Linguistic Theory/MA lecture

The role of speech perception in sound patterns

Zoltán G. Kiss

Dept. of English Linguistics, ELTE

topic outline

- relationship between sound patterns and speech perception
- laryngeal/voicing contrast & neutralization
- traditional, syllable-based approach
- cue based approach: Licensing by Cue

compulsory text

Donca Steriade (MIT):
Phonetics in Phonology: The Case of Laryngeal Neutralization (1997)

terms

- syllable, syllabification (onset, nucleus, coda, rhyme)
- licensing of contrast
- neutralisation
- cue
- VOT

licensing: by cue or by prosody?

An example voicing neutralization pattern (e.g., Polish, Lithuanian, Slavic, Sanskrit)

- Obstruents are distinctively voiced or voiceless before vowels and consonantal sonorants.
- Obstruents are neutralized (devoiced) word finally.
- Obstruents are neutralized before any obstruent: they surface assimilated in voicing to the following obstruent.

licensing: by cue or by prosody?

The prosodic/syllable based analyses:

a. The [voice] feature is unlicensed in the coda, licensed in onset.
b. The [voice] feature is licensed in a segment by a following **tautosyllabic** sonorant.

licensing: by cue or by prosody?

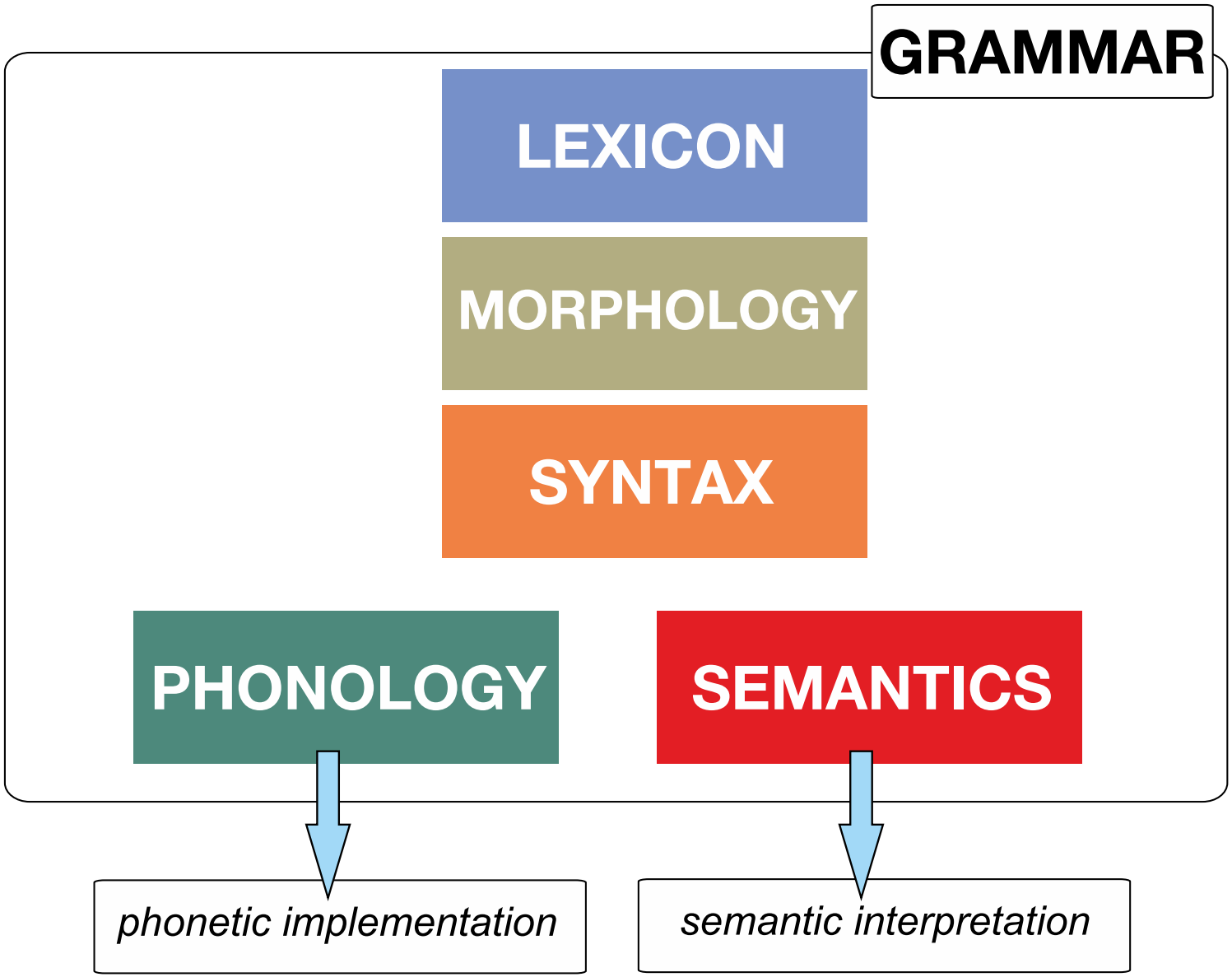
An alternative view: **licensing by cue**

