Continuous and discrete aspects of phonological cognition

Štefan Beňuš

sbenus@ukf.sk

http://www1.cs.columbia.edu/~sbenus/

Constantine the Philosopher University in Nitra
Institute of Informatics, Slovak Academy of Sciences, Bratislava

AI Seminar, Comenius University, Bratislava, 24.3.2014

The big issue addressed today…

- A theory of cognition must provide tools for studying **the relation between the qualitative and the quantitative aspects of cognitive systems**
- In this talk
  - Qualitative = phonology
  - Quantitative = phonetics
  - Cognitive system = speech

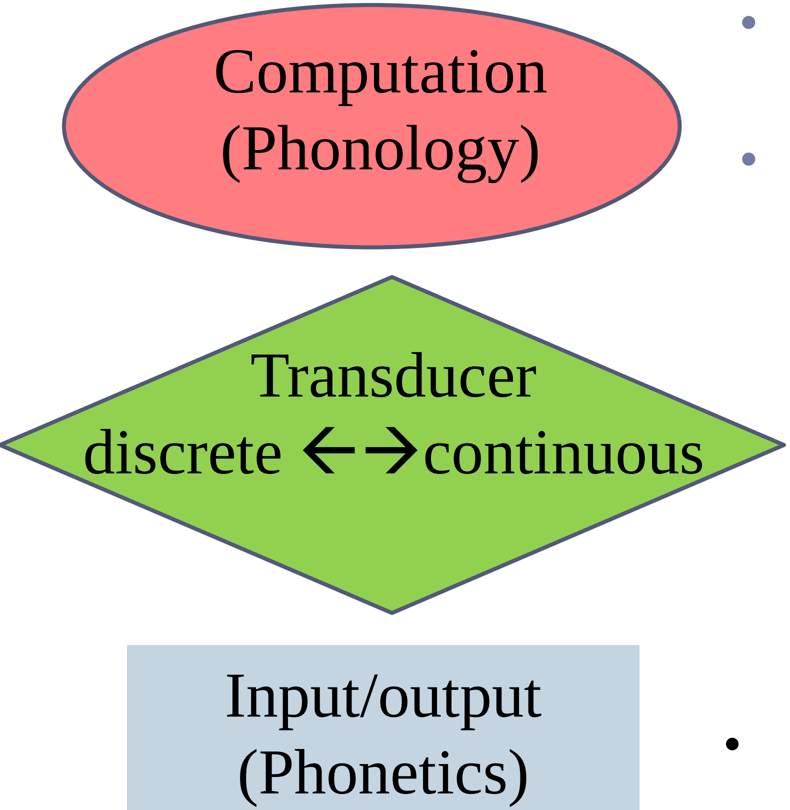
sbenus@ukf.sk

Speech

- Displays and should account for:
  - more stable qualitative discrete-like aspects
    - English: do[gz], ca[ts] but not *do[ks], *ca[dz]
    - German: do[ks], ca[ts], *do[gz], *ca[dz]
    - Slovak:
      - do[ks], ca[ts] /__ {#, C[-voice]}   (drozd, chvost)
      - do[gz], ca[dz] /__ {V, C[+voice]}   (chvost {mojej, alebo})
    - systematic sound patterns can be described as combinatorial systems of symbolic linguistic units such as features, segments, and higher prosodic constituents (a.k.a. Phonology)
  - more noisy continuous aspects
    - If you ever look as word-final /s/, /z/, or other obstruents, you see "mess"
    - But essentially, hypotheses based on phonology are typically supported statistically when investigating continuous vocal tract activity and acoustic events (a.k.a. Phonetics)

sbenus@ukf.sk

‘Traditional’ view



- Abstract symbolic representations
  - /t/, [+voice], σ,…
- discrete logical operations
  - /d/ → /t/ /__ $]_{\sigma}$

“the realization component ... maps symbolic categories –things that can be described using discrete mathematics – onto physical parameters – things that can be described using continuous mathematics” (Ladd 02)
Grounding (Harnad), Interpolation (Pierrehumbert, Keating)

- Continuous articulatory movements and their acoustic consequences in real space and time

“We believe that phonology consists of a set of formal properties (e.g., organization into syllables and feet, feature spreading processes) that are modality independent and thus not based on phonetic substance. The goal of phonological theory should be to discover these formal properties. Failure to appreciate this goal has resulted in rampant ‘substance abuse’ in the phonological community.“ (Halle & Reiss 01)

sbenus@ukf.sk

GOFAI

- Cognitive systems
  - mental representations with a crucial property: they are systematic, structured combinations of discrete constituents (Fodor & Pylyshyn, 1988; Pylyshyn, 1984)
  - Discrete computation as operations (transformations) over these representations
- Today's question(s)
  - How does this type of cognitive system (competence) relates to the measurable aspects of speech (performance), and to other variable and continuous aspects of speech?
  - What if native speakers produce "intermediate" renditions?
    - And they do! Systematically!
  - Is/should be this included in our modeling of cognition?
  - What is the formal system that best describes our knowledge about speech?

sbenus@ukf.sk

Traditional approach: Pros

- Notions of contrast, modality independence
- Attempts to combine the computation and substance in feature theory
- Analysis-by-synthesis
  - Successes in (rule) speech synthesis when phonetics ‘acts on’ the output of phonology
- …

sbenus@ukf.sk

Traditional approach: Cons

- Social
  - two formal approaches, two communities (phonetics sometimes further divided into the 'production' and 'perception' people)
  - "Phonological forms are not constrained to be producible in a vocal tract, descriptions of vocal tract activities need not be, and are not, descriptions of phonological forms, and neither phonological forms, nor vocal tract activities need to be perceivable." (Goldstein & Fowler)
- The question of time
  - A-temporal nature of the phonological representations prevents the explorations of patterns (=cognition) regarding timing of articulatory actions (*abi* vs. *iba*)
- Methodological
  - Kosslyn's (1967) "inference problem"
- Architectural
  - phonetics ←→ phonology effects

sbenus@ukf.sk

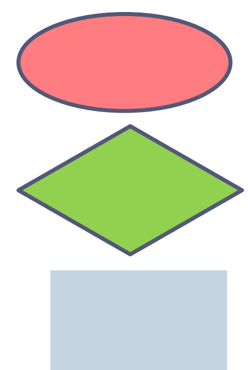
Kosslyn's "inference problem"

- Data = F (*Competence*) + Noise

core of linguistic inquiry

but nature of F is not explicitly studied

- Computation (Competence) is embedded in a continuously varying environment. To understand it, we must use inferences based on surface, performance data extracted from specific contexts.
- Abstracting away from contextual or environmental factors requires an understanding of how computation adapts to different contexts.
- This, in turn, assumes an understanding of computation.



sbenus@ukf.sk

Phonetics ← → phonology

- Inventory interdependence
  - Phonetic (contextual) variability constrained by the number of phonological categories
  - The number of categories constrained by phonetic considerations (e.g. b, d, *g)
- Incomplete neutralization
  - German: consonants like /d/ and /z/ are devoiced word-finally but the trace of the voicing distinction remains and is accessible for other pragmatic tasks (e.g. Port & Crawford 89)
    - *Rad* ‘wheel’ /rad/ → [rat] → ? [t+]
    - *Rat* ‘advice’ /rat/ → [rat] → [t]
  - Problem: the phonetic implementation transducer cannot deliver the difference observed in the final consonants after the phonology wipes the contrast out.

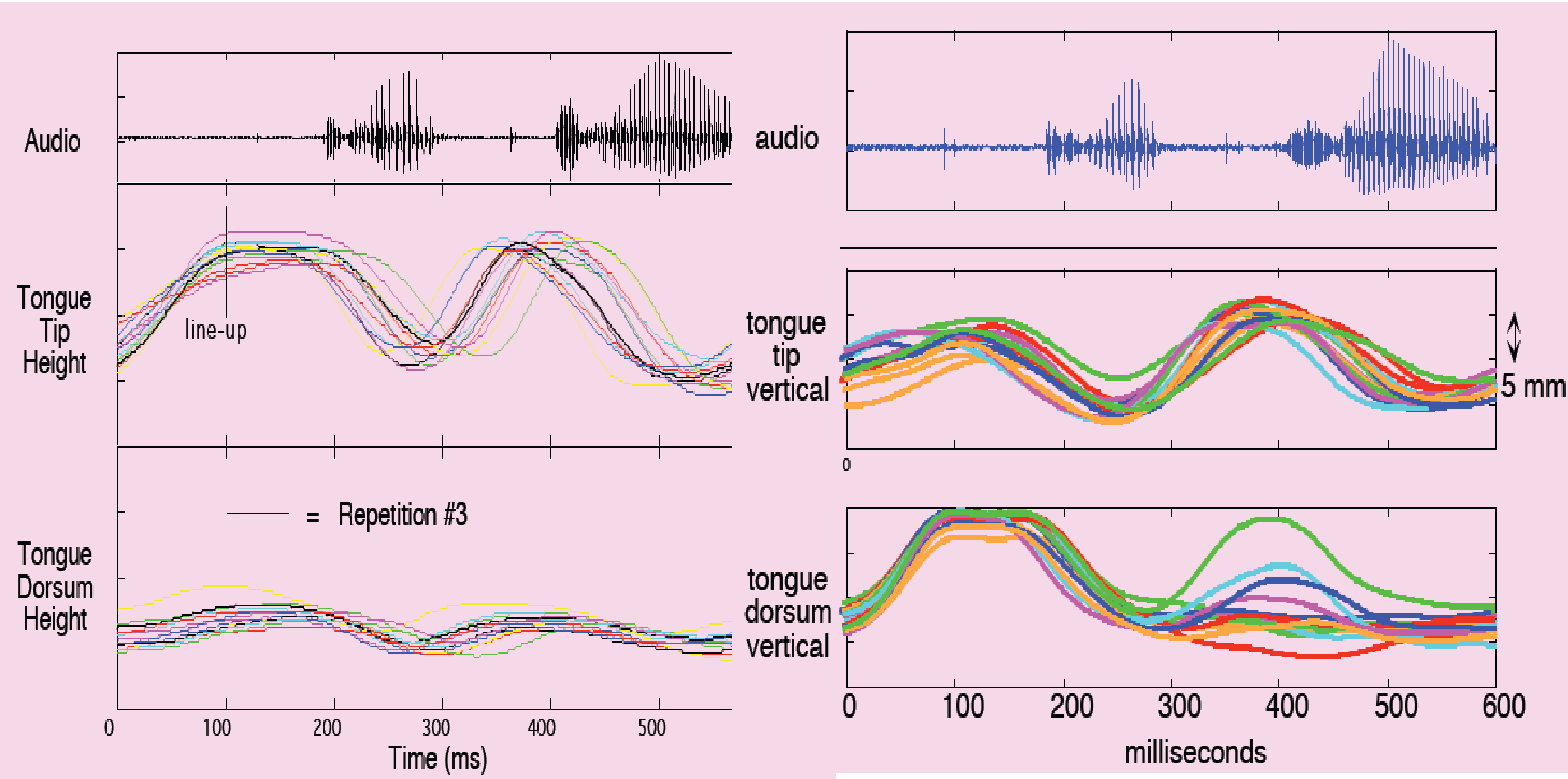
Phonetics ← → phonology

- Speech errors
  - Traditionally assumed to support the division between ‘form’ and ‘substance’ (here planning and implementation, Fromkin 71 and others)
    - Errors happen during planning and are implemented ‘normally’, e.g. [p^h]ick a s[p]oon → s[p]ick a [p^h]oon
  - But careful phonetic studies of error productions reveal systematic effects of lexical factors (Goldrick) and phonetic context (Pouplier)
    - Planning cannot be devoid of implementation

sbenus@ukf.sk

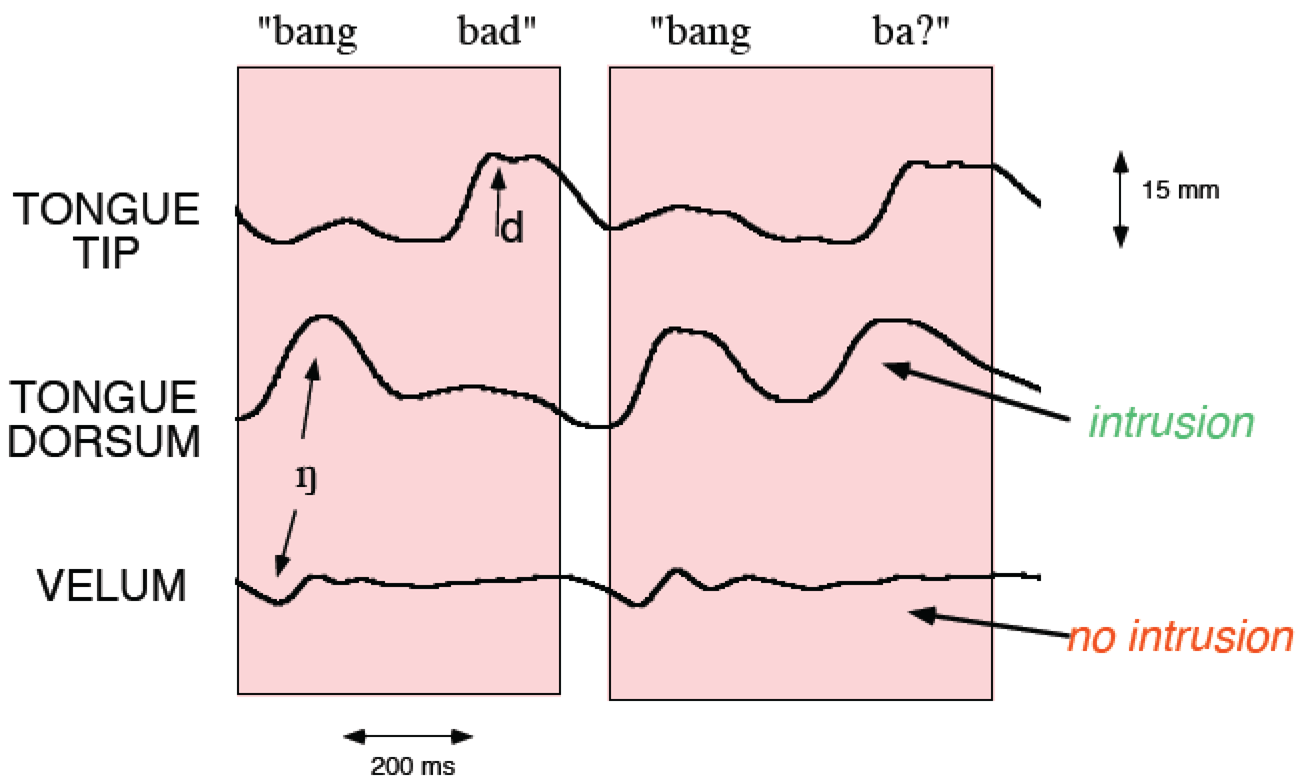
Intrusion errors (Pouplier & Goldstein)

