

“SOUTHERN MIN CHINESE TONE SANDHI AS POLAR BOUNDARY TONES”

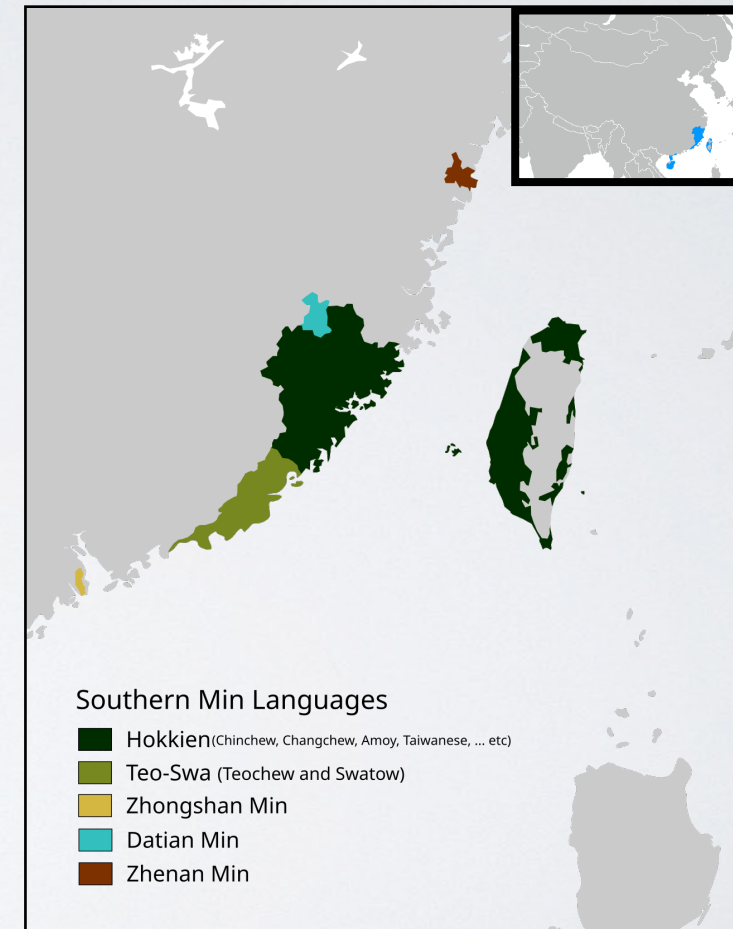
Andrew Nevins and Nicholas Rolle
University College London / Princeton University

Atelier de Phonologie March 2026

(For manuscript, including extensive references: <https://ling.auf.net/lingbuzz/009182>)

SOUTHERN MIN CHINESE (SMC) TONE SANDHI

- Southern Min (Minnan)
- Hokkien (Hoklo, Hagu, Holo)
- Taiwanese (Taigi)
- Xiamen (Amoy)
- Fujian (Fujianese, Fukienese)

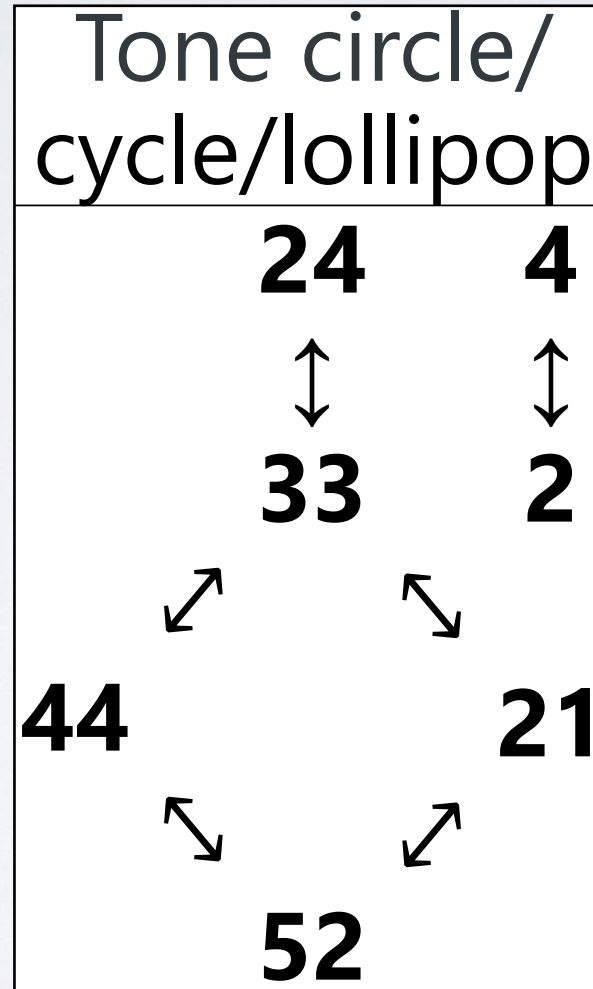


SOUTHERN MIN CHINESE (SMC) TONE SANDHI

Tone	Non-Final	↔	Final/Iso.
T5 <i>sî</i>	33 (mid)	↔	24 (rise)
T2 <i>sí</i>	44 (high)	↔	52 (high-fall)
T1 <i>sì</i>	33 (mid)	↔	44 (high)
T7 <i>sī</i>	21 (low-fall)	↔	33 (mid)
T3 <i>sì</i>	52 (high-fall)	↔	21 (low-fall)
T8 <i>sik</i>	2 (checked-low)	↔	4 (checked-high)
T4 <i>sik</i>	4 (checked-high)	↔	2 (checked-low)

Non-final position	Final position
<i>si</i> ³³ <i>kan</i> ⁴⁴ 'time'	<i>sui</i> ³³ <i>si</i> ²⁴ 'anytime'