- one of the major cues to the distinction between voiced/voiceless obstruents is the **voice onset time (VOT)** value observable on a following segment
- different VOT values – indicating different [voice] categories in the preceding obstruent – can be observed on a following vowel or sonorant but not on obstruents
- pre-obstruent obstruents necessarily lack at least this one bit of information about their laryngeal category
- word-finally: the VOT cue is missing (nothing follows)

licensing by cue: the gist

- absence of a major cue – or articulatory difficulties in implementing it – represent the main factor responsible for voicing (and other) neutralization
- this analysis promises to explain the grammar of neutralization, by showing how independently known facts about the **perception** and **production** of speech interact with grammatical conditions to yield sound patterns

traditional generative model of grammar



traditional role of phonology and phonetics

- **phonological component** consists of various entities and rules/conditions (feature set, sonority sequencing conditions, aspiration rule, etc.) whose interaction determines which contrasts a language will have and where
- **phonetic implementation component**: laws that map phonological representations onto articulatory instructions, and laws that compute the acoustic and perceptual consequences of articulatory gestures
- downward arrow connecting phonology to phonetics = decision to have a contrast and have it in a specific position is taken in phonology
- phonology cannot be affected by external factors, i.e. by physical conditions under which the contrast will be implemented (articulation, perception)

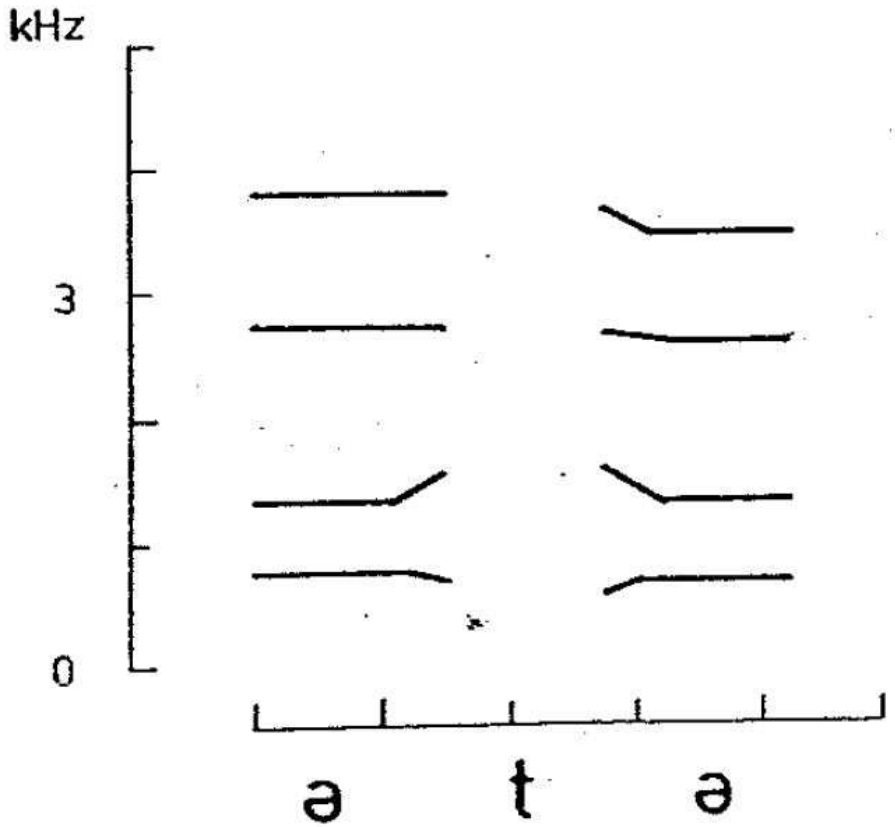
phonetically-grounded phonology

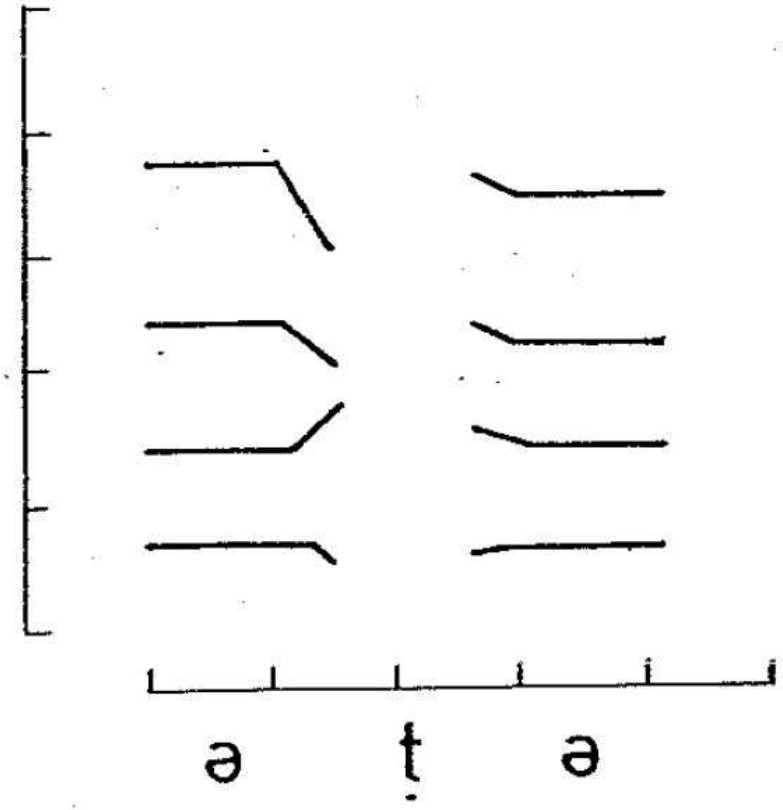
- speakers possess knowledge of the relative ease with which different types of contrasts can be implemented
- e.g., /k/–/g/ contrast is more easily detectable in intervocalic position than in inter-obstruent position