top-top (control) cop-top



Source: L. Goldstein's notes on Articulatory Phonology

Errors may be sub-segmental



Source: L. Goldstein's notes on Articulatory Phonology

sbenus@ukf.sk

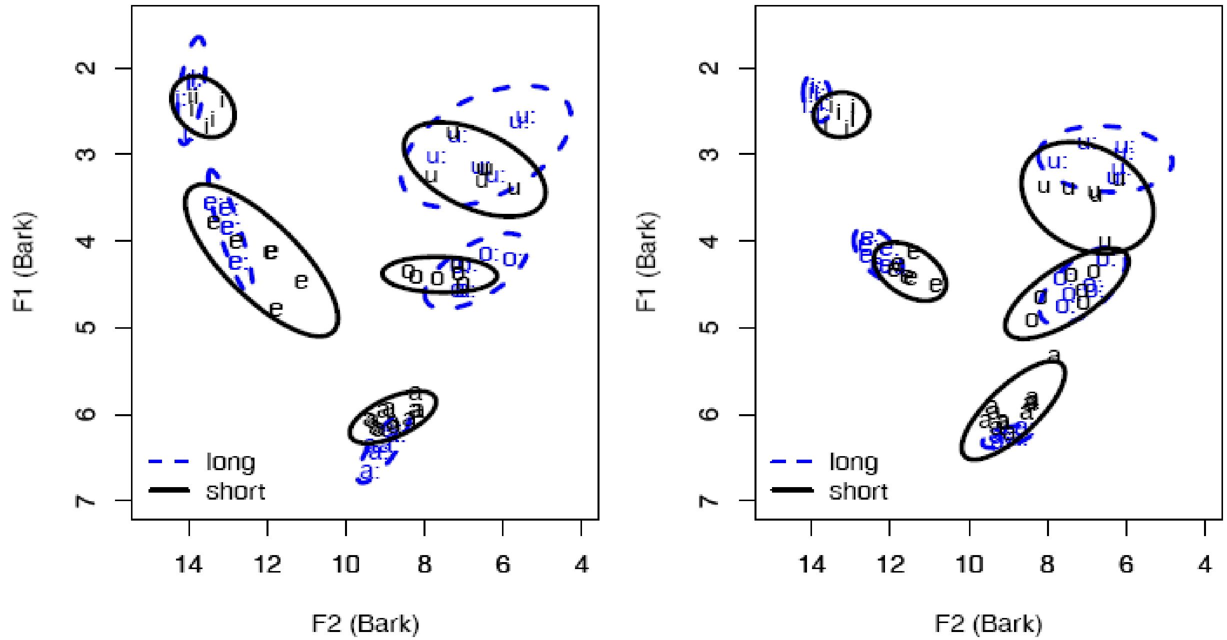
Intermediate summary

- In domains relatively “close to surface/performance” a strict boundary between phonetics and phonology is questionable
- These post-lexical processes are typically variable, depend on style/speed, do not have exceptions
- What about processes that are traditionally lexical, deep, do not depend on style/speed, have exceptions?

Case study #1
Slovak Yers

Slovak vowels & Yers

- Slovak is a typical 5-vowel system
- Yers are vowels that alternate with zero



Nom. Sg.	Gen.Sg.	Instr.Sg.	Gloss
pal**e**c	palc-a	palc-om	‘thumb’
lak**e**ť	lakť-a	lakť-om	‘elbow’
p**e**s	psa	psom	‘dog’
kot**o**l	kotl-a	kotl-om	‘cauldron’
pár**o**k	párk-a	párk-om	‘sausage’

Yer background

- Diachronically
  - Present in Common Slavic (and Old Church Slavonic)
  - Represent IE short lax /i/ and /u/ ([ĭ], [ŭ])
    - *vьdova Lat. vidua widow
    - *dъva Lat duo two

(examples from Scheer 03)

  - Became centralized and shortened, and eventually lost altogether in some positions (cca. 10th century); e.g. *vdova*, *dva* in Slovak
- Synchronically
  - Present in some positions in most modern languages
    - In some the front-back distinction maintained, e.g. Slovak
    - In others they merged , e.g. Czech (kotel, párek)
    - Can develop also in non-slavic borrowings; e.g. sveter-svetra

sbenus@ukf.sk

Phonology

- The alternations could be treated as insertions or deletions but the environment cannot be specified
  - Deletion: kotol-kotla-*kotola vs. kostol-*kostla-kostola (‘stl’ is an ok cluster in Slovak)
  - Syllabically-based insertion
    - *kotl, *metr (ok in Czech), but
      - Which vowel to insert in languages like Slovak?
      - Some word-final clusters exist, e.g. *park* (c.f. *párok*)
- Hence, yer and nonyer vowels must be different lexically (underlyingly)
  - Various ways of achieving this (yers in UR inventory, yers as unassociated floating segments, etc.)
  - Followed by rules/constraints for their vocalizations and subsequent lowering to /e/ or /o/
    - Vocalize a yer if another yer follows and lower to a mid vowel (kot**x**l-**x** → kot**ŭ**l**x** → kot**o**l)

Phonologically, yers are "deficient" compared to non-yers

- Phonological weakness of yers
  - Underlyingly specified with [tense], not supported on the surface
  - Unassociated to the melodic tier (e.g. Rubach 93)
  - More marked (e.g. Jarosz 06)
    - simultaneously [+high] and [-tense] or [-high] and [+tense]
  - In government models of phonology formalized as dependent (i.e. incapable of government, Scheer 06)
- Yers are also less frequent as types
- Prediction of all phonological accounts: this phonological deficiency does not carry over to phonetics

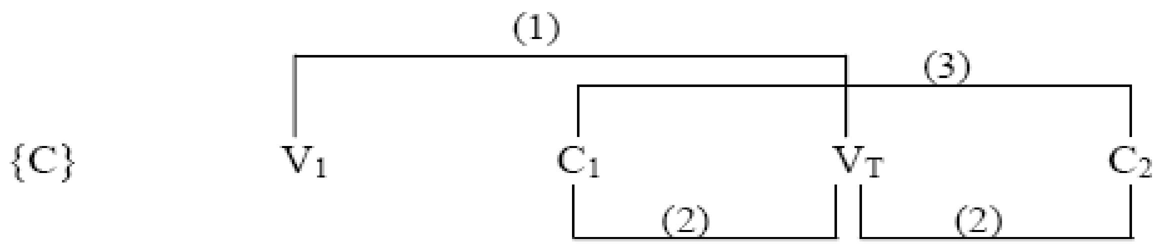
sbenus@ukf.sk

Phonetic weakness of yers

Beňuš (2012, JPhon)

- Prosodic weakness
  - Shorter duration, more centralization (undershoot)
  - Present in Slovak
- Prediction: if weaker
  - yer V_T should be shorter, more centralized, and less resistant to coarticulation from surrounding sounds than non-yer V_T
  - yers should behave similarly to non-yers in fast rate

  - Measure the similarity of V_T with V_1 (1)
  - Test the degree of coproduction between V_T and lingual C_1 or C_2 (2)
  - Test the degree of overlap between C_1 and C_2 (3)



sbenus@ukf.sk

Predictions

- No phonetic differences whatsoever
  - No perceptual difference for /e/ and /o/ pairs, one category since about the 10^{th} century
  - Same orthography, short unstressed mid vowels
  - Evidence for 'purely' phonological alternations
- If the differences are present
  - another phonetics-in-phonology effect in deep morpho-phonological alternations
  - they might signal the incomplete merger of the diachronic contrast

sbenus@ukf.sk

Experiment: Material

YER			NON-YER		
kábel (5564)	[ka:bel]	‘cable’	Ábel (381)	[a:bel]	‘Name’
Čapek (940)	[tʃapek]	‘Name’	papek (181)	[papek]	‘twig’
cumel (5)	[tsumel]	‘pacifier’	čumel (143)	[tʃumel]	‘he stared’
obec (128656)	[obets]	‘village’	obed (20240)	[obet]	‘lunch’
rámec (116986)	[ra:mets]	‘frame’	námet (8892)	[na:met]	‘idea’
párok (1523)	[pa:rok]	‘sausage’	nárok (34915)	[na:rok]	‘requirement’
nebol (>100000)	[ɲebol]	‘he wasn’t’	jebol (14)	[jebol]	‘he fell (curse)’
kufor (7796)	[kufor]	‘suitcase’	humor (13630)	[humor]	‘humor’
kapor (2510)	[kapor]	‘carp’	mramor (2323)	[mramor]	‘marble’
smútok (15699)	[smu:tok]	‘sadness’	sútok (600)	[su:tok]	‘confluence’

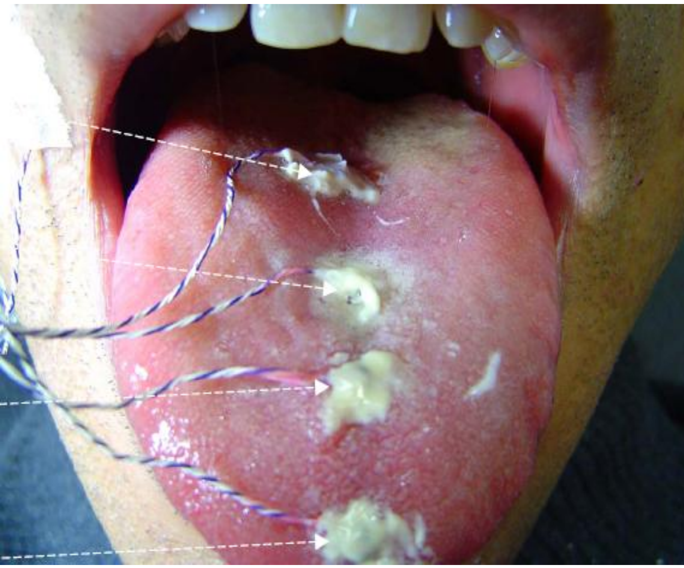
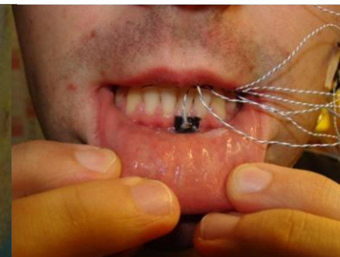
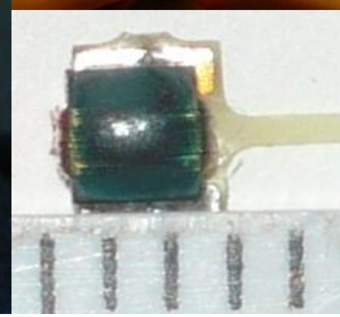
sbenus@ukf.sk

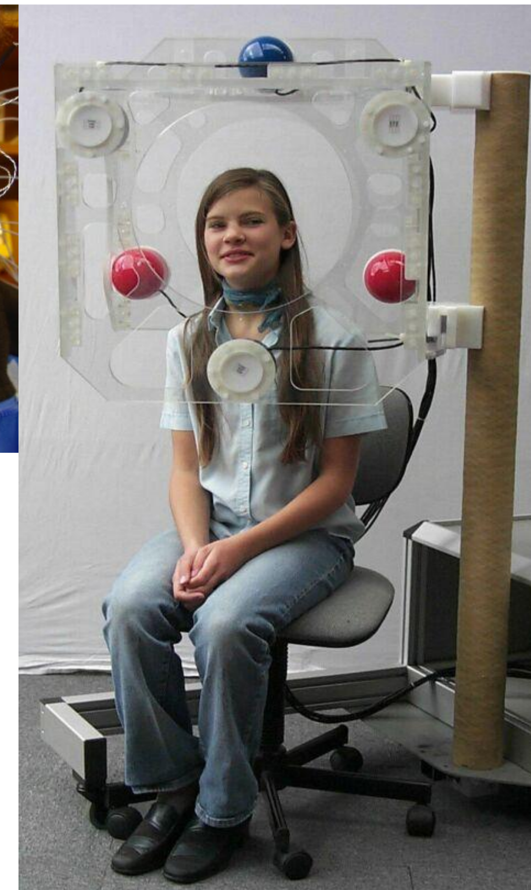
Experiment: procedure

- 5 subjects (2F, 3M) read sentences in alternating blocks of normal and fast speech rate
- Frame sentences had the target word conjugated in the 1st part, and the yer/nonyer word w/o a suffix in the 2nd part. The first part indirectly cues whether the word has a yer or not.
  - *Čítame s mramorom a mramor parádne.* NY
  - *Čítame s kaprom a kapor parádne.* Y
- 200 tokens per subject (20 target words, 5 reps in normal and 5 in fast rate)

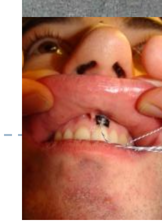
Electromagnetometry (EMA)

(IPS LMU Mníchov)