SOUTHERN MIN CHINESE (SMC) TONE SANDHI



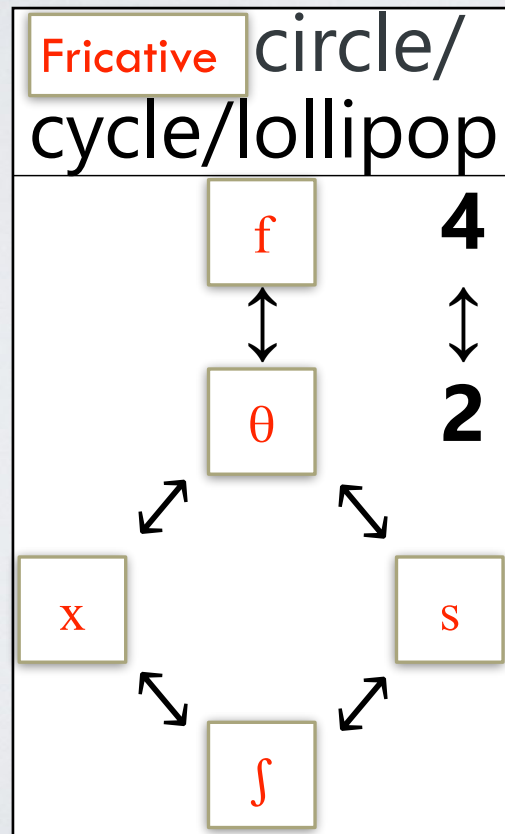
IMAGINE IF THIS WERE SEGMENTS...
 (SCIENCE FICTION PHONOLOGY)

*	Input		Output
Alt 1	ta f	→	ta θ
Alt 2	ta θ	→	ta s
Alt 3	ta s	→	ta f
Alt 4	ta f	→	ta x
Alt 5	ta x	→	ta θ

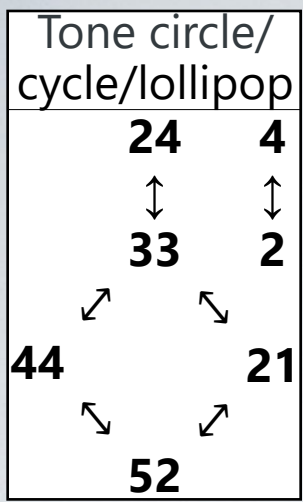
*	Input		Output
Alt 1	f at	→	θ at
Alt 2	θ at	→	s at
Alt 3	s at	→	f at
Alt 4	f at	→	x at
Alt 5	x at	→	θ at

IMAGINE IF THIS WERE SEGMENTS... (SCIENCE FICTION PHONOLOGY)

- Fricative non-phrase final → Fricative phrase-final



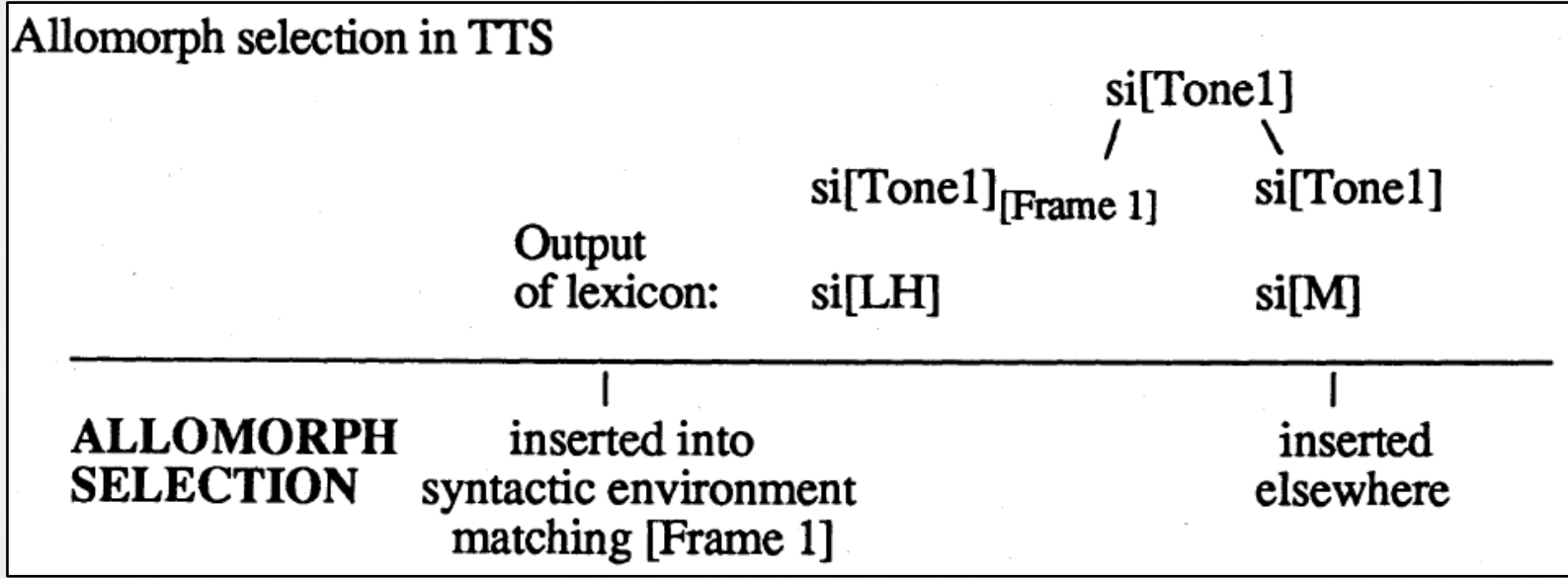
- a **taf** I like → a likable **taθ**
- a **taθ** I like → a likable **tas**
- a **tas** I like → a likable **taf**
- a **taf** I like → a likable **tax**
- a **tax** I like → a likable **taθ**

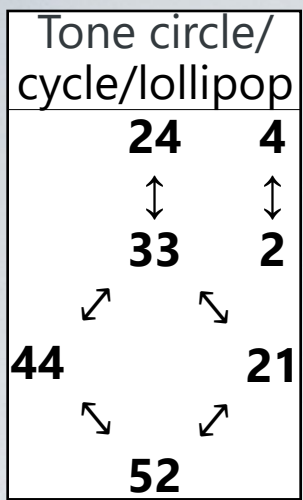


WHAT IS A POSSIBLE PHONOLOGICAL ALTERNATION?

OR, IS SMC TONE SANDHI EVEN 'REAL' PHONOLOGY?