- knowledge of this sort enters the grammar in the form of **implementational constraints**, there is now an upward arrow from phonetics into phonology, too
- interaction of these conditions with the rest of grammar determines whether the language maintains a given contrast in a given position

a brief example: Gujarati apical contrast

- alveolar /t/ and retroflex /ʈ/ (/aʈa/) contrast between vowels: /ata/ ≠ /aʈa/
- but the apical contrast is neutralized word-initially and after consonants:
  /ta/ but */ʈa/ and /apta/ but */apʈa/

stylized spectrograms of Gujarati alveolar & retroflex stops





role of perception

- phonetic fact: only the V–C transitions differentiate /t/ and /ṭ/; the C–V transitions are essentially identical in the two cases
- phonological fact 1: if a language neutralizes the contrast between alveolars and retroflexes, then it does so first in **word-initial or post-consonantal position**
- = contexts where the **helpful V–C transitions are missing**
- phonological fact 2: if a language allows the contrast between alveolars and retroflexes, it does it **after vowels**
- = contexts where the **helpful V–C transitions are present**
- perception-based hypothesis: in contexts where the retroflex–alveolar contrast is **hard to perceive**, it is categorically suppressed, because **it would be difficult to implement there**

prosodic approach fails

- no connection between the syllabic or word-position of the consonant and its ability to carry distinctive retroflexion
- these neutralized word initial or postconsonantal stops are **all onsets** but then so is the distinctively retroflex intervocalic stops!
- there is a directly observable connection between the distribution of cues to a contrast, the phonetic implementation fact, and the phonological distribution of the contrast

from phonetic fact to phonological constraint

- the phonetic fact: retroflex contrast is difficult perceive in non-V-to-C transitions
- from this we can establish a phonological constraint in the grammar of Gujerati: 'ban apical contrast in contexts lacking V–C transitions'