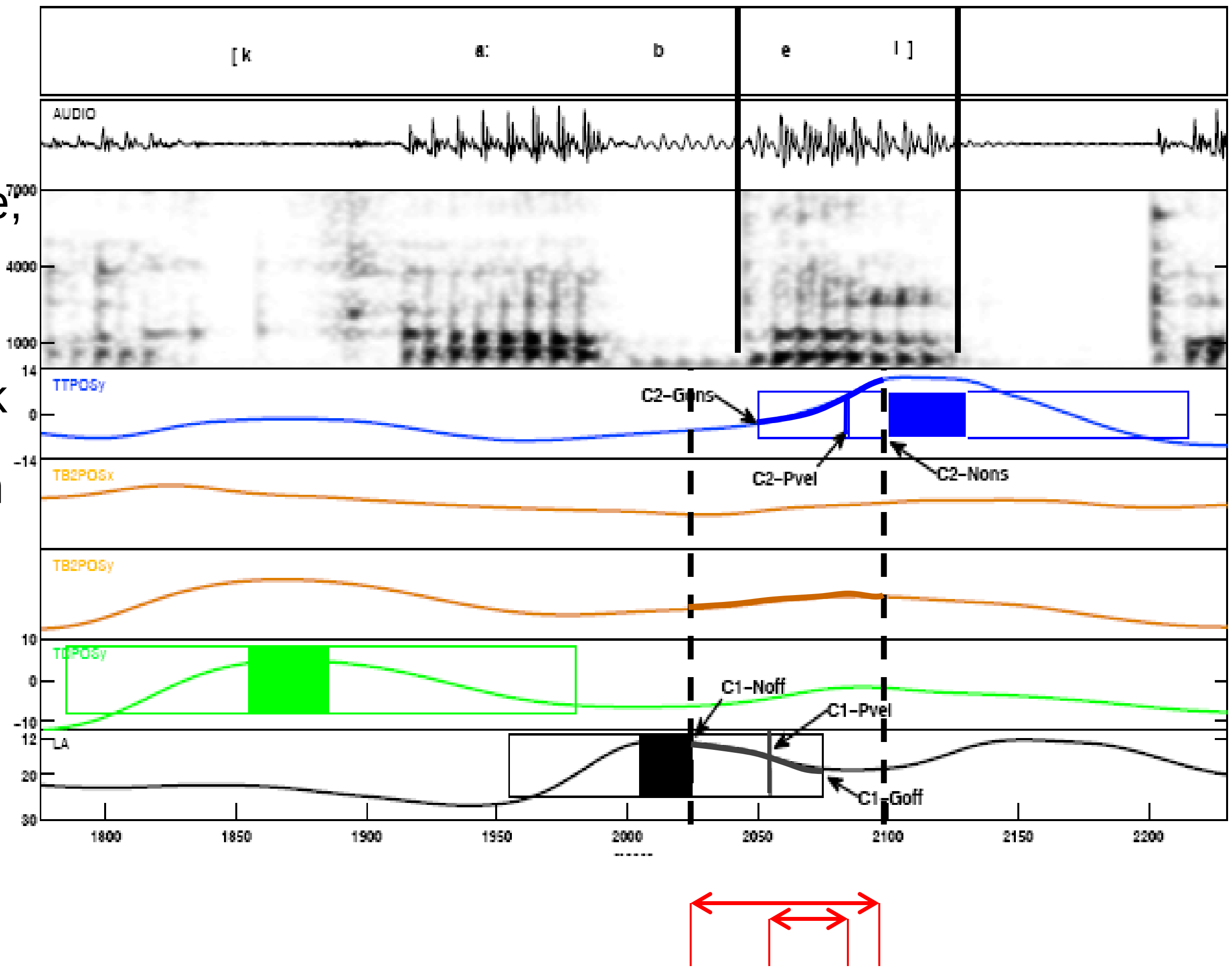


- Small receiver coils are attached on the active articulators (tongue, lips, jaw)
- due to the electro-magnetic field generated by the transmitter coils, we can record the movement of these small coils with high precision (up to 500 Hz).
- Reference sensors help with correction of non-articulatory movements



Measures and dep. variables

- Duration
- Quality
- Coarticulatory characteristics
  - V_1-V_T Euclidean distance; smaller distance => less coarticulatory resistance of V_T => weaker V_T
  - C-C coart.: Peak-to-peak ratio; smaller ratio => more truncation between the consonants flanking V_T (Harrington et al. 95, Hoole & Mooshammer 02) => weaker V_T
  - DurNuc also measures the overlap of the consonants; greater overlap => weaker V_T
  - C-V coart. : slope & curvature from DCT



sbenus@ukf.sk

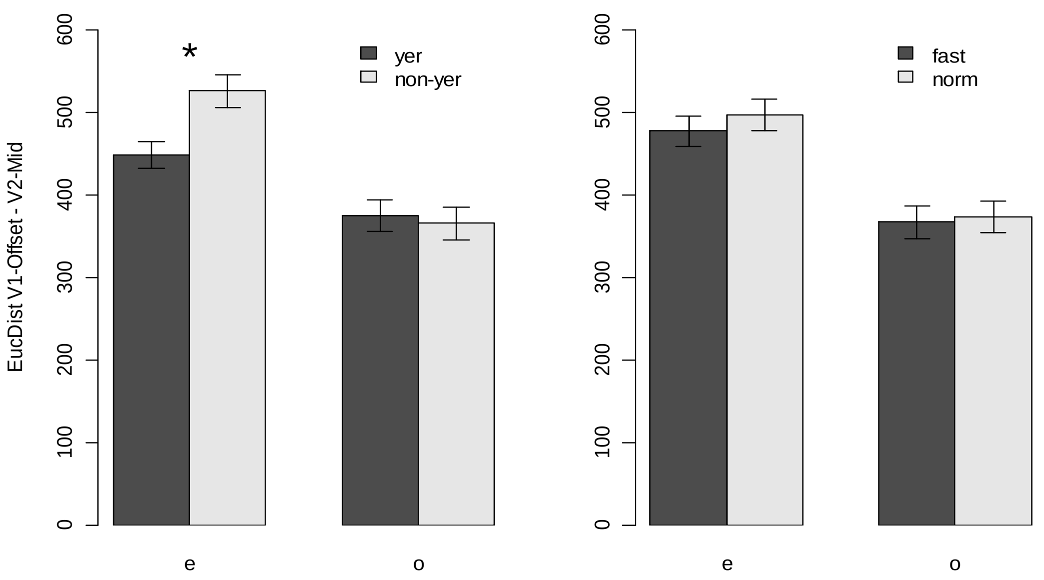
Results: speech rate

<table>
<tr><th>V_F = fast, V_N = normal</th><th>Hypothesis</th><th colspan="2">Measure</th><th>/e/</th><th>/o/</th></tr>
<tr><td rowspan="2">V-duration</td><td rowspan="2">V_F shorter than V_N</td><td colspan="2">DurAc</td><td colspan="2">☑</td></tr>
<tr><td colspan="2">DurArt</td><td colspan="2">☑</td></tr>
<tr><td rowspan="2">V-quality</td><td rowspan="2">V_F more centralized than V_N</td><td colspan="2">Ac (F2)</td><td>☑</td><td>☑</td></tr>
<tr><td colspan="2">Art (TB{1,2}-x)</td><td colspan="2"></td></tr>
<tr><td>V_1-V_T coarticulation</td><td>$|V_1-V_F| < |V_1-V_N|$</td><td colspan="2">V_1-V_TEucDist</td><td colspan="2"></td></tr>
<tr><td rowspan="4">(C)V_TC coproduction</td><td rowspan="2">slope V_F < slope V_N</td><td rowspan="2">DCT2 {TB1,TB2}</td><td>hor.</td><td colspan="2">☑</td></tr>
<tr><td>vert.</td><td colspan="2">☑</td></tr>
<tr><td rowspan="2">curvature V_F < curvature V_N</td><td rowspan="2">DCT3 {TB1,TB2}</td><td>hor.</td><td colspan="2">☑</td></tr>
<tr><td>vert.</td><td colspan="2">☑</td></tr>
<tr><td rowspan="2">C_1C_2 coproduction</td><td rowspan="2">$|\mathbf{C_1}V_F\mathbf{C_2}| < |\mathbf{C_1}V_N\mathbf{C_2}|$</td><td colspan="2">Peak-to-Peak Ratio</td><td colspan="2">☑</td></tr>
<tr><td colspan="2">DurNuc</td><td colspan="2">☑</td></tr>
</table>

sbenus@ukf.sk

Example Results in more detail: V_1-V_T coarticulation

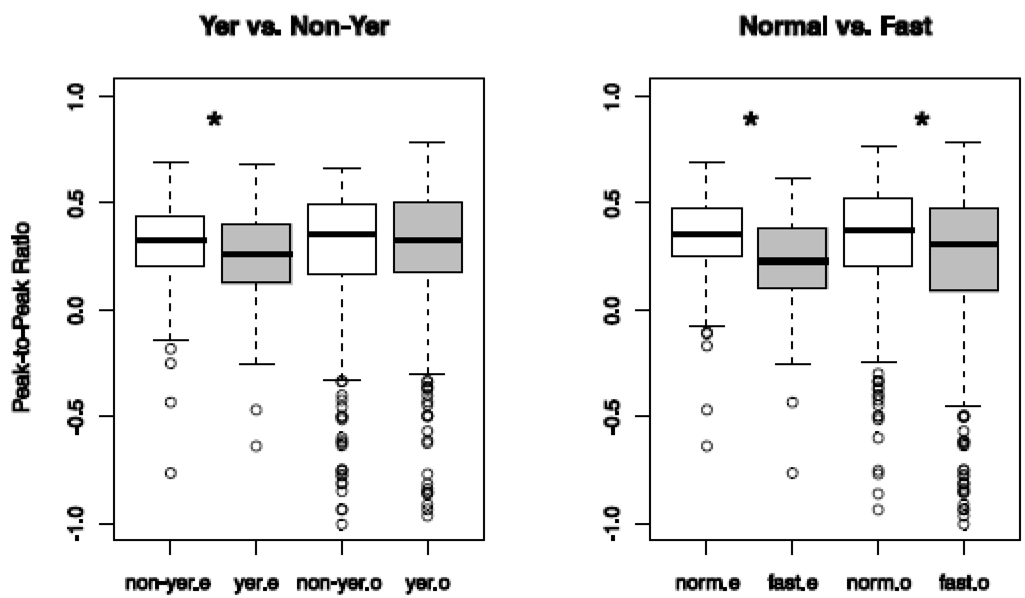
- yer /e/ significantly more coarticulated with V_1 than non-yer /e/ ($F = 25.8$)
- Speech rate similar but non-significant effect
- yer /e/ less resistant to coarticulation than non-yer /e/



sbenus@ukf.sk

Results: coarticulation between flanking Cs

- C-opening before and C-closing after V_T were more coproduced for yer /e/ than non-yer /e/
- Similar weakening observed for speech rate



- Cs flanking yers more co-produced on DurNuc measure than Cs flanking non-yers, similar weakening in speech rate

sbenus@ukf.sk

Results: yer vs. non-yer

<table>
<tr><th>V_Y = yer, V_{NY} = non-yer</th><th>Hypothesis</th><th colspan="2">Measure</th><th>/e/</th><th>/o/</th></tr>
<tr><td rowspan="2">V-duration</td><td rowspan="2">V_Y shorter than V_{NY}</td><td colspan="2">DurAc</td><td colspan="2">☑??</td></tr>
<tr><td colspan="2">DurArt</td><td></td><td>☒</td></tr>
<tr><td rowspan="2">V-quality</td><td rowspan="2">V_Y more centralized than V_{NY}</td><td colspan="2">Ac (F2)</td><td>☑</td><td></td></tr>
<tr><td colspan="2">Art (TB{1,2}-x)</td><td></td><td>☑</td></tr>
<tr><td>V_1-V_T coarticulation</td><td>$|V_1 - V_Y| < |V_1 - V_{NY}|$</td><td colspan="2">$V_1$-$V_T$EucDist</td><td>☑</td><td></td></tr>
<tr><td rowspan="4">(C)V_TC coproduction</td><td rowspan="2">$slope_Y < slope_{NY}$</td><td rowspan="2">DCT2 {TB1,TB2}</td><td>hor.</td><td>☑☒</td><td>☑</td></tr>
<tr><td>vert.</td><td colspan="2">☑</td></tr>
<tr><td rowspan="2">$curv._Y < curv._{NY}$</td><td rowspan="2">DCT3 {TB1,TB2}</td><td>hor.</td><td colspan="2">☑</td></tr>
<tr><td>vert.</td><td></td><td></td></tr>
<tr><td rowspan="2">C_1C_2 coproduction</td><td rowspan="2">$|\mathbf{C}_1 V_Y \mathbf{C}_2| < |\mathbf{C}_1 V_{NY} \mathbf{C}_2|$</td><td colspan="2">Peak-to-Peak Ratio</td><td>☑</td><td></td></tr>
<tr><td colspan="2">DurNuc</td><td colspan="2">☑</td></tr>
</table>

sbenus@ukf.sk

Summary

- no single result provides conclusive evidence and some inconsistencies were found
- BUT: the results converge in supporting the hypothesis that yers are phonetically weaker than non-yers
- Problem for traditional accounts
  - Differences should have been wiped out by “phonology”

sbenus@ukf.sk

Case study #2
Transparency in Hungarian

Hungarian vowel inventory

	Front		Back	
	[–Round]	[+Round]	[–Round]	[+Round]
High	i[i] í[i:]	ü[y] ű[y:]		u[u] ú[u:]
Mid	é[e:]	ö[ø] ő[ø:]		o[o] ó[o:]
Low	e[ɛ]		á[ɑ:]	a[ɔ]

- (e.g. Ringen & Vago 98)

(subset of) Hungarian vowel harmony

	Dative	Adessive	Notes
a. ház 'house'	ház-nak	ház-nál	regular harmony
b. tök 'pumpkin'	tök-nek	tök-nél	regular harmony
c. radír 'eraser'	radír-nak	radír-nál	/í/ is transparent
d. víz 'water'	víz-nek	víz-nél	TVs usually trigger front harmony
e. híd 'bridge'	híd-nak	híd-nál	TVs exceptionally trigger back harmony
f. nüansz 'nuance'	nüansz-nak	nüansz-nál	back vowels are opaque
g. parfüm 'perfume'	parfüm-nek	parfüm-nél	front round vowels are opaque
h. aszpirin 'aspirin'	aszpirin-nak/nek		adding TVs decreases transparency
i. hotel 'hotel'	hotel-nak/nek		/e/ is less transp. than /i/ but more than /ü/

sbenus@ukf.sk

Stem-final front vowels

I. A + {i, í, é} ↓ back suffix	papír-ban/*ben buli-ban/*ben kávé-ban/*ben	‘paper.Iness’ ‘party.Iness’ ‘coffee.Iness’
II. A + ü ↓ front suffix	parfüm-*ban/ben	‘perfume.Iness’
III. A + e ↓ vacillation	hotel-ban/ben Ágnes-ban/ben	‘hotel.Iness’ ‘Agnes.Iness’
IV. {i, í, é} ↓ front/back suffix	híd-ban/*ben víz-*ban/ben	‘bridge.Iness’ ‘water.Iness’
V. A+{i, í, é}+{i, í, é} ↓ front suffix/vacillation	aszpirin-ban/**ben**	‘aspirin.Iness’

sbenus@ukf.sk

Challenges in the data

- The <u>set of transparent vowels</u> is {i, í, é, e}: common properties resulting in transparency, and differences within the set
- The notion of <u>locality</u>: front vowels in the back harmony domain, e.g. ‘radír-nak’
- <u>Exceptions</u>(?): transparent vowels may also select a back suffix, e.g. ‘víz-nek’ vs. ‘híd-nak’
- The nature of <u>vacillation</u>, e.g. ‘hotel-ban/ben’

sbenus@ukf.sk

Motivation for the articulatory study

- Well-accepted assumption in phonology:
  - Transparent vowels do not participate in vowel harmony, at least not on the surface.
  - This seems right, [i] in Tomi seems identical to [i] in Imi, neither phonologists/phonetitians nor naïve native speakers perceive them as different.
  - But we knew little about the articulatory characteristics of these vowels