- SMC Tone Sandhi: “**Not derivable**” (Moreton 1999); “Paradigmatic **replacement**” (Schuh 1978); “Arbitrary **substitution**” (Chen 2000)
- “The ... alternations are **suppletive**” (Tsay & Myers 1996)

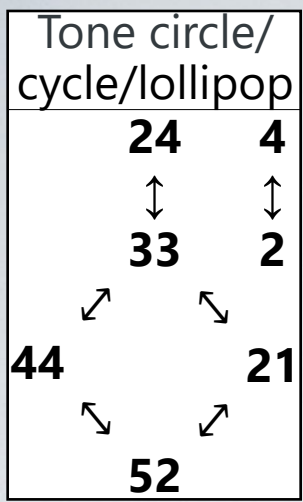




WHAT IS A POSSIBLE PHONOLOGICAL ALTERNATION?

OR, IS SMC TONE SANDHI EVEN 'REAL' PHONOLOGY?

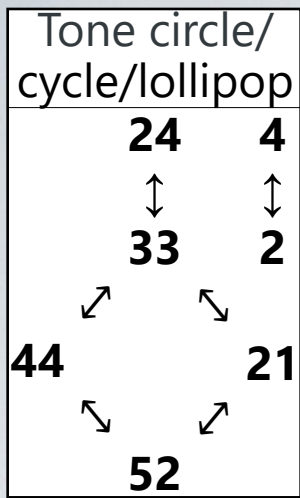
- Presumably these processes of “arbitrary substitution” and universal “suppletion” are available to all types of phonology — **So why do we only find it with tone** (and if then, very rarely indeed)?
- Tonal **exceptionalism**: “Tone can do everything that segmental or metrical phonology can do, but the reverse is not true” (Hyman 2011)
- But the question is ... **why?**



WHAT IS A POSSIBLE PHONOLOGICAL ALTERNATION?

OR, IS SMC TONE SANDHI EVEN 'REAL' PHONOLOGY?

- Tone gives you the right ingredients to generate these effects
- 1) Overt (i.e. phonologically contentful) tonal units frequently act as boundary marking for prosodic constituency (cf. rarity of segmental boundary features)
- 2) Tones can combine and co-occur in the same position (i.e. non-identical tones on two moras of the same vowel)
- 3) Tone often shows binary dissimilation: **H H → H L**
(cf. rarity of consonantal dissimilation **k k → g k**)

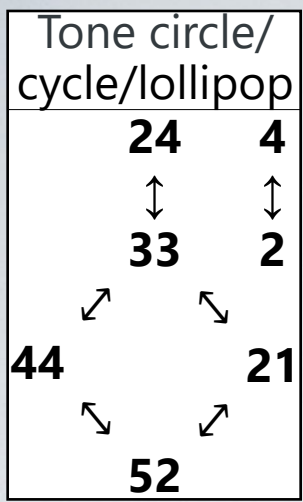


OUR PROPOSAL: SMC TONE SANDHI IS
 DERIVABLE AND NATURAL PHONOLOGY!

Our proposal: (i) XT tones
 decomposed into features,
 with (ii) tone sandhi due to
 a phrasal polar register
boundary tone ***R%***

PART I:

CITATION FORMS ARE DERIVED,
“SANDHI FORMS” ARE UNDERLYING



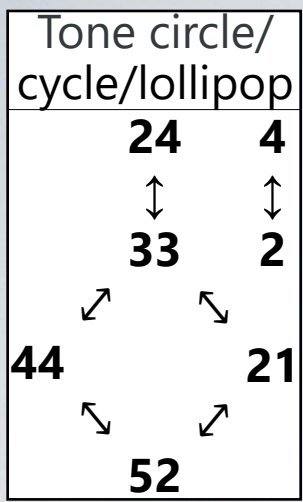
A MODEST REVISION: THE CITATION FORM IS DERIVED, NOT UNDERLYING

See Wright 1983 UM

Tone	Non-Final	↔	Final/Iso.
T5	<i>sî</i>		
	33 (mid)	↔	24 (rise)

Position 1: Final-as-basic /**24**/ → [**33**] / __ ...

Position 2: Non-final-as-basic /**33**/ → [**24**] / __ #



A MODEST REVISION: THE CITATION FORM IS DERIVED, NOT UNDERLYING

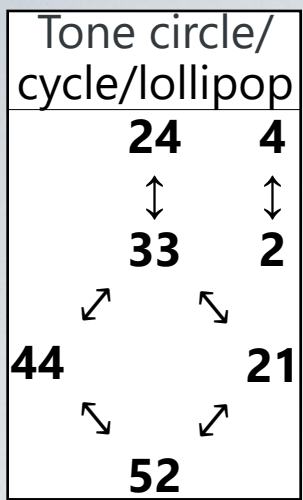
- **Simplicity argument:** Non-final-as-basic requires only **one** tone change

[44 24 24 21 52 44 21 21 52 44 52 2]
 yi kiong kiong kio gua ke k'uah puah tiam tsing ku ts'eq
 (he by force cause I more read half hour long book)
 'He insisted that I read for another half an hour'

- Alternate (the prevailing and default view): **All** tones but one change

[44 24 24 21 52 44 21 21 52 44 52 2]
yi kiong kiong kio gua ke k'uah puah tiam tsing ku ts'eq

- Considerations from language acquisition and speech planning (along the lines of Wagner 2012) would favor such a re-analysis, where it is only the last syllable that is changed and requires a special rule

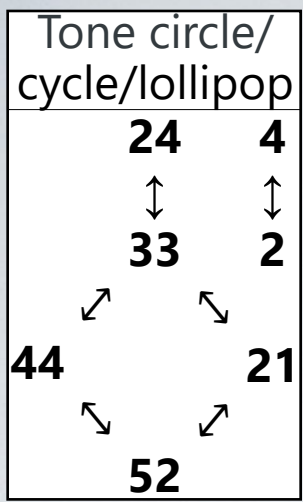


A MODEST REVISION: THE CITATION FORM IS DERIVED, NOT UNDERLYING

- Domain-final addition: Precedence in interpreting Mandarin 3rd tone realization

- Phrase-final [214] **ni**²¹⁴ ‘you’
- Non-final [21] **ni**²¹ kan⁵¹ ‘you look’
- Dissimilation [35] **ni**³⁵ hao²¹ ma⁰ ‘you good (QUES)’