relative perceptibility

- a contrast is directly related to the **relative perceptibility** of that contrast in a given phonetic position
- certain positions are optimal for the contrast, others less so
- optimal: lot of quality cues
- a correlation between positions of poor perceptibility and sites of neutralization

perception cues to obstruent voicing in various positions

1. after a V and before a sonorant: e.g. *a**b**ra*, *a**b**a*, *a**p**ra*, *a**p**a*
   cues: closure voicing, closure duration; V1 duration; F1 values in V1; burst duration and amplitude; VOT value; F0 and F1 values at the onset of voicing in V2
2. word initially or after an obstruent and before a sonorant: e.g. ***b**ra*, ***b**a*, ***p**ra*, ***p**a;* and *as**b**ra*, *as**b**a*, *as**p**ra*, *as**p**a*
   cues: closure voicing, closure duration (for post C obstruents only); burst duration and amplitude; VOT value; F0 and F1 values at the onset of voicing in the following V
3. after V at end of the word: e.g. *a**b***, *a**p***
   cues: closure voicing, closure duration; V duration; F1 values in V ; burst duration and amplitude

perception cues to obstruent voicing in various positions

4. after V and before obstruent: e.g. *a**b**sa*, *a**p**sa* cues: closure voicing, closure duration; V1 duration; F1 values in V1
5. between obstruents: e.g. *as**b**ta*, *as**p**ta*
   cues: closure voicing, closure duration
6. after an obstruent at the end of the word: e.g.: *as**b***, *as**p***
   cues: closure voicing, closure duration
7. before obstruent word initially: e.g.: ***b**sa*, ***p**sa*
   cues: closure voicing, closure duration

ranking of positions based on cues

- as we go down the list of contexts (from 1. to 7.), the set of typically available cues to voicing progressively shrinks
- the positions where the identification of voicing categories emerges as the most difficult (4., 5., 6., & 7.) are in fact positions where such contrasts have seldom been documented
- the cases in 5., 6., 7. are highly significant: they involve obstruent clusters that are rather well attested, yet only one language – Khasi (India/Bangladesh) – is known to allow distinctively voiced obstruents in sequences like ***b****sa*
- cue-based approach: a single factor - relative poverty of cues - induces neutralization in all the contexts

ranking of positions based on cues

- Position 4, V__obstruent (*ab**s**a*): voicing of an obstruent can be identified more reliably due to the preceding V, and we can find more languages that maintain contrast here
- these languages do not preserve the voicing contrast in 5., 6., 7. but they do maintain it when the obstruent is either left or right adjacent to a vowel

ranking of positions based on cues

- Position 3, V__# (*ab*): is more favourable for the contrast than 4. bc. release burst here is possible
- in 4. V1O1O2V2, O2 is better/more strongly cued (a V is after it!) and so the categorization of O1 with respect to voice is likely to be **influenced by that of O2**
- related to this is the fact that voicing neutralization never occurs finally **without also occurring in pre-obstruent position**

cue-based voicing patterns

- the claim (again): there is a link between the relative likelihood of neutralization of a feature and the relative perceptibility of that feature in a given context
- voicing contrast will be maintained in some context as a direct function of the cues available there: all else equal, the better the cue package, the greater the likelihood of contrast preservation
- the sites of neutralization have **no uniform characterization in terms of prosodic (esp. syllabic) organization**
- **no language maintains the voicing contrast in a less informative context, unless it also does so in the more informative contexts**

patterns of voicing neutralization

(6) **Patterns of [voice] neutralization** (O = obstruent, R = sonorant, incl. vowel)

fewer cues <----------> **more cues**

	#_ 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 (Crawford 1964)	no voice contrast	no voice contrast	no voice contrast	no voice contrast	**contrast**
Lithuanian (Senn 1966)	no voice contrast	no voice contrast	no voice contrast	**contrast**	**contrast**
French (Dell 1995)	no voice contrast	no voice contrast	**contrast**	**contrast**	**contrast**
Shilha (Applegate 1958)	no voice contrast	**contrast**	**contrast**	**contrast**	**contrast**
Khasi (Nagaraja 1985)	**contrast**	(sequence missing)	**contrast**	**contrast**	**contrast**

cue quality, cue weighting

- quantity of cues matters
- but also the **quality**: e.g. C-V transition cues (onset cues) have primacy over V-C offset cues
- in V1O1O2V2 O2 determines the voicing of O1 (and not the reverse): regressive voicing assimilation (and not progressive)
- this correlates with the observation: the most common sites of voicing neutralization are __# and __O – where the onset cues are absent

scale of obstruent voicing perceptibility acc. to context

V_R ▶ V_# ▶ V_O ▶ {O_O, O_#, #_O}

▶ indicates that voicing in one context is more perceptible than in the context listed to its right

from perception to constraints

- the perceptibility scale is fixed, and a ranked constraint system can be generated from it:

(i) *αvoice / [-son] __[-son], [-son] __#, #__[-son]

↓

(ii) *αvoice/ V_ [-son]

↓

(iii) *αvoice/ V_ #

↓

(iv) *αvoice/ V_ [+son]

- e.g. (i) = 'no voicing in: between obstruents, after obstr. word-finally, word-initially before obstr.'

implicational hierarchy

- it is important to remember the **scale is fixed**
- e.g. no language is predicted where voicing is neutralized finally but not before obstruents – this would violate the fixed ranking, there would be a gap
- generally: no language is predicted in which voicing is licensed in a less informative context than the one where it is neutralized

Example: Lithuanian

- distinctive voicing preserved before sonorants:
  au**kl**e ‘governess’ – au**gl**inas ‘fruitful’
  vi**kr**us ‘sent’ – e**dr**us ‘glutton’
  sil**pn**as ‘weak’ – sko**bn**is ‘table’
  smagus ‘cheerful’ – **žm**ogus ‘man’
- voicing is neutralized word-finally:
  daug [dau**k**] ‘much’; kad [ka**t**] ‘that’
- voicing is neutralized before obstruents
  at-gal [a**dg**al] ‘back’, dirb-ti [dir**pt**i] ‘work-inf’
  spalva ‘colour’

Lithuanian: syllable or cues?

- 'voicing neutralization in the coda': *dau**g**, a**t**-gal, dir**b**-ti*
- but coda-based analysis fails
- syllabification of CC is **always** C.C, including obstr.+liquid clusters
- e.g. a**t.n**e.se, ir**k.l**as, rak**š.t**is, cy**p.l**ys, dum**p.l**es, dum**b.l**as, krem**b.l**ys, krem**z.l**e
- there is no voicing neutralization in all codas!

Lithuanian: licensing by cue

- there is no justification for characterizing the site of licensing or neutralization in terms of syllabic position
- there are licensed onsets: ***s****magus* 'cheerful' – ***ž****mogus* 'man' but also neutralized onsets: ***s****palva* 'colour'
- there are neutralized codas: *daug* [dau**k**] 'much' but also licensed codas: *au**k**.le* 'governess'– *au**g**.linas* 'fruitful'
- Licensing by Cue: voicing in Lithuanian obstruents is neutralized in all and only the positions where the main cues are missing

Example: Hungarian

- distinctive voicing preserved before sonorants:
  tréfa ‘joke’ – **dr**ukkol ‘cheer’
  pa**pl**an ‘duvet’ – a**bl**ak ‘window’
- distinctive voicing preserved word-finally:
  há**t** ‘back’– ká**d** ‘tub’; kala**p** ‘hat’ – ra**b** ‘prisoner’;
  mé**sz** ‘whitewash’ – mé**z** ‘honey’
- voicing is neutralized before obstruents
  kú**t-b**an [**db**] ‘in a well’, ká**d-t**ól [**tt**] ‘from a tub’

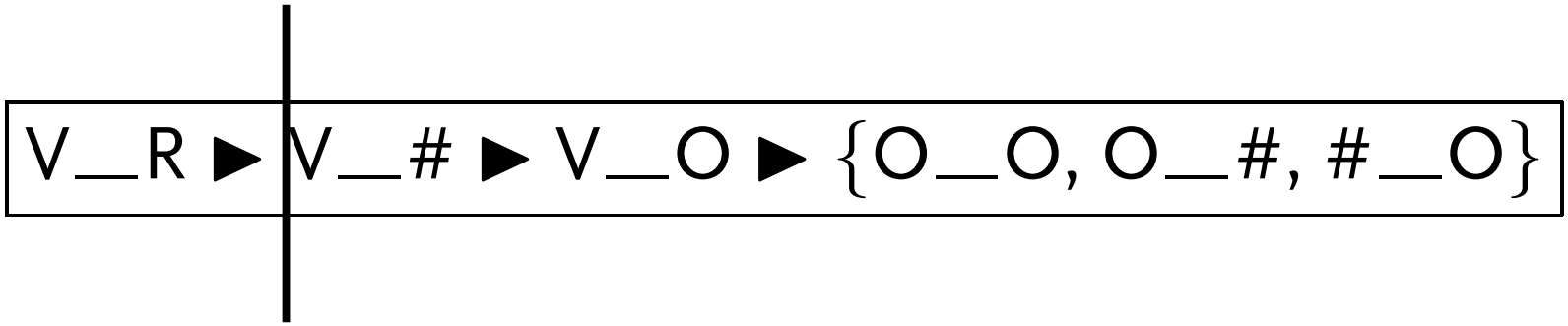
Hungarian: syllable or cues?

