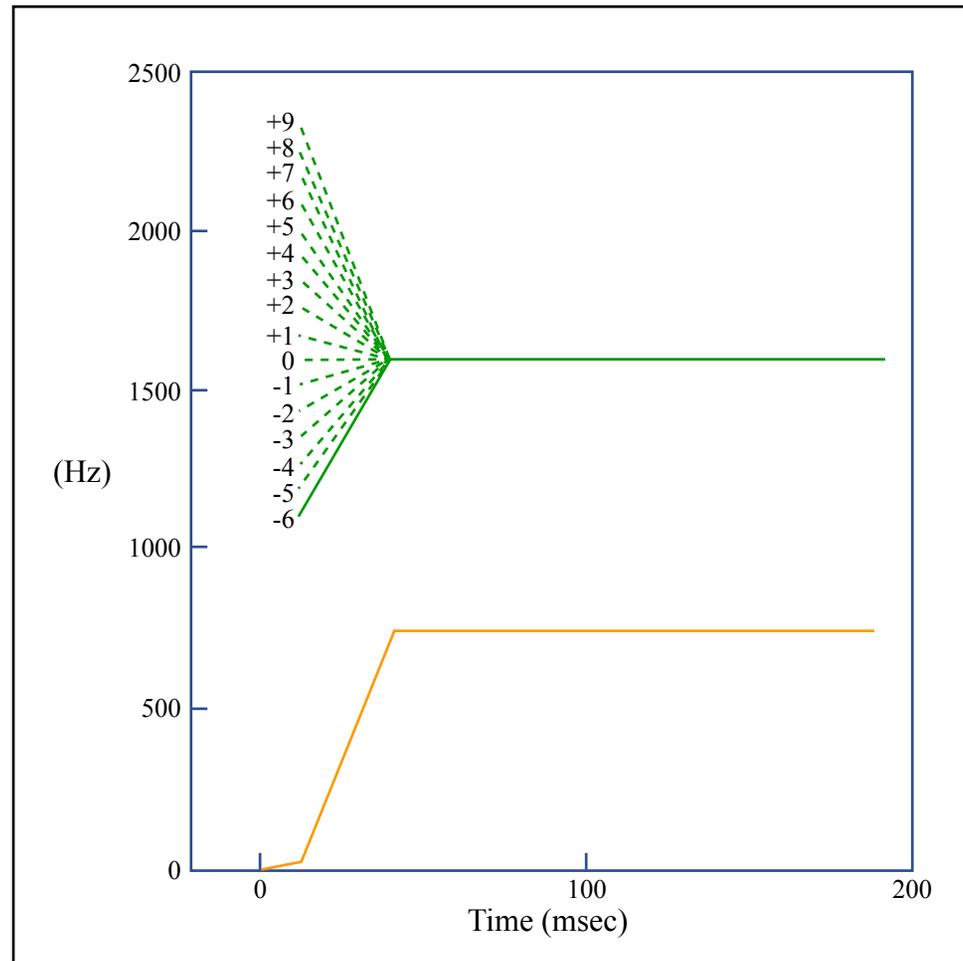
- Compare discrimination/identification of contrasting sounds in different contexts under the same conditions.
 - E.g. Wright (2003) compares perception of [b][d][g] before vowels and after vowels in noise [ba/da/ga], [ab/ad/ag].
 - We want to know about [ba] vs. [pa], [ab] vs. [ap].
- Conflicting cues: Create stimuli with conflicting cues to a contrast and see which prevails.
 - E.g. Fujimura et al: cross-spliced [ab-] from [aba] and [-da] from [ada]
 - closure transitions cue [b], release transitions+burst cue [d]
 - What do listeners perceive? I.e. which cues dominate?
- Try this for voicing in stops and fricatives.
 - create four stimuli each, 2 stops, 2 fricatives.
 - send the sound files to me.

Categorical perception

- Strict categorical perception is said to occur where discrimination performance is limited by identification performance, i.e. listeners only have access to category labels, so stimuli can only be distinguished if they are identified as belonging to different categories.
- Tested in two stages:
 - Identification of a synthetic continuum
 - Discrimination of stimuli from the continuum