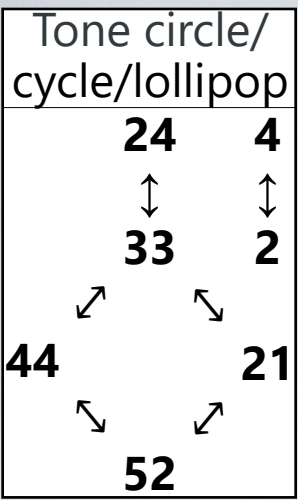
- [21] → LL: found in non-final position, before any other tone (except for another LL)
- [35] → LH: found as the result of dissimilation before another LL, whereby the sequence LL LL becomes LH LL
- [214] → LL(H): found in phrase-final position due to the contribution of an additional boundary tone (H)



A MODEST REVISION: THE CITATION FORM IS DERIVED, NOT UNDERLYING

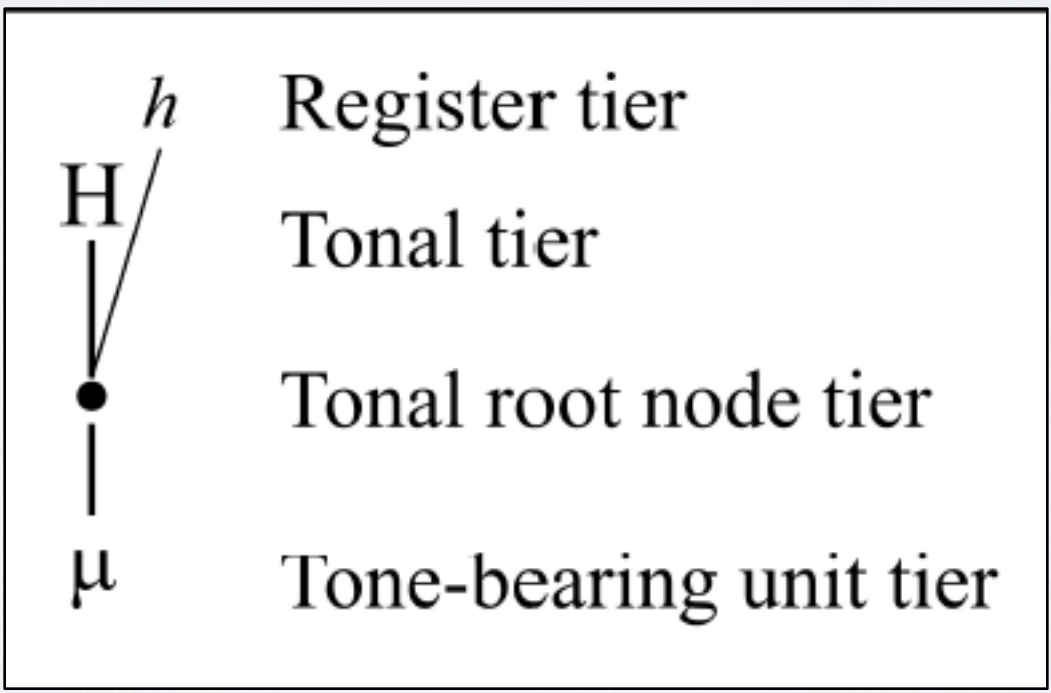
- Additional arguments (see Nevins & Rolle 2025)
 - Synchronic tonal alternations involving adjectival retriPLICATION, and nominalization with *-a*
 - Diachrony (proto-forms closer to “sandhi” forms)

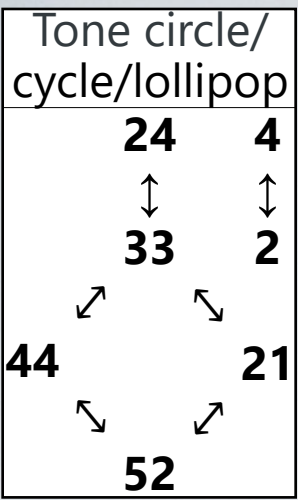
PART II:
UNDERLYING REPRESENTATIONS OF
SMCTONE CONTRASTS



UNDERLYING VALUES OF SEVEN BASIC TONEMES IN SMC

- Geometric representation of tone features (Yip 1980, Pulleyblank 1986, Snider 1999, Bao 1999, *inter alia*)

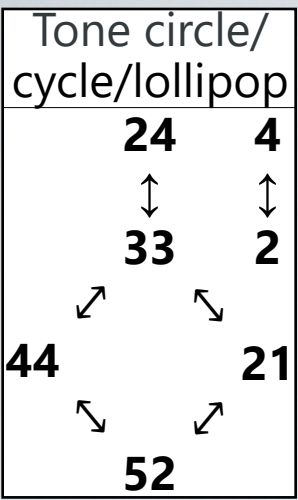




UNDERLYING VALUES OF SEVEN BASIC TONEMES IN SMC

- Syllables are **bimoraic** ($\mu\mu$), unless checked (end in obstruent — T8 and T4 below) which are monomoraic (μ)
- All representations begin with tonal feature [H]
- Tonal features only associate to single root node (\bullet)

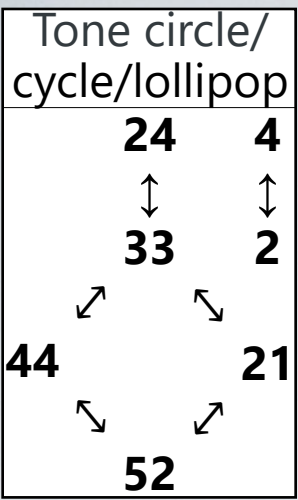
	T5	T2	T1	T7	T3	T8	T4
	33 ↔ 24	44 ↔ 52	33 ↔ 44	21 ↔ 33	52 ↔ 21	2 ↔ 4	4 ↔ 2
/UR/							



UNDERLYING VALUES OF SEVEN BASIC TONEMES IN SMC

- There is only **one** register feature per tone ([*h*] or [*l*])
- Register features only associate to first root node (●)
- Root node (●) only associates to **one** feature per tier (i.e. one tonal feature and one register feature), though it may associate to multiple moras (cf. T5 vs T1)