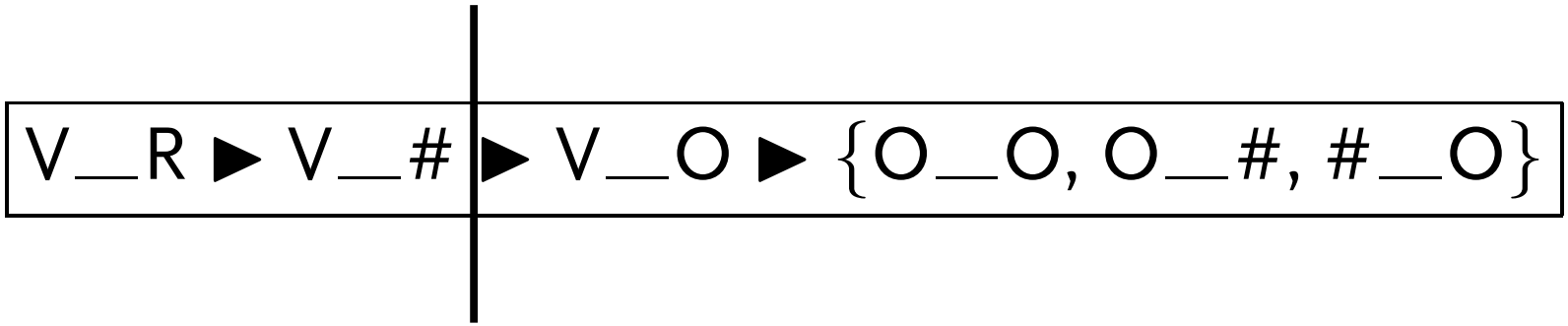
- 'voicing neutralization in the coda' fails in Hungarian because not all codas are neutralized
- word-final codas: *há**t*** – *ká**d***
- word-medial codas: *pa**p.**lan* – *a**b.**lak***
- only pre-obstruent codas are neutralized:
  *kú**t.**ban* [**db**] 'in a well', *ká**d.**tól* [**tt**]

Hungarian: licensing by cue

- the cue-based approach gives a uniform analysis
- voicing in Hungarian obstruents is neutralized in all and only the positions where the main cues are missing: before obstruents
- the difference between Lith. & Hung.? – where the language marks the point where contrast is allowed in the cue scale

Lithuanian vs. Hungarian





further voicing patterns predicted by Licensing by Cue

- V__R ▶ #__R
- Lac Simon (Quebec) and Totontepec Mixe (Mexico): neutralization word-finally, before/after obstruents but also word-initially; contrast: only between sonorants
- LS: loanword *banana* is [pa:na:n]
- TM: *na**s**oya* 'embroidered' – *wa**z**oy* 'shirt'; but only *s* initially, finally, before/after obstruents: ***s**u**s**pa* 'musician', *tadapu**s*** 'he already cut it', *mnahk**s**up* 'you're going'
- this pattern of neutralization is also impossible to characterize in syllabic terms: what neutralizes in LS or TM are all the codas plus a subset of the onsets!

further voicing patterns predicted by Licensing by Cue

- word-initial neutralization can be straightforwardly analyzed by observing the difference between V__R and #__R contexts on the perceptibility scale, with #__R inferior to #__R
- $\boxed{\text{V}\_\_\text{R} \blacktriangleright \#\_\_\text{R} \blacktriangleright \text{V}\_\_\# \blacktriangleright \text{V}\_\_\text{O} \blacktriangleright \{\text{O}\_\_\text{O}, \text{O}\_\_\#, \#\_\_\text{O}\}}$