- Prediction (of the traditional view):
  - the phonetic properties of these vowels in different harmonic contexts should be the same and any potential difference should be only due to coarticulation

sbenus@ukf.sk

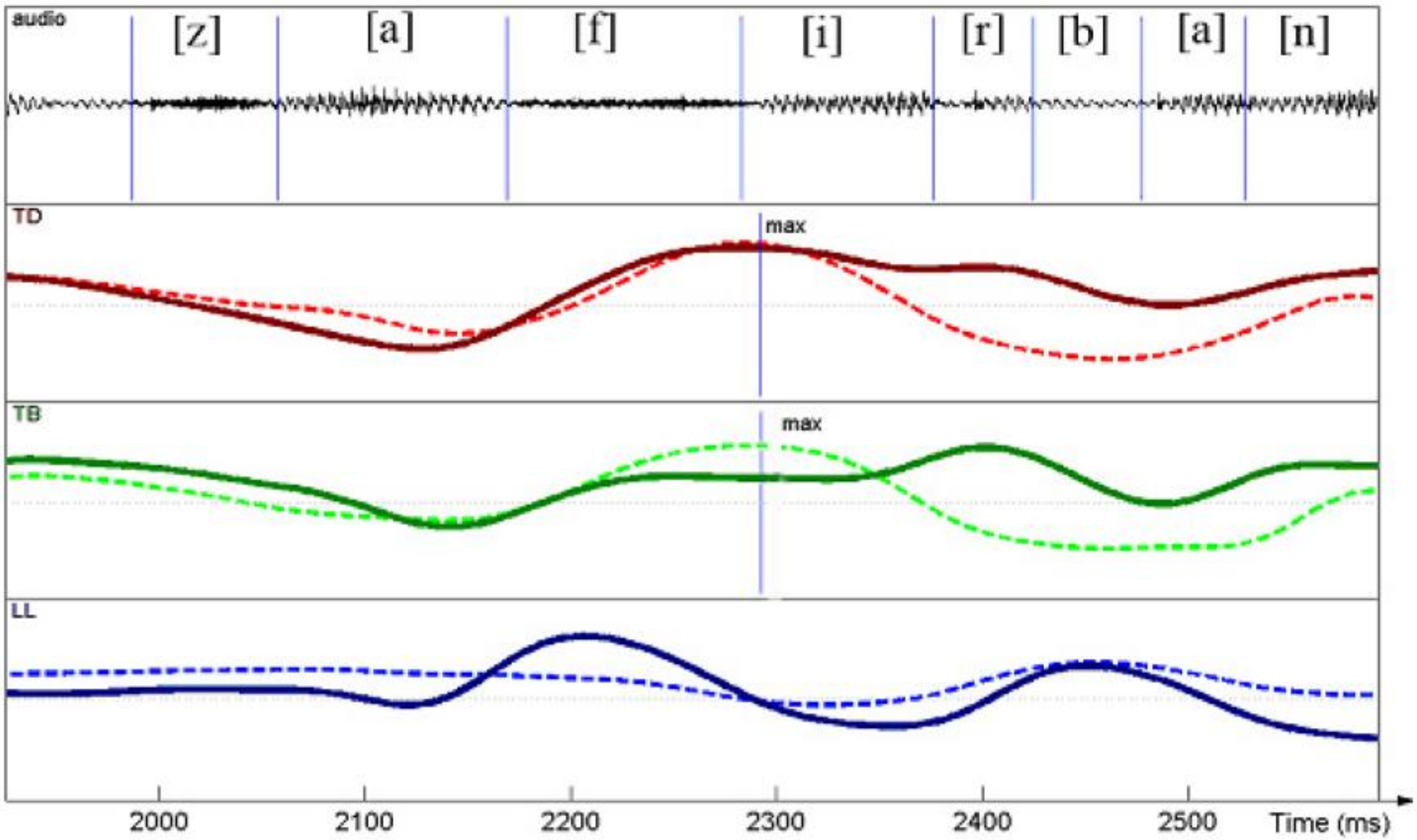
Experiment: TVs in [$\pm$back] context

Stimuli	Back context		Front context	
trisyll.	[kaːbiːtom]	‘daze’	[repiːtɛm]	‘let fly’
	[bulivɔl]	‘party’	[bilivɛl]	‘pot’
	[boːdeːtoːl]	‘hut’	[bideːtøːl]	‘bidet’
monosyll.	[ʃiːp]	‘whistle’	[tsiːm]	‘address’
	[tseːl]	‘aim’	[seːl]	‘wind’

- **Methodology**: EMMA (3 subjects), Ultrasound (1 subject)
- **Measured:** maximal advancement of the tongue, quantified with 2 dependent variables (plus one based on pair-wise area comparisons)

sbenus@ukf.sk

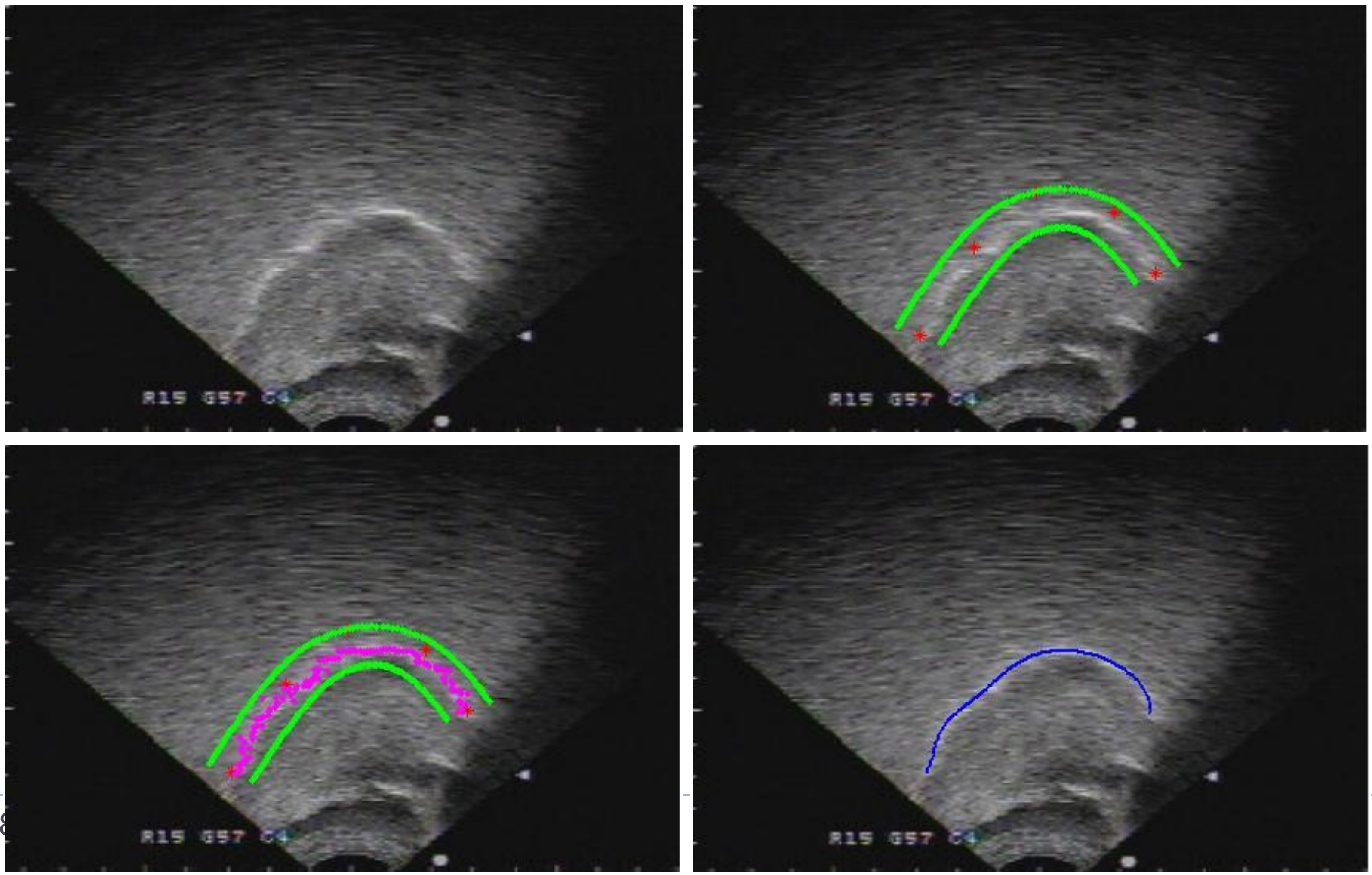
EMMA: horiz. position of lingual receivers



- Determine the value at the peaks of the time functions representing the kinematic trajectories of the receivers attached sagittally on the tongue (Tiede et al. 1999)

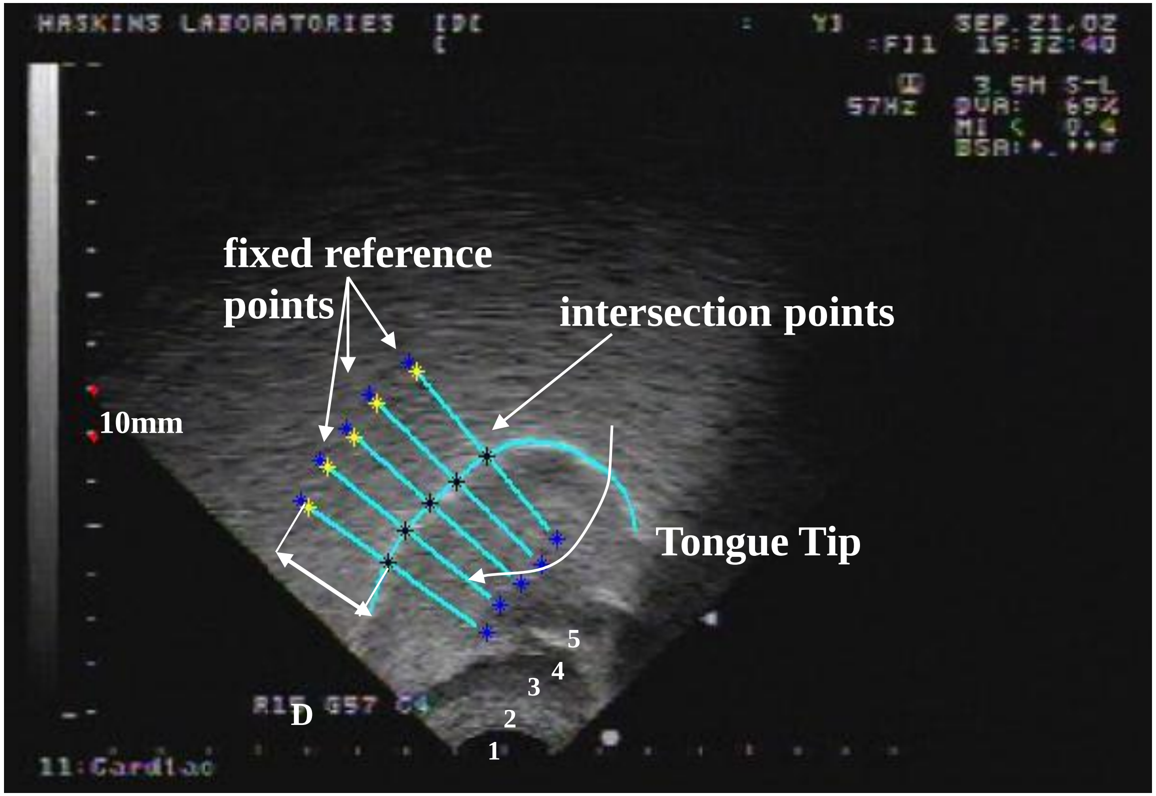
Ultrasound: edge-tracing

- Determine the frame with the most extreme front position
- Edge tracing: B-spline snakes (Iskarous 2005)

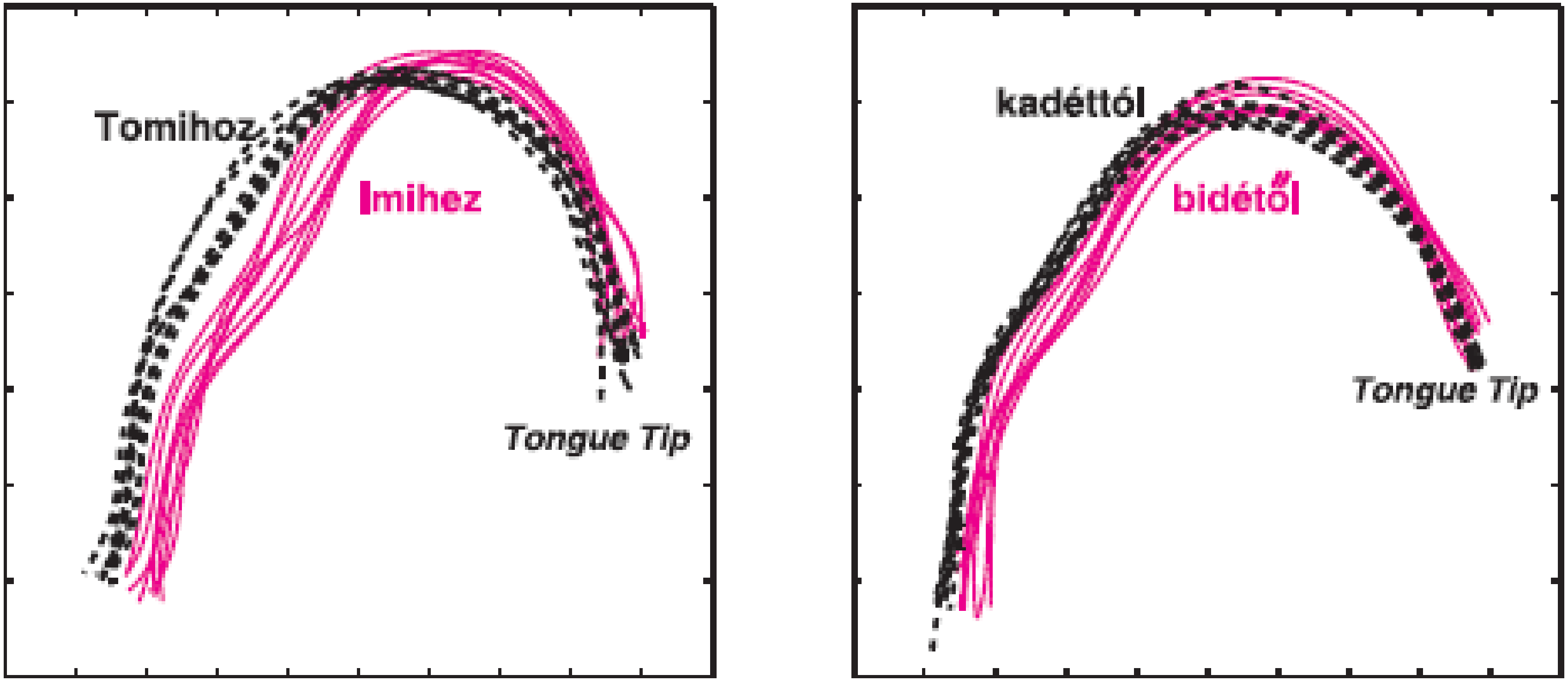


Degree of dorso-pharyngeal constriction

- Determine the distance between the fixed point on the line and the point where the line intersects the tongue surface