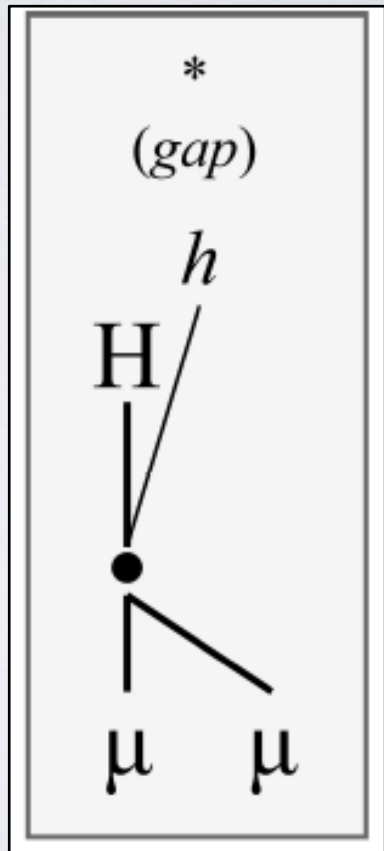
	T5	T2	T1	T7	T3	T8	T4
	33 ↔ 24	44 ↔ 52	33 ↔ 44	21 ↔ 33	52 ↔ 21	2 ↔ 4	4 ↔ 2
/UR/							



UNDERLYING VALUES OF SEVEN BASIC TONEMES IN SMC

- There is an incidental gap: No {H,h} linked to two moras (i.e. no “T1 equivalent” of T2)
- This is an underlying gap — May be derived (to be shown)

	T5	T2	T1	T7	T3	T8	T4
	33 ↔ 24	44 ↔ 52	33 ↔ 44	21 ↔ 33	52 ↔ 21	2 ↔ 4	4 ↔ 2
/UR/							



PART III:
DERIVING SMCTONE SANDHI

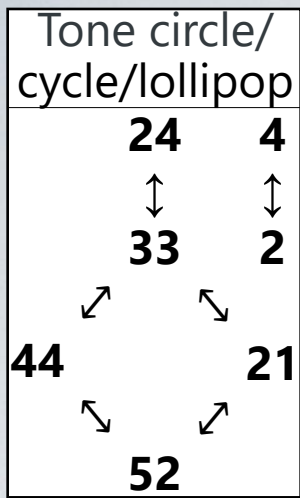
Tone circle/ cycle/lollipop	
24	4
33	2
44	21
52	

UR-TO-NON-FINAL FORM

- **Non-final forms:** Specifying underspecified material (*via* feature spreading)
- SMC “mids”: Phonetically identical but phonologically distinct

	T5	T2	T1	T7	T3	T8	T4
	33↔24	44↔52	33↔44	21↔33	52↔21	2↔4	4↔2
/UR/							
Non-final	[33]	[44]	[33]	[21]	[52]	[2]	[4]

- Moras map onto a pitch range
- {H,h}-moras map to [4] by default, but [5] before [2]
- Both {H,l}- and {L,h}-moras map to [2] by default, but to [3] if on adjacent μ s
- {L,l} always maps to [1]

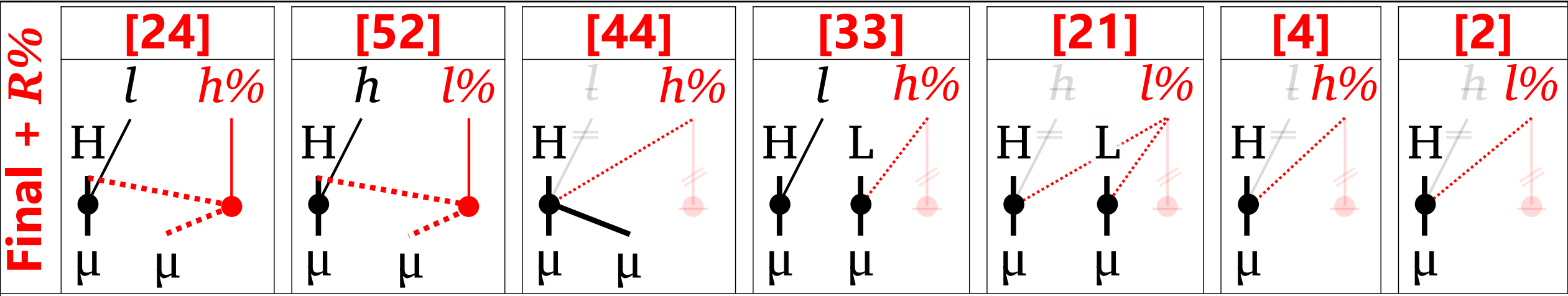


POLAR REGISTER BOUNDARY TONE AND ITS EFFECTS ON FINAL FORM

Final forms: Polar register boundary tone ***R%*** at end of phonological phrase (Polar boundary tone – Hyman 2007a; Register boundary tone – Gjersøe *et al.* 2019)

Tone circle/ cycle/lollipop	
24	4
33	2
44	21
52	

POLAR REGISTER BOUNDARY TONE AND ITS EFFECTS ON FINAL FORM



Polar *h%* docks on μ_2
H spreads to μ_2
First {H,*l*}: [2]
Second {H,*h*}: [4]

Polar *l%* onto μ_2
H spreads to μ_2
First {H,*h*}: [5]
Second {H,*l*}: [2]

Polar *h%*
No free μ
Register spreads to TRN (●), Then onto $\mu_1 + \mu_2$
{H,*h*}: [44]

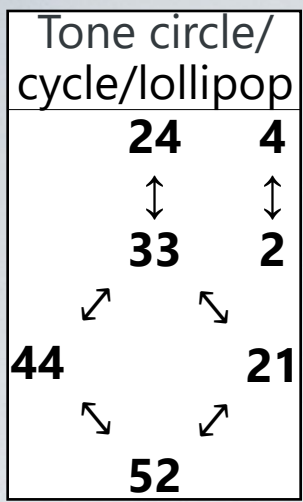
Polar *h%* onto μ_2
First {H,*l*}: [3]
Second {L,*h*}: [3]

Polar *l%* onto μ_1 and onto μ_2
*{H,*h*} {L,*l*}
First {H,*l*}: [2]
Second {L,*l*}: [1]

Polar *h%*:
{H,*h*}: [4]