Categorical perception

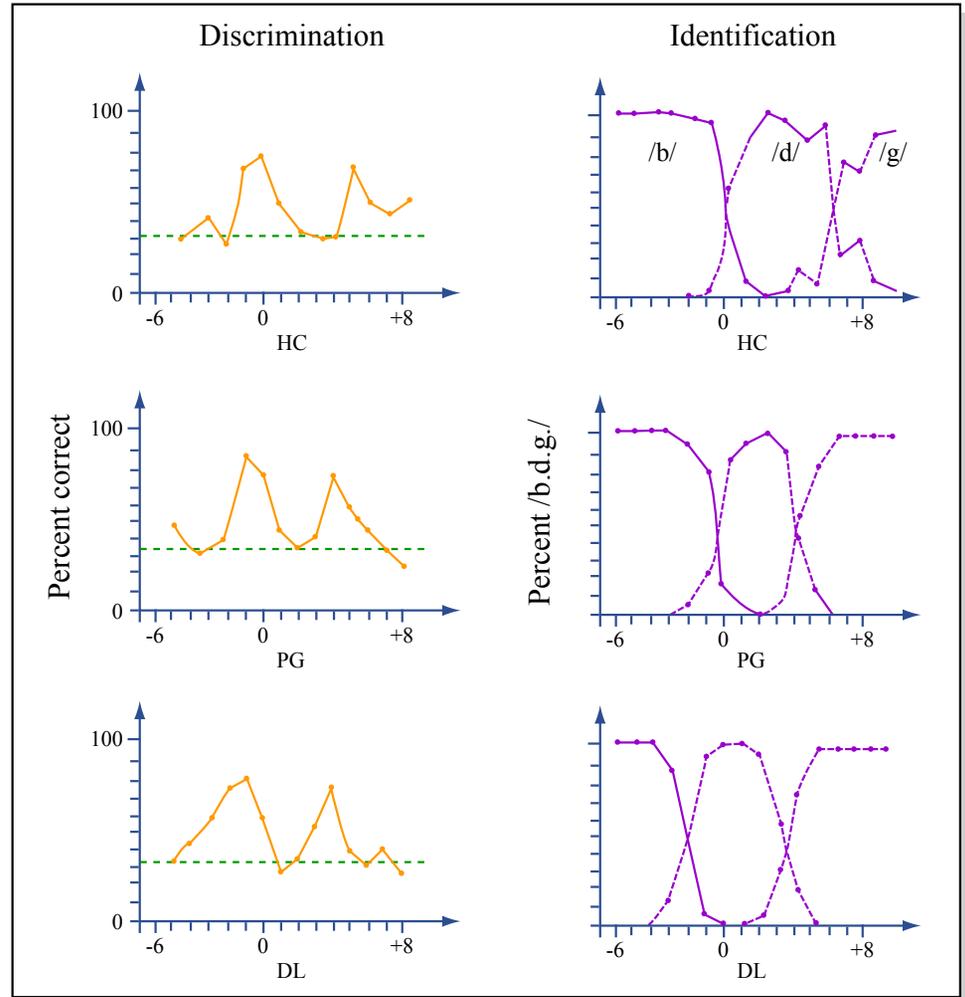
- E.g. Liberman (1970) place of articulation F2 transition continuum, b-d-g.



Categorical perception

- Identification: Subjects identify stimuli as b, d, g
- Discrimination: Subjects are presented with pairs of stimuli and asked to judge whether they are the same or different.

- Relatively abrupt transitions in identification functions.
- Peaks in discrimination function at the category boundary



Categorical perception

- Discrimination has never been found to be precisely predictable from identification - Discrimination is always better than predicted.
- More loosely, categorical perception is sometimes said to be exhibited where there is a discrimination peak at the category boundary determined by identification, even if the relationship is not precisely as predicted.
- A sharp transition in the 'identification function' for a stimulus continuum is *not* categorical perception in any technical sense.

Why is categorical perception significant?

- The (loose) categorical perception pattern contrasts with the pattern observed in psychophysical experiments using non-speech stimuli:
 “Typically, nonspeech stimuli that vary acoustically along a single continuum are perceived continuously, resulting in discrimination functions that are monotonic with the physical scale” (Luce and Pisoni, p.31).
- This contrast was used by Liberman and others to argue that speech perception is ‘special’ – i.e. it uses special mechanisms, not the general mechanisms of non-speech auditory perception.

Why is categorical perception significant?

- Vowels are not usually perceived categorically, even in the loose sense (Luce and Pisoni and refs there).
- The argument for specialness from categorical perception has been weakened by:
 - Evidence for categorical perception of non-speech sounds (noise-buzz, Miller et al 1976).
 - Evidence that Chinchillas perceive a VOT continuum categorically (Kuhl and Miller 1975).
- Recent data from eye-tracking studies suggest that perception may depend gradiently on the speech signal - categorical effects are introduced by decision processes (McMurray et al 2002).