Results-preview



Main results (Benus & Gafos 2007)

- TVs in front harmony contexts were **<u>slightly</u>** less retracted than in back harmony contexts.
- This effect was robust and highly significant for all 3 subjects and both methodologies with trisyllabic words

- With monosyllabic words, the effect was less robust but still significant for some measurements
  - the result cannot be due to coarticulation from adjacent vowels, this difference has to be stored
  - Hence, in effect, TVs have a "back" counterpart, which combines with back vowels for the purposes of harmony and a "front" counterpart that combines with the front vowels

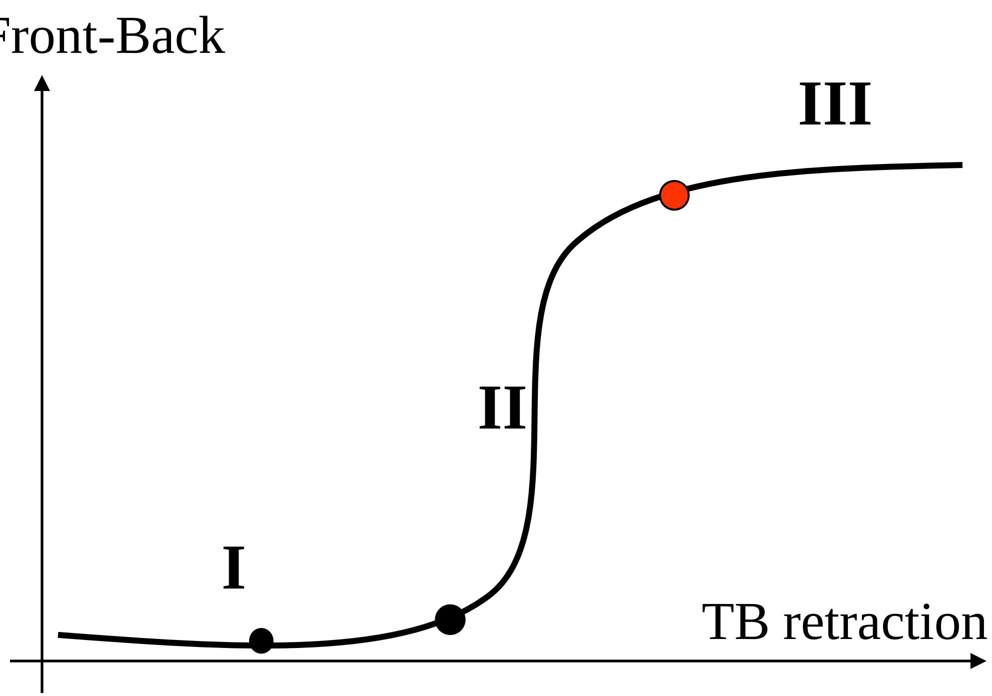
sbenus@ukf.sk

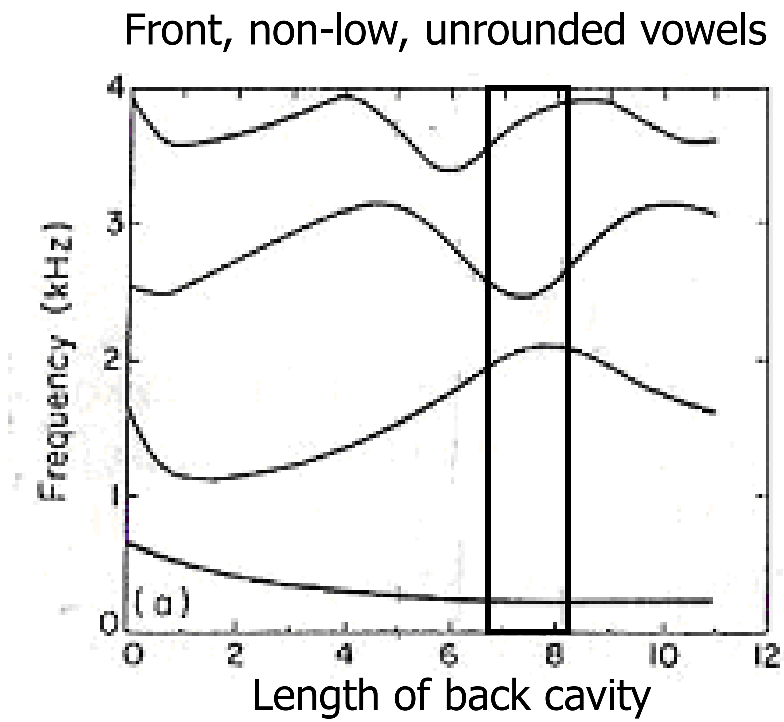
Phonetic basis of transparency

- Our experiments: transparent vowels are <u>prone</u> to <u>articulatory</u> coarticulation
- Proposal: phonological transparency correlates with the quantal nature of vowels
- High front vowels like /i/ are <u>resistant</u> to <u>perceptual</u> coarticulation (many studies)
- Degree of phonetic retraction of a V is linked to its phonological behavior

sbenus@ukf.sk

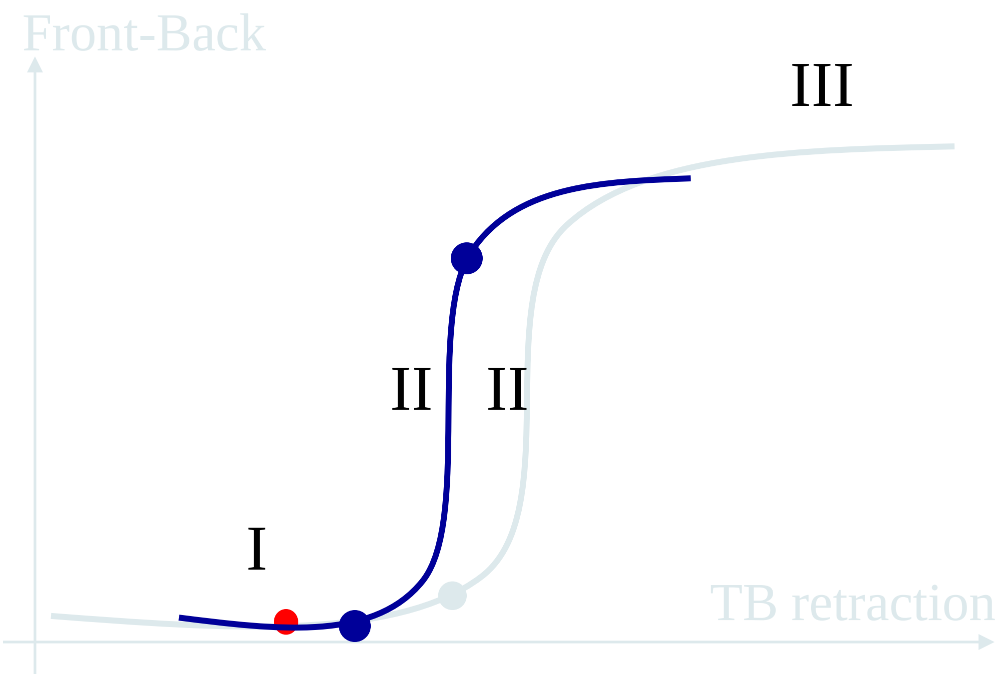
Phonetic Basis of Transparency

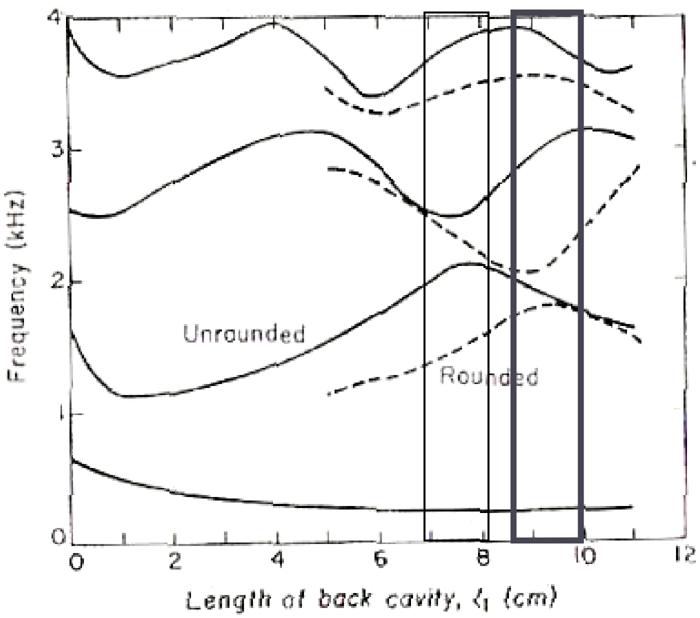




- [i] can be retracted significantly without corresponding acoustic consequences (Stevens 1972, Wood 1979).
- [i] is most likely to be followed by a [+back] suffix.

Opacity

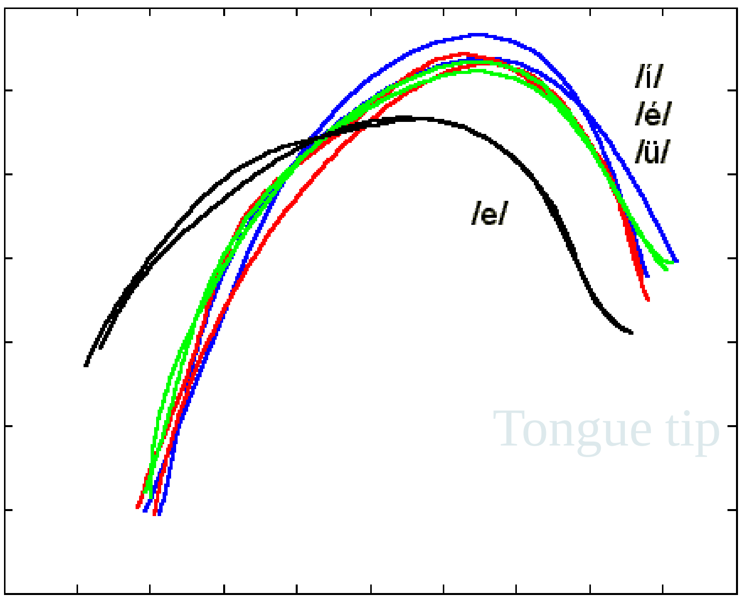




- /ü/ cannot be retracted to the same degree as /i/ without losing its perceptual identity (Wood 1986).
- /ü/ is most likely to be followed by [–back] suffixes in B-ü stems.

Low /e/: medial retraction

- The acoustic output of the front unrounded low vowels is more sensitive to articulatory perturbations in the horizontal position of the tongue body than the acoustic output of the non-low vowel.



Support:

Välimaa-Blum (1999), and indirectly in Stevens (1989) and Wood (1986)

- /e/ can be medially retracted
- /e/ is followed by either [+back] or [–back] suffixes in B-e stems.

Summary of Hungarian TVs

- Degree of phonetic articulatory retraction in stem-final vowel, constrained perceptually, correlates with suffix choice
  - Greater phonetic retraction => greater chance of a back suffix
- Why problematic for traditional accounts?
  - Due to uni-directionality, differences should be wiped out by phonology
  - The systematic and phonetically meaningful relationship between stem-final retraction and suffix form cannot be used to model the cognitive system of speech (competence)

sbenus@ukf.sk

Summary of Hungarian TVs and Slovak YVs

- Traditional Phonetics > || > Phonology not supported even in 'deep' morpho-phonological patterns
- If relaxed, cognitive "phonological" systems enriched with "phonetics" might provide better explanation of the patterns (VH model)
- Super minute differences: why would the system keep them?
  - Assuming Phonology ←→ Phonetics, phonetic differences *enhance* complex abstract phonological patterns (mutually!)

sbenus@ukf.sk

An alternative (*mode of inquiry*): Nonlinear Dynamics

- A formal language and a research paradigm that allows to
  - Express *both* qualitative and quantitative aspects of a complex system within a unified framework
  - Do away with the temporal metaphor of precedence between the qualitative and the quantitative, ***without losing sight of the essential distinction between the two***.
- Moreover:
  - "… dynamics […] happens to be the single most widely used, most powerful, most successful, most thoroughly developed and understood descriptive framework in all of natural science. It is used to explain and predict phenomena as diverse as subatomic motions and solar systems, neurons and 747s, fluid flow and ecosystems. Why not use it to describe cognitive processes as well?" (van Gelder & Port 1995: 4).