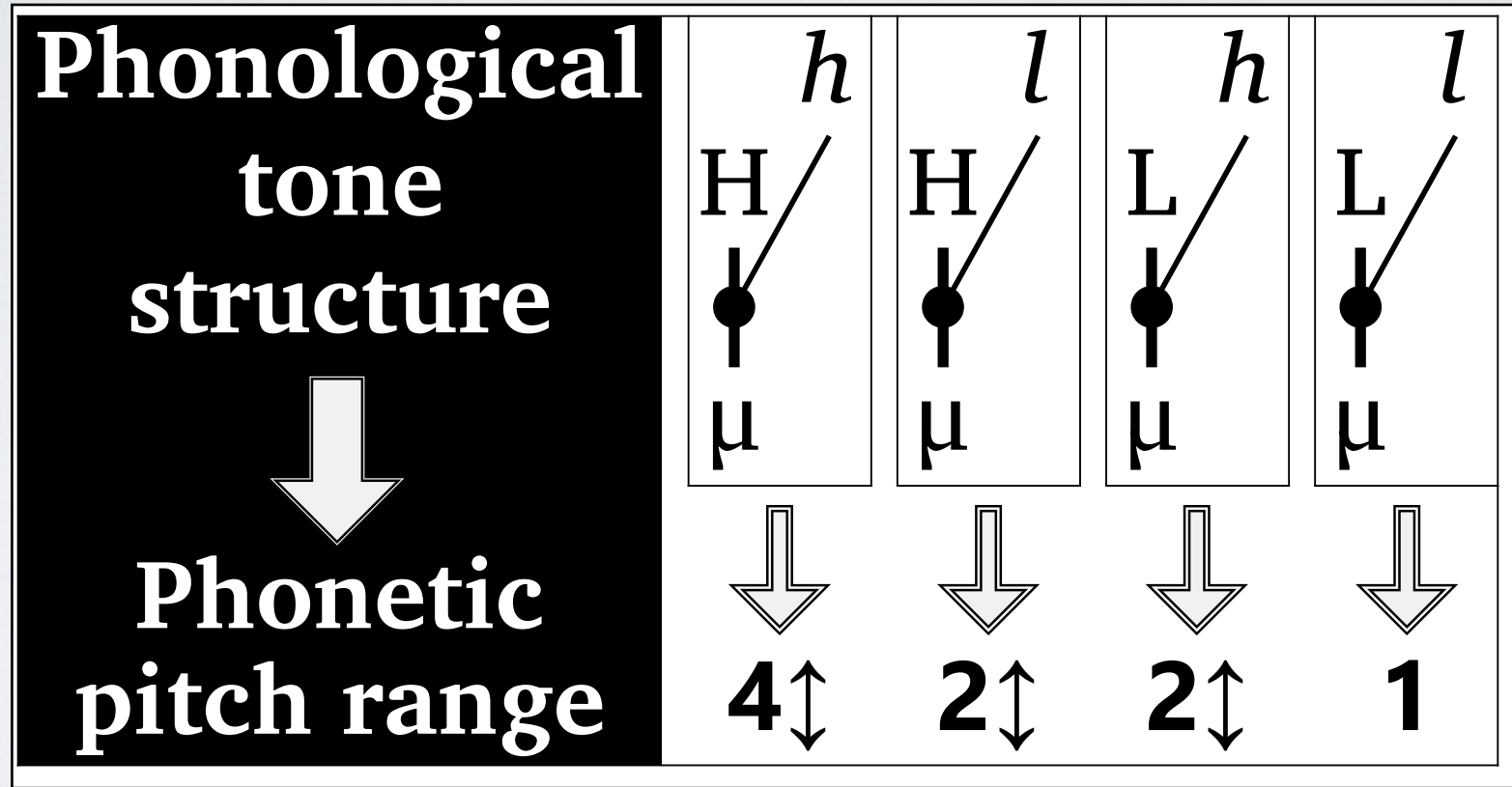
Polar *l%*:
{H,*l*}: [2]

PART IV: DISCUSSION

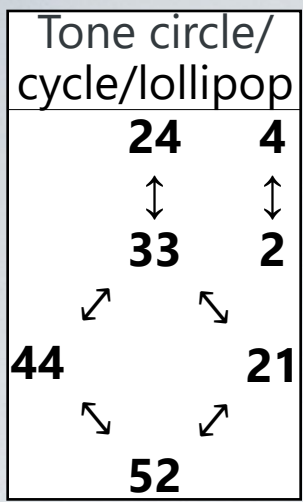


“CHAO NUMBERS” ARE PHONETIC AND CONTEXT-SENSITIVE

TONE & REGISTER ARE BINARY;
 “CHAO NUMBERS” ARE PHONETIC AND
 CONTEXT-SENSITIVE

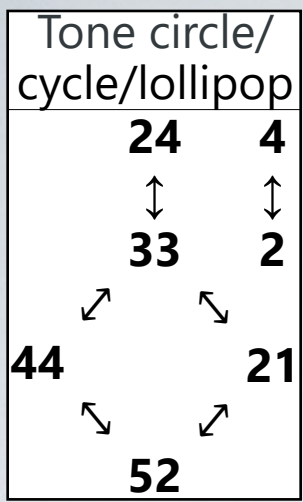


- See also Keating’s “Window Model of Phonetic (Under)Specification”



FINAL PARTICLES AND SO-CALLED “NEUTRAL” TONES

- When we talk about the **domain of the boundary tone**:
 - **R%** does not necessarily dock to the rightmost edge of the entire utterance, but to the final syllable with an overt tone of the **Tone Sandhi domain** (same for Mandarin as well)
 - We may call this Tone Sandhi domain the **Phonological Phrase** (or the **Accentual Phrase**), a prosodically intermediate constituent above the phonological word but below the utterance/intonational phrase (IP)



FINAL PARTICLES AND SO-CALLED “NEUTRAL” TONES

- Particle *kong* PART (< ‘say’) Simpson & Wu 2002)

• {a-hui} {kong a-sin} {m **lai**} ←— *Sentence-final words undergo tone sandhi*
 name say name NEG **come**
 ‘A-hui said A-sin is not coming’

• {a-sin} {m lai} **kong** ←— *Particle kong does not undergo tone sandhi!*
 name NEG come **PART**
 ‘A-sin is not coming’

Tone circle/ cycle/lollipop	
24	4
↕	↕
33	2
↙ ↘	
44	21
↘ ↙	
52	

FINAL PARTICLES AND SO-CALLED “NEUTRAL” TONES

- Particle *kong* PART (< ‘say’) Simpson & Wu 2002)

- $\left[\begin{array}{c} R\% \\ | \\ \{a-hui\}\varphi \end{array} \right. \begin{array}{c} \mathbf{kong} \\ | \\ a-sin \end{array} \left. \begin{array}{c} R\% \\ | \\ \{m \ lai\}\varphi \end{array} \right] \updownarrow$

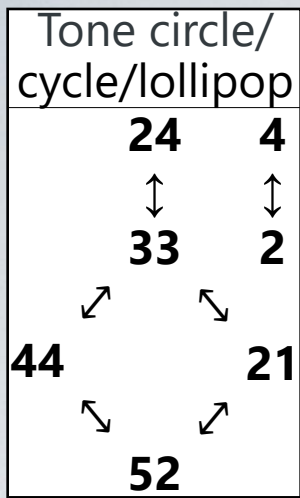
name **say** name NEG come

‘A-hui said A-sin is not coming’

- $\left[\begin{array}{c} R\% \\ | \\ \{a-sin\}\varphi \end{array} \right. \begin{array}{c} R\% \\ | \\ \{m \ lai\}\varphi \end{array} \left. \begin{array}{c} \mathbf{kong} \\ | \\ \mathbf{PART} \end{array} \right] \updownarrow$

name NEG come **PART**

‘A-sin is not coming’



FINAL PARTICLES AND SO-CALLED “NEUTRAL” TONES

- What kind of prosodic properties do these sandhi-external particles display?
 - One pattern is to simply continue the previous pitch trajectory (remaining phonologically ‘neutral’ or ‘unspecified’ into the mapping to phonetics)
 - But they may also bear **L%** (or more rarely **H%**) boundary tones / BTs (Wang & Fon 2011)

Tone circle/ cycle/lollipop	
24	4
↕	↕
33	2
↙ ↘	↙ ↘
44	21
↘ ↙	↘ ↙
52	

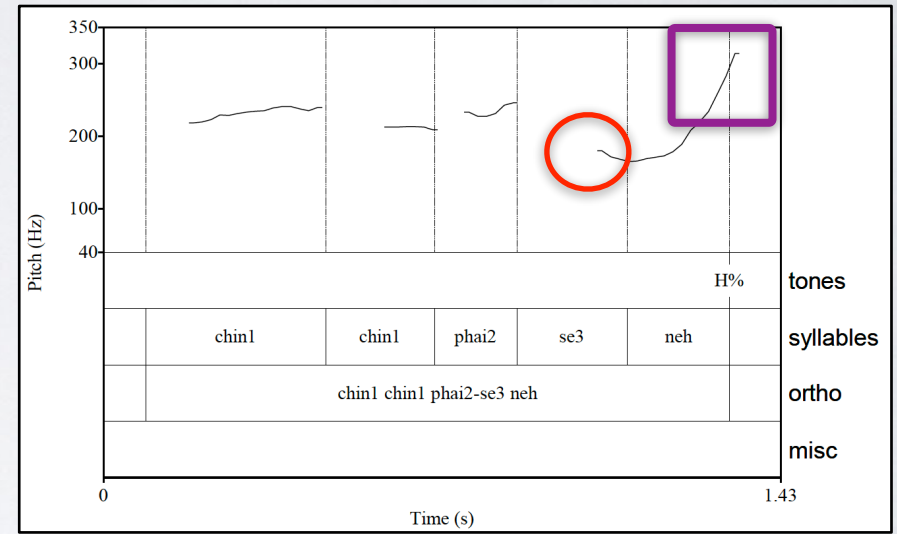
FINAL PARTICLES AND SO-CALLED “NEUTRAL” TONES

- Complementarity of phrase-level BTs (*R%*) and clause-level BTs (e.g. *H%*)

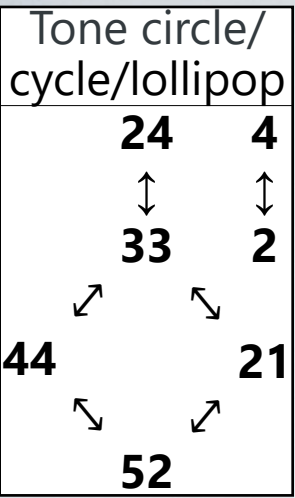
• [{ chin chin pháiⁿ-sè } φ **neh**] l

true true sorry **PART**

‘it’s so sorry’ (Wang & Fon 2011)

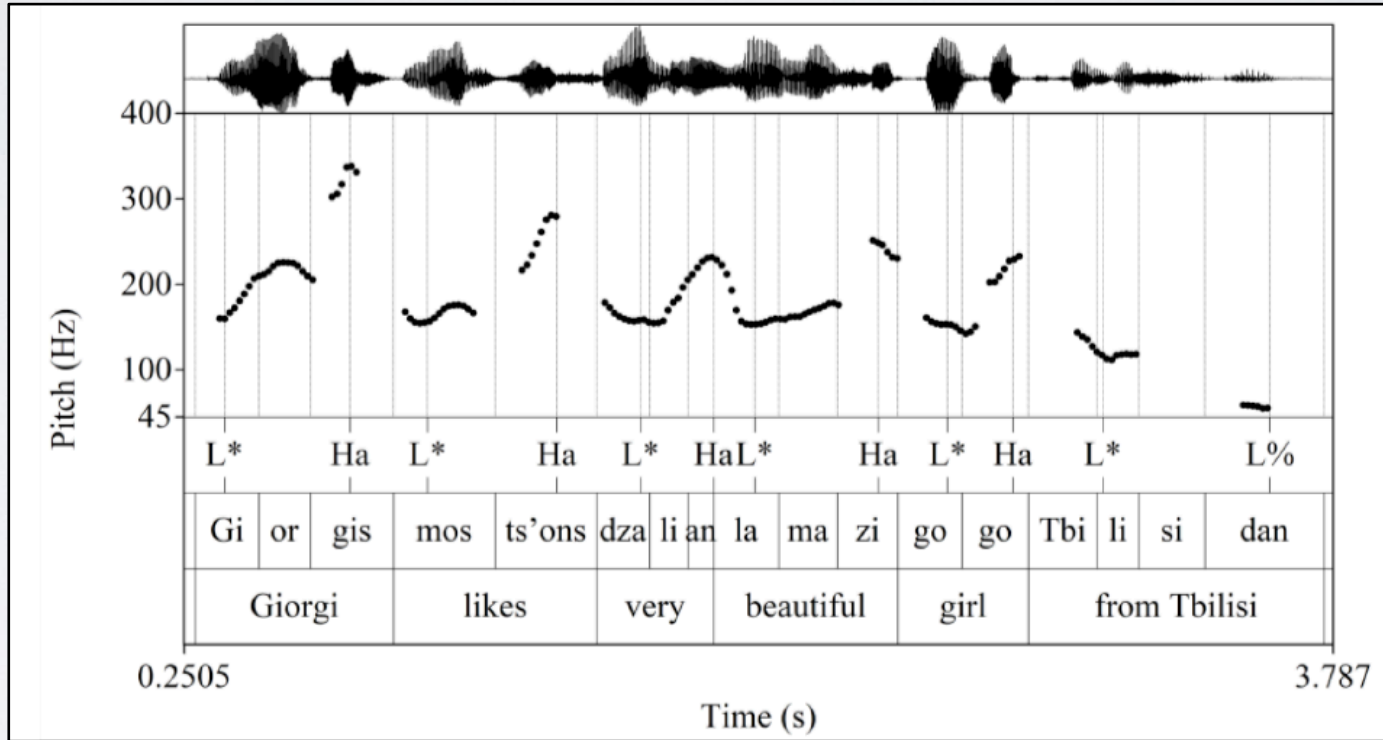


- “*H%* adds a “cute” connotation to the expression” ... “this type of usage is only found in the speech of one of the young speakers, it is possible that this usage has an idiosyncratic nature”

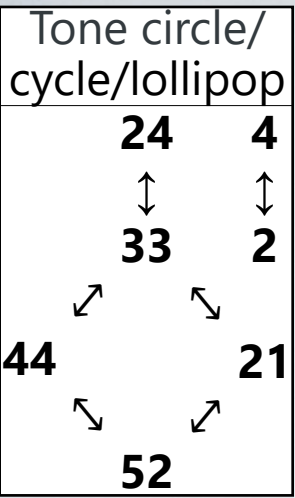


MULTIPLE BOUNDARY MARKING AKIN TO **INTONATION**

- Georgian (Borise 2023): BTs at **aP**-level vs. **iP**-level

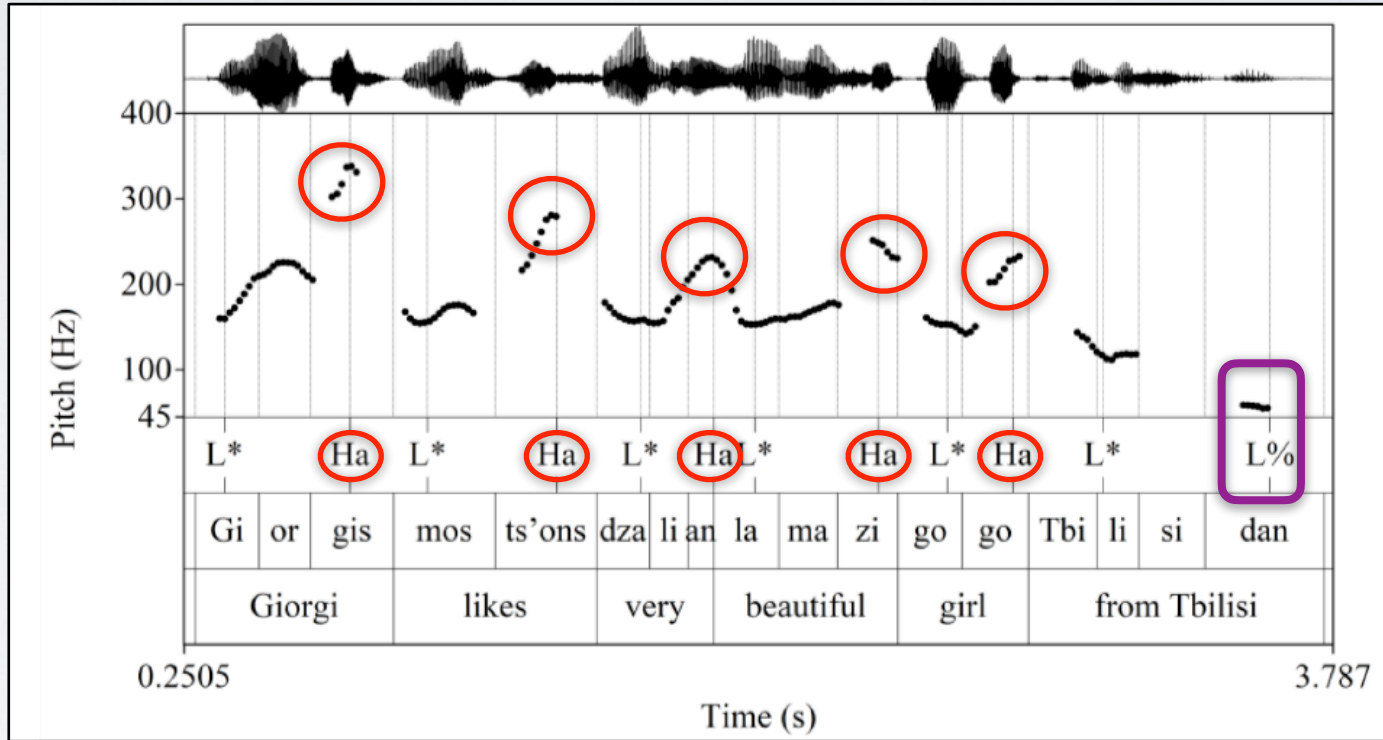


(Giorgi-s)_{AP} (mosts'on-s)_{AP} (dzalian)_{AP} (lamaz-i)_{AP} (gogo)_{AP} (Tbilisi-dan)_{AP}.
 Giorgi-DAT like-PRS.3SG very beautiful-NOM girl.NOM Tbilisi-from
 'Giorgi likes a very beautiful girl from Tbilisi.'



MULTIPLE BOUNDARY MARKING AKIN TO INTONATION