Dynamics: basic notions

- Dynamics models the motion (change) of systems in time
- This is precisely what the articulators do
- Speech can be conceptualized as achievements of target vocal tract constrictions, similar to reaching motions.
  - Noise: effect of environment, conditions
  - Attractors: “phonological”, discrete states
    - Representations
    - Processes (rules (SPE)/constraints (OT))

sbenus@ukf.sk

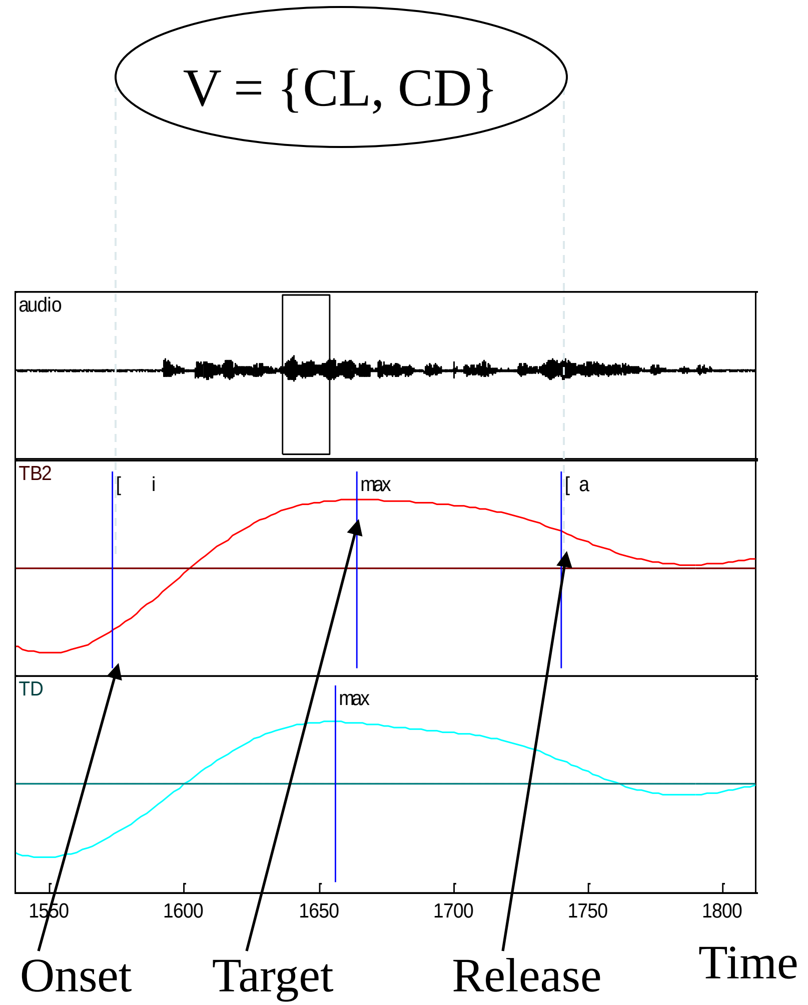
Non-linear dynamics model for Hungarian transparency

Gestures as phonological representations (Gafos & Goldstein 2011)

- Specify target constriction **abstractly** using constriction location (CL) and constriction degree (CD) variables
  - E.g. /p/ ≈ make lip aperture (LACD) = 0
- The articulator motions are **context-dependent** (e.g. bite-blocks), but the task description guiding them is **invariant** (motions ≠ gestures)
- While the state (e.g. LACD) is changing continuously, the (differential) equations that give rise to the time-varying state are fixed and represent the same level of abstraction as symbols in the traditional approach
- Combinatory power stems from differences in primary articulators and from discretizing CD and CL continua (possibly through quantal (=non-linear) relationship between articulation and perception)

Gestures as phonological representations

- Phonological representations are dynamically defined spatio-temporal gestures (Browman & Goldstein 1995).
- Each vowel is represented as a gesture with a specified constriction location (CL) and constriction degree (CD) variables (Wood 1986).



sbenus@ukf.sk

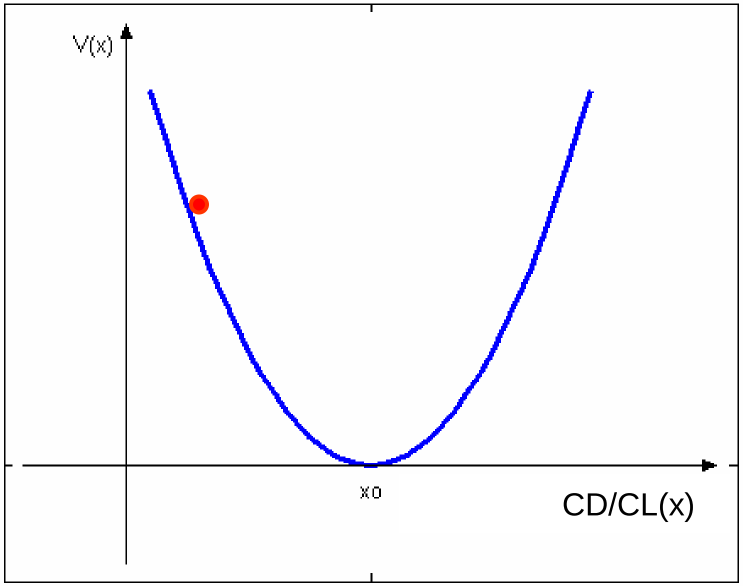
Conceptually…

- Most speech sounds can be modeled dynamically as point-attractors with a simple mass-spring system

$$m*d^2x/dt^2 + b*dx/dt + k(x - x_o) = 0 \qquad dx/dt = -k/b(x - x_o)$$

Geometrically…

- the movement of an articulator toward a target can be imagined as a ball moving in a potential landscape V(x)
  - dx/dt= -dV(x)/dt
  - For $dx/dt = -k/b(x-x_0)$, $V(x) = k/2b(x-x_0)^2$



Stability as resistance to noise

- In natural systems, attractive states exhibit small fluctuations around their mean values.

- Fluctuations are due to noise. Noise is present because behavior is complex and includes parallel involvement of different faculties and necessary coupling between them.

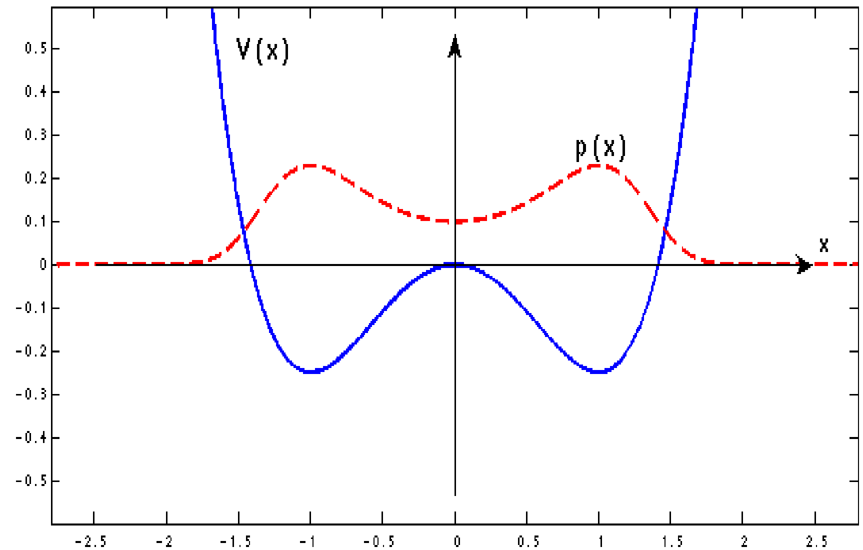
sbenus@ukf.sk

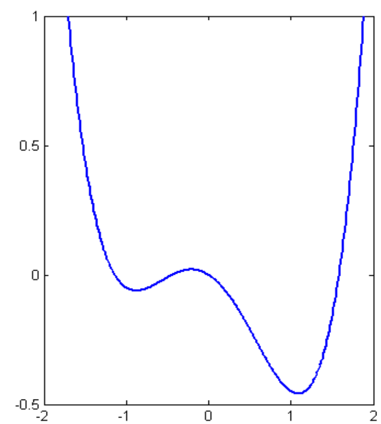
Stochastic dynamical systems

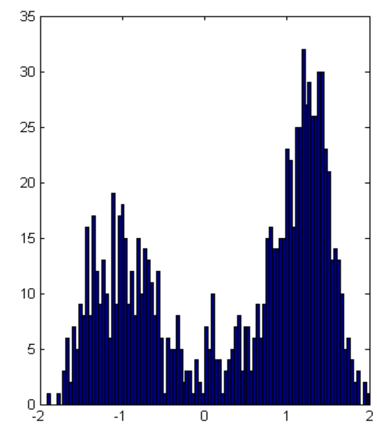
$$\dot{x} = f(x) + Noise = -dV(x)/dx + Q\sqrt{\xi_t}$$

- We can compute the probability of finding *x* within a given region of values using the probability density function *p(x)* (Freidlin & Wentzell 84)

- We can also estimate pdf by numerically simulating the asymptotic behavior of *x* by solving the DE for random initial conditions and with added noise and then plot the histogram of these solutions.







sbenus@ukf.sk

Stability coexists with change

- Attractors are stable in that they are resistant to noise in a probabilistic sense.
- But in behavioral systems this stability coexists with the flexibility to change.
- At a formal level, the ability to change requires that we relax the notion of dynamic stability.

sbenus@ukf.sk

How to relax dynamic stability?

$$\dot{x} = f(x) + Noise = -dV(x)/dx + Noise$$

$$\downarrow \quad \text{via parameterization}$$

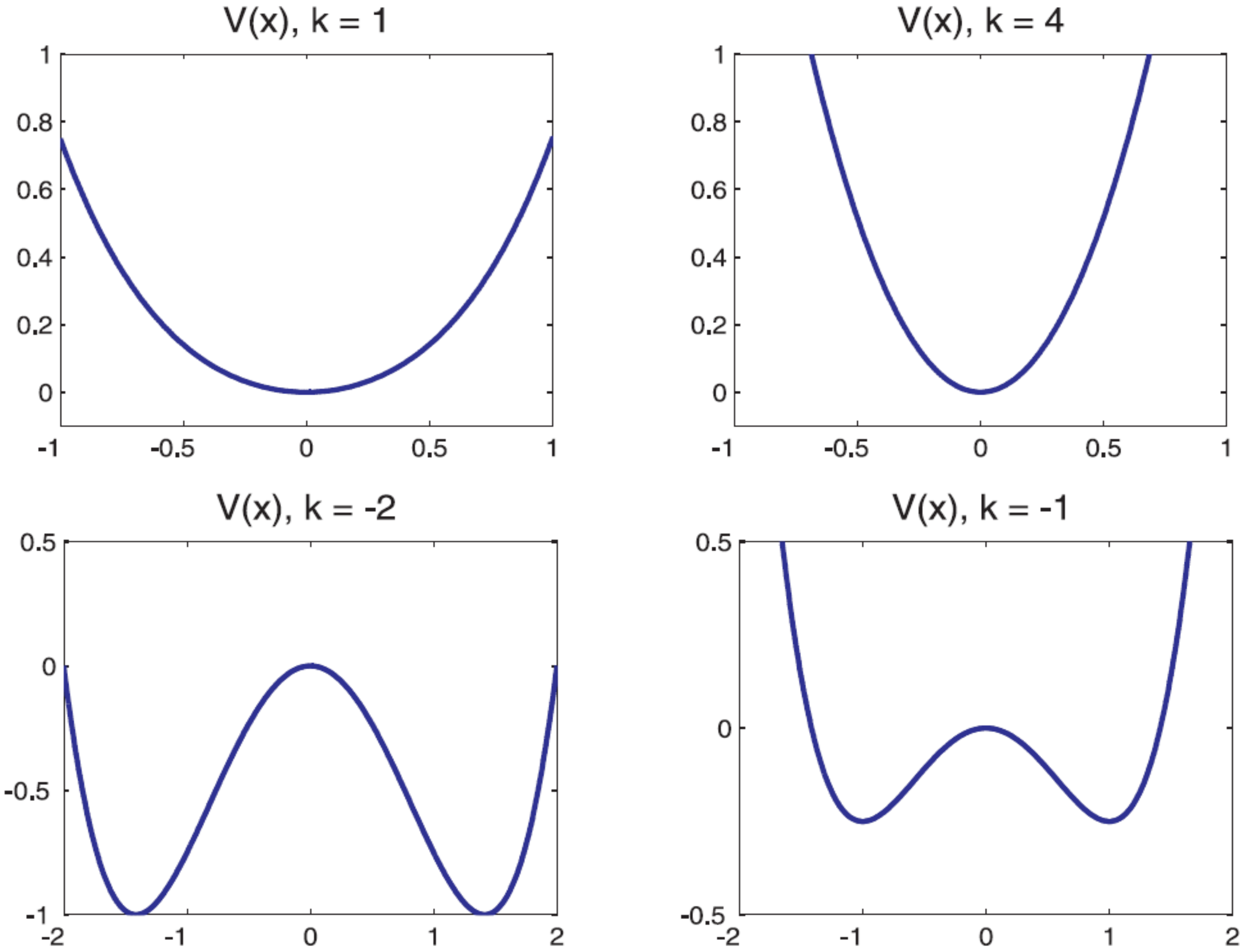
$$\dot{x} = f(x, P) + Noise = -dV(x)/dx + Noise$$

- In general, as *P* changes continuously, the corresponding solutions to our equation also change continuously. But, when *P* crosses a critical value the system may change qualitatively or discontinuously.

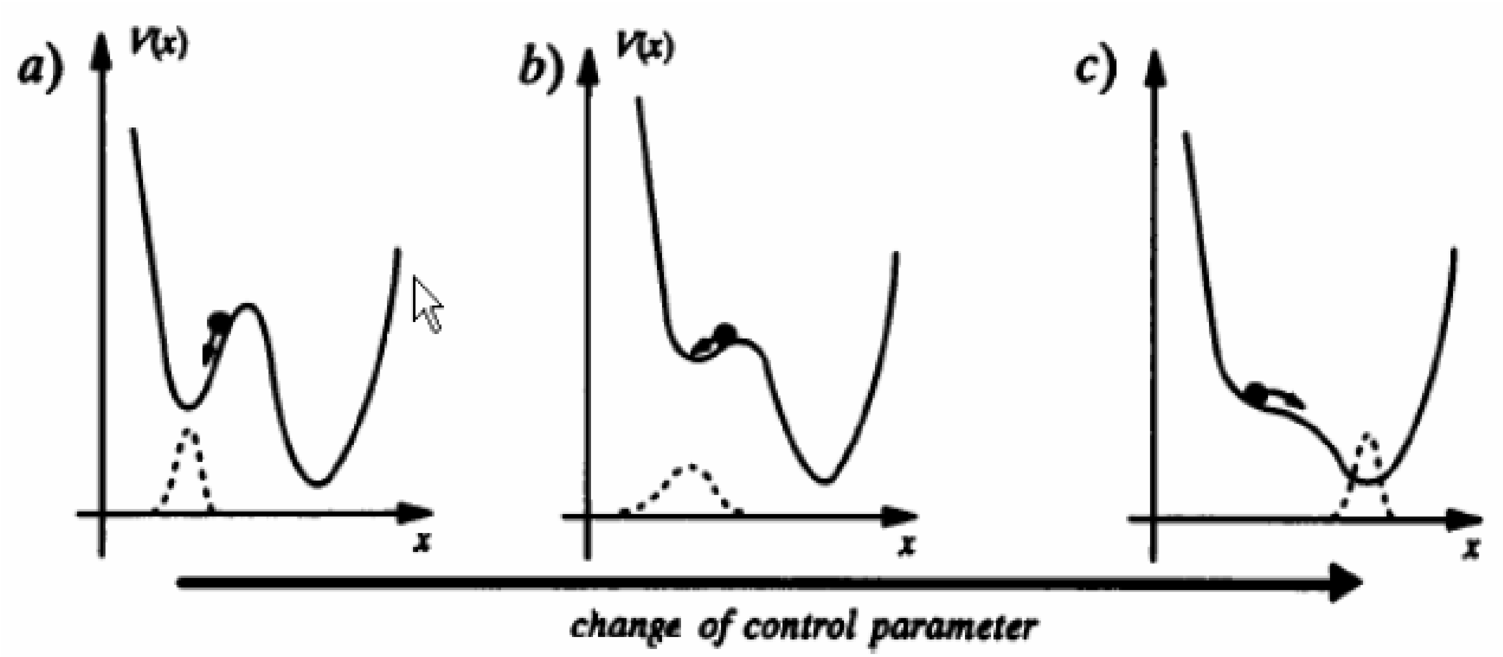
sbenus@ukf.sk

Example: $f(x) = -kx - x^3$

- By integration: $V(x) = -\int f(x)dx = \frac{1}{2}\mathbf{k}x^2 + \frac{1}{4}x^4$ (+ C)



Loss of an attractor as control parameter changes



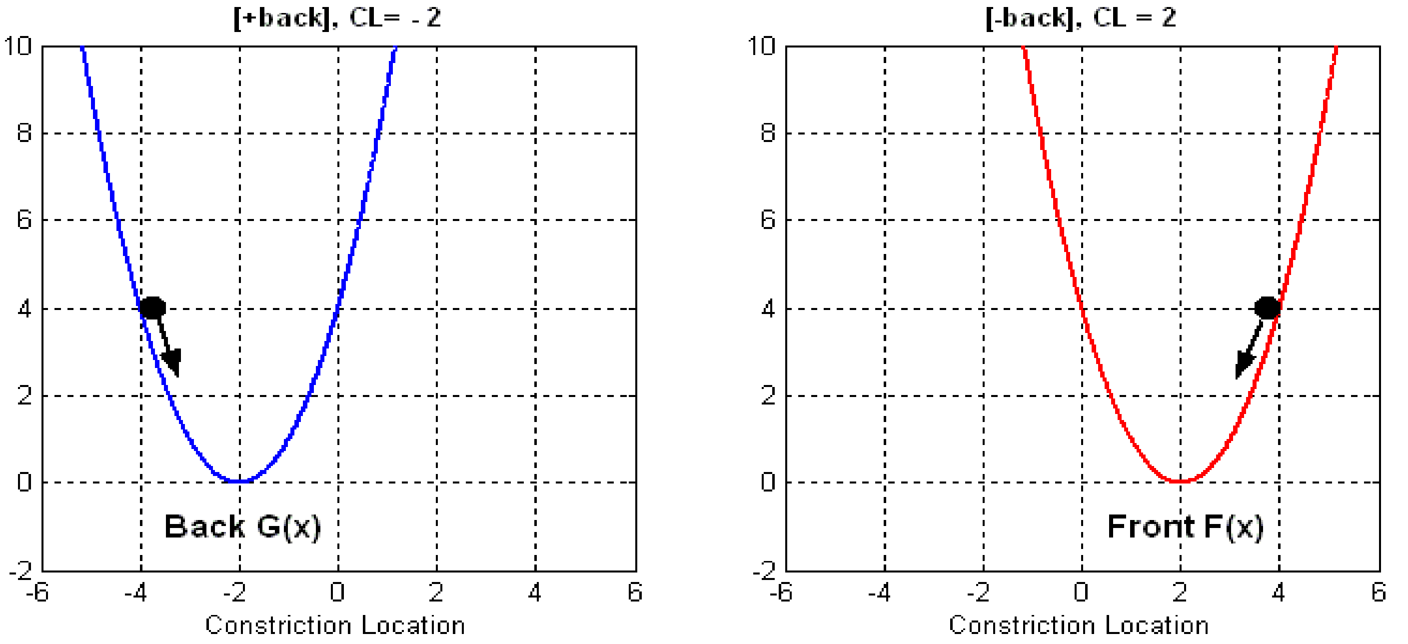
Kelso et al. 2003

e.g. finger-wagging

sbenus@ukf.sk

Dynamics of vocalic targets

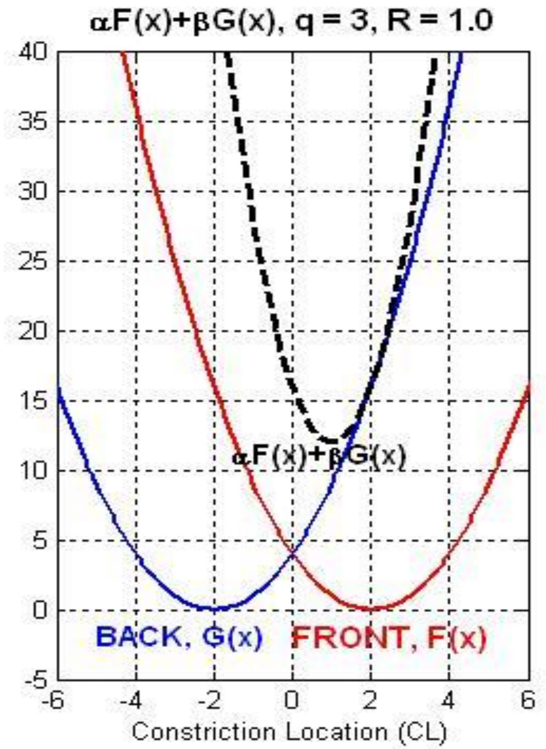
- Monostable landscape $V(x) = \alpha(x - x_0)^2$, where x_0 represents the CL target value, front or back.



sbenus@ukf.sk

Articulatory blending formally

- Simplest working hypothesis: linear combination of input potentials, αF(x) + βG(x), α β are the weights of the individual gestures, q = α/β.



Perturbations of vowel constriction location due to blending are captured with the degree of retraction **R**

sbenus@ukf.sk

Working hypothesis for suffix dynamics

- Since suffixes alternate between a front and a back version, the suffix dynamics must afford at least two attractors.
- Given this requirement (Arnold 2000), a good candidate for f(x, R) is the function

$$f(x, R) = R + x - x^3$$

- For our purposes:

$$f(x, R) = (3R - 2) + x - x^3$$

$$V(x, R) = (2 - 3R)x - \tfrac{1}{2}x^2 + \tfrac{1}{4}x^4$$

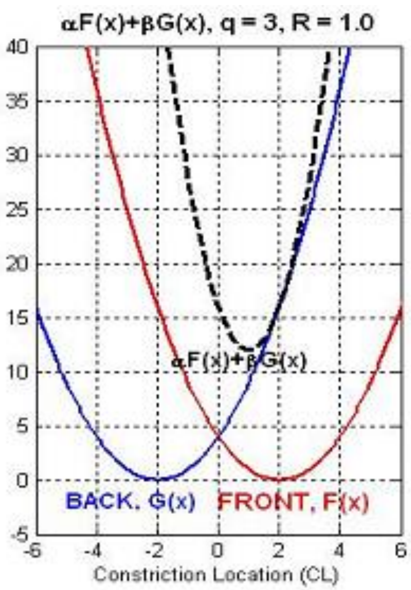
Transparent vowels: significant retraction

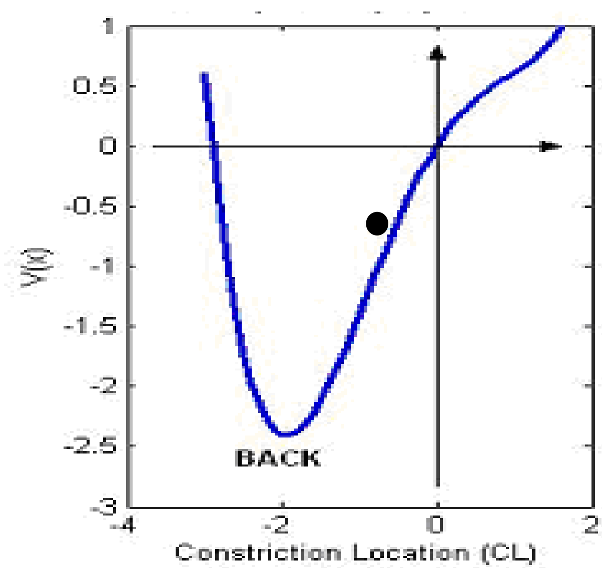
'papír-nak'

a={uvul., wide}
$CL_a = -2$

í={pal.,nar.}
$CL_í = 2$

V={___, wide}
CL = ??(a/e)





$$V(x,R) = (2-3R)x - x^2/2 + x^4/4 = -x - x^2/2 + x^4/4$$

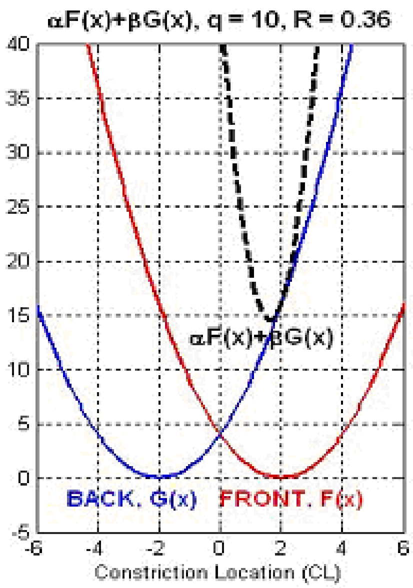
Opaque vowels: small retraction

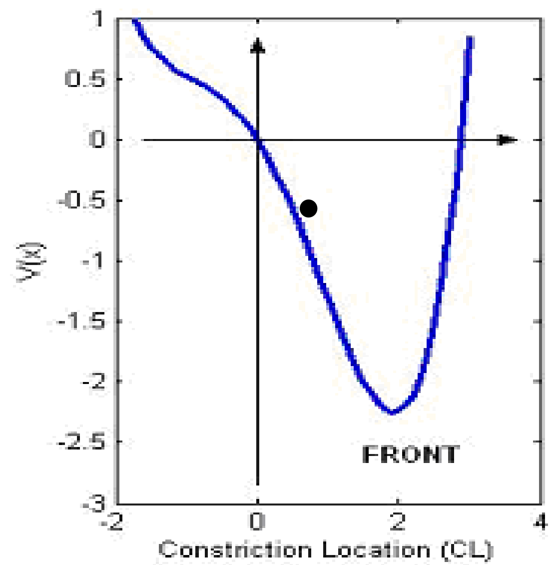
‘parfüm-nek’

a={uvul., wide} CL_a = -2

ü={pal.,nar.} CL_i= 2

V={____, wide} CL = ??(a/e)





$V(x,R) = (2-3R)x - x^2/2 + x^4/4 = .9x - x^2/2 + x^4/4$

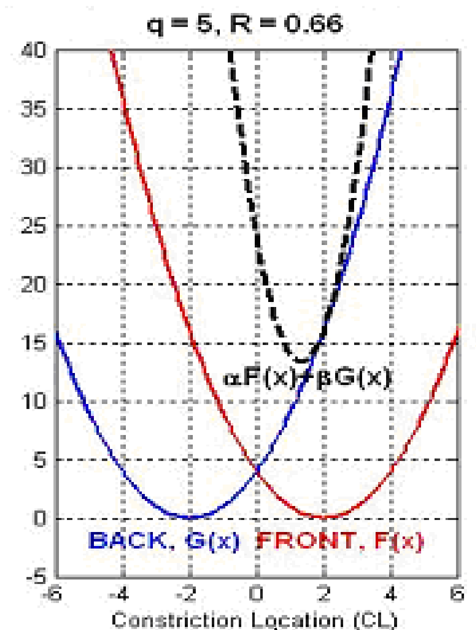
Less transparent /e/: intermed. retraction

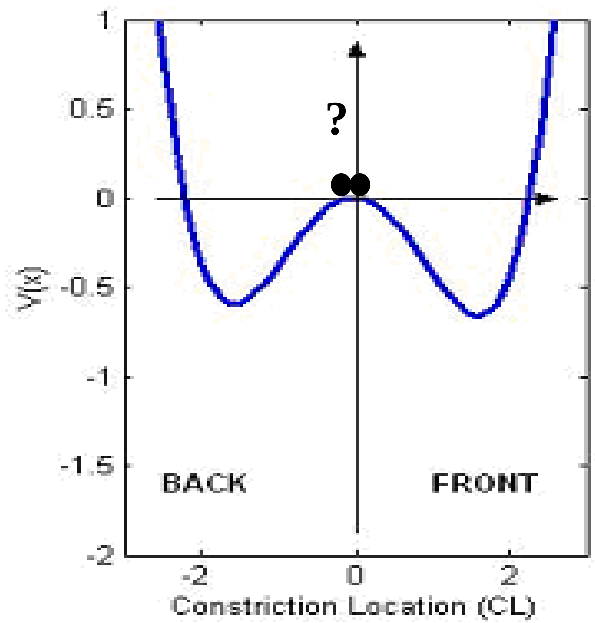
‘hárem-n{a/e}k’

á={uvul., wide} CL_a = -2

e={pal.,wide} CL_e= 2

V={___, wide} CL_0 = ??(a/e)





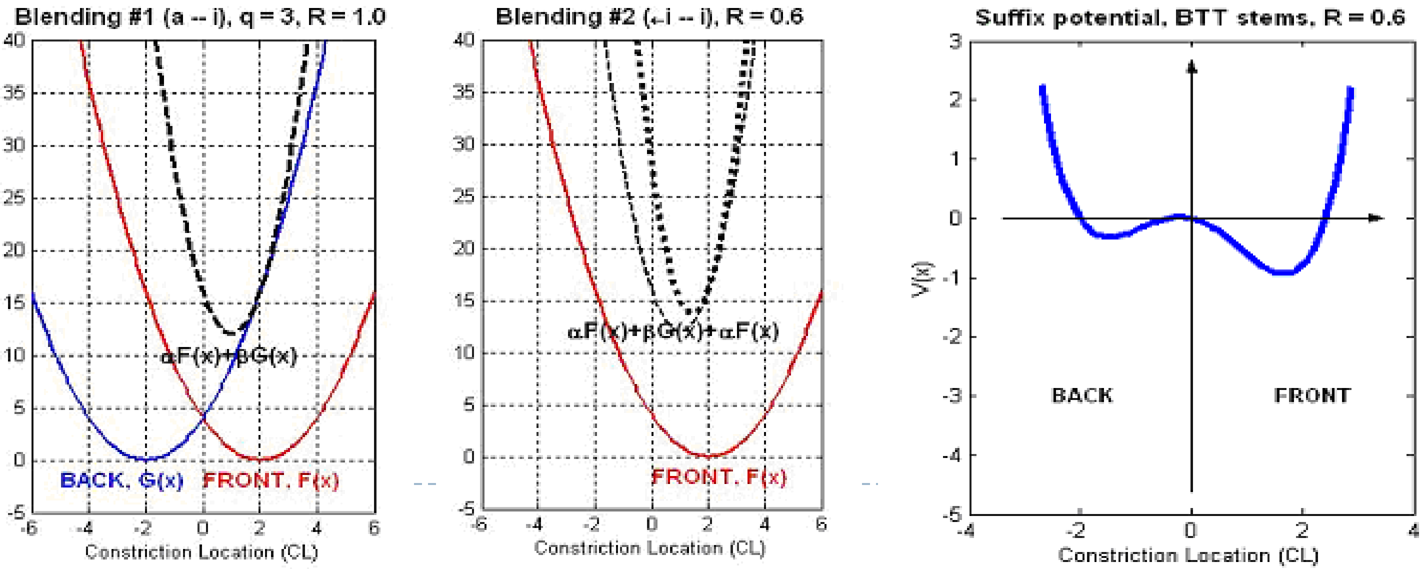
$$V(x,R) = (2-3R)x - x^2/2 + x^4/4 = .1x - x^2/2 + x^4/4$$

Multiple transparent vowels

- BTT stems are more likely to vacillate or take front suffixes than BT stems (*mami-nak* vs. *aszpirin-n{a/e}k*)
- This is predicted by the model

‘aszpirin-n{a/e}k’

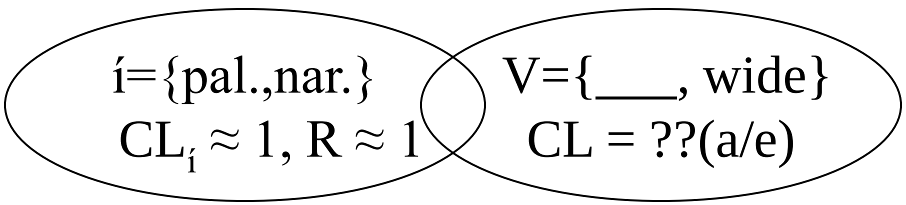
a={uvul., wide} CL_a = -2	i={pal.,nar.} CL_i= 2	i={pal.,nar.} CL_i= 2	V={___, wide} CL = ??(a/e)

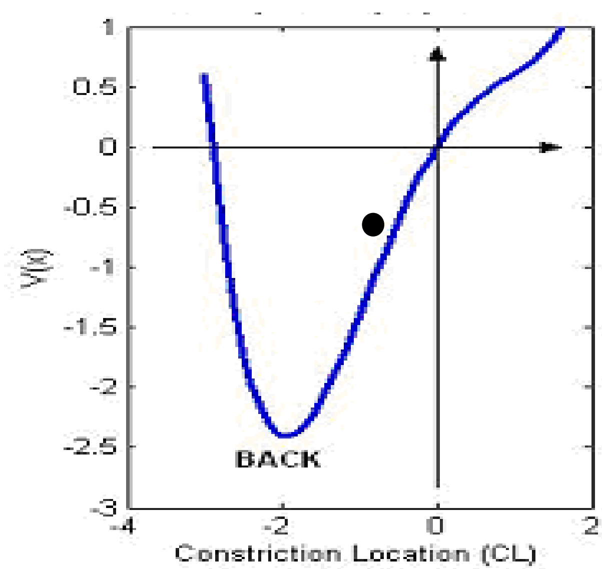


Exceptional mono-syllabic stems

Retraction is lexically specified, suffix selection does not proceed ‘on-line’, rather, the relationship between the retraction degree and the suffix is phonologized.

‘híd-nak’

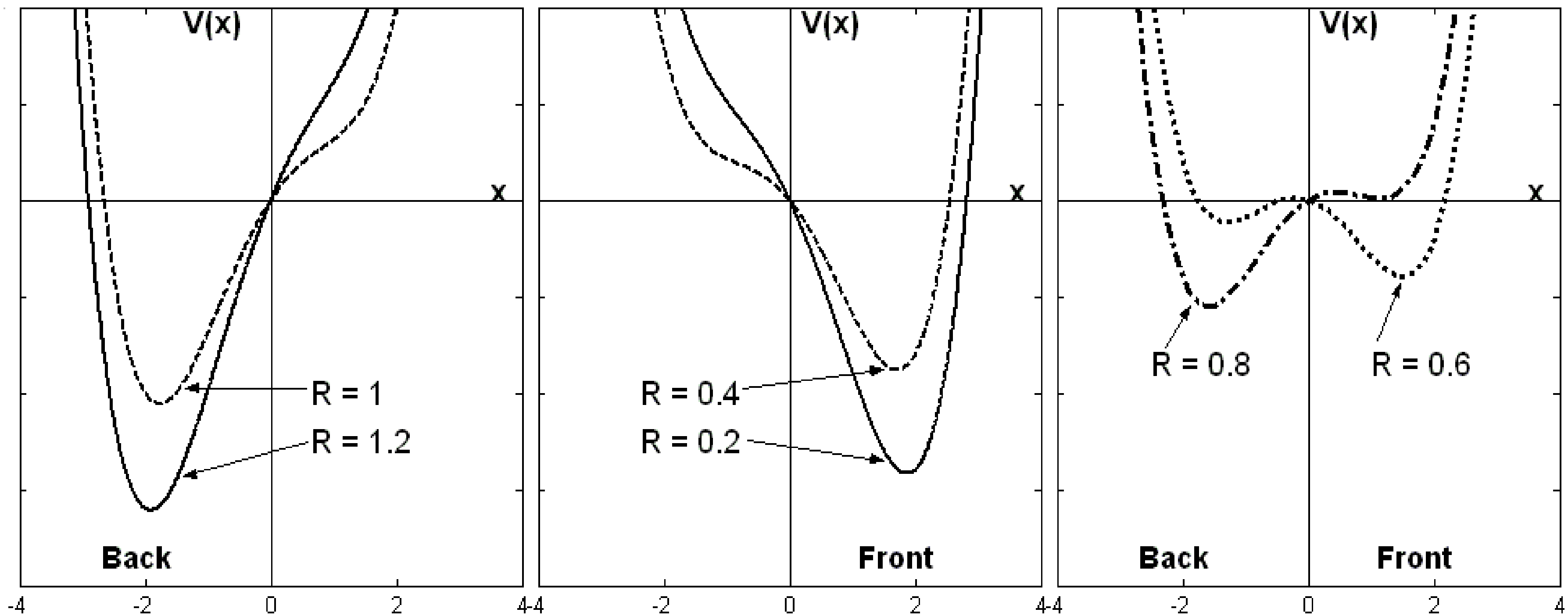




$$V(x,R) = (2-3R)x - x^2/2 + x^4/4 = -x - x^2/2 + x^4/4$$

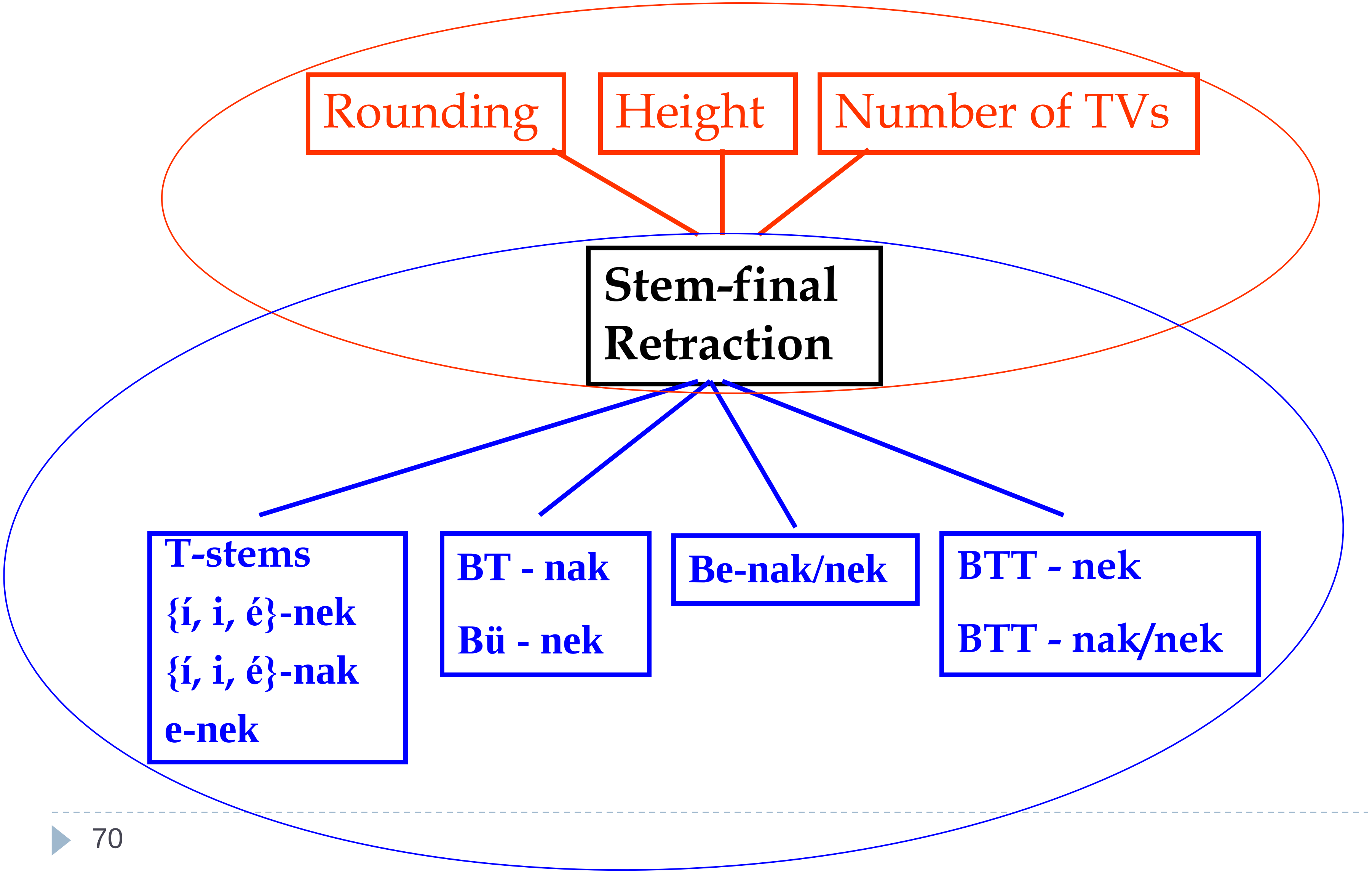
Suffix form as a function of *R*

- maximal (*papír-nak*)
- minimal (*emir-hez*)
- intermediate (*aszpirin–nak/nek*)

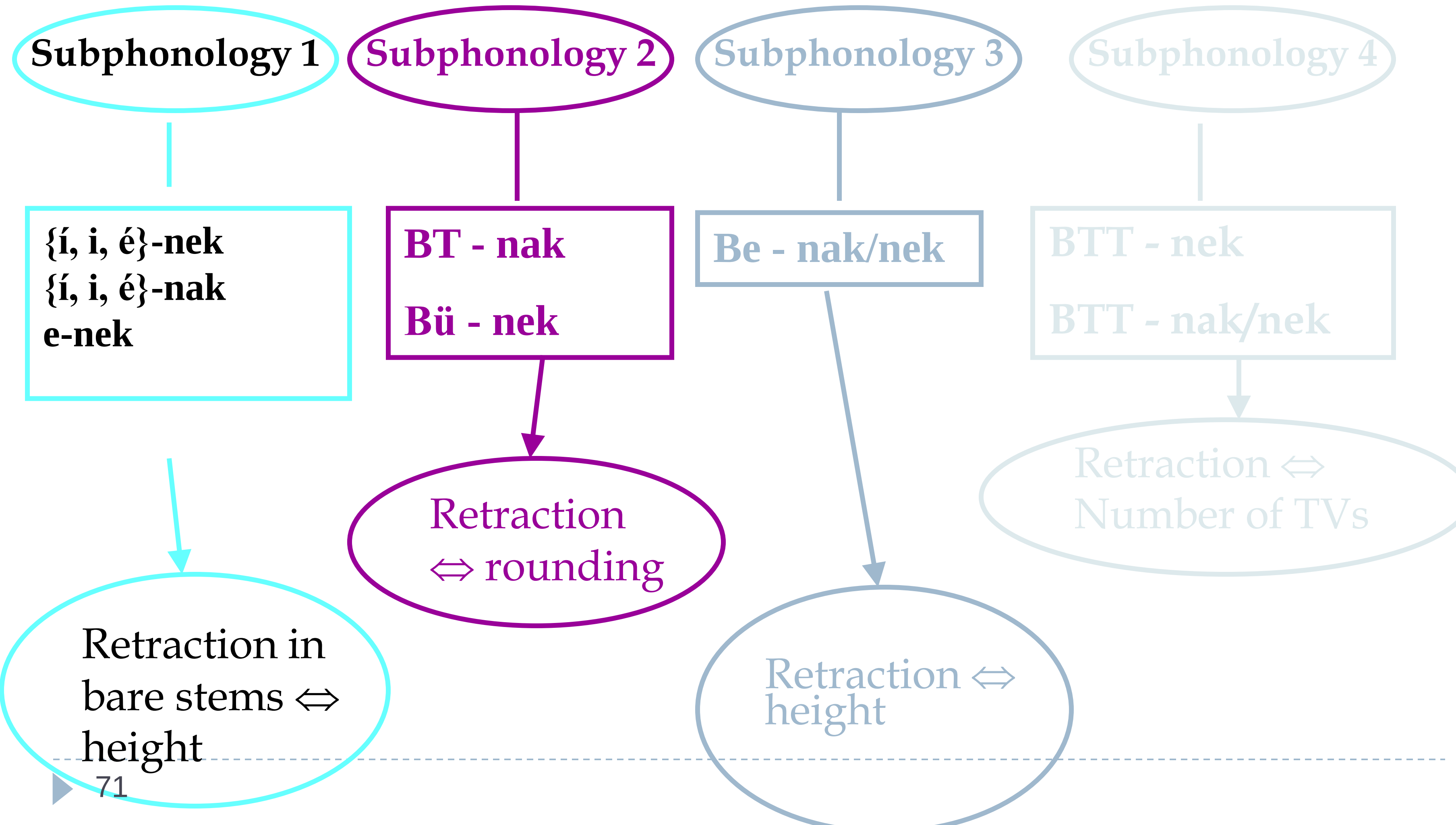


Variability as scaling and bifurcation

Transparency in an integrated phonology-phonetics



Transparency in a segregated phonology-phonetics



Phonological cognition in non-linear dynamics

- Grammar/cognition is construed as the attractor landscape defined by differential equations
  - Stable attractors in this landscape describe the stable coarse-grained generalizations observed in the phonological descriptions
  - ‘Phonetic’ substance is inseparably linked to ‘phonological’ form, obviating thus problems with transduction and grounding
  - Variability in speech: discrete-like as bifurcations, continuous-like as scaling, both linked to changes in continuous control parameters
  - Spatio-temporal domain accessible both in the representation as well as in the grammar, potential to analyze development at various time scales
  - Suitable for speech perception and effects of frequency or type of experience on perception (e.g. Nguyen et al. 07) and has potential to account for effects from other speech domains (social structures, pragmatic intensions, etc)
- Relation of stability-variation to sound changes and system-reconstruction into another stable state as a result of increased variation
- Are there discrete-like operations here? YES! We suggest that discrete operations provide a close-enough approximation of various instances of NL grammar.
- Issues: how is this system constrained?
  - The notion of phonological contrast
  - Computation is in continuous interaction with the environment in which it is embedded: embodiment as a constraining factor

Ďakujem za pozornosť !

sbenus@ukf.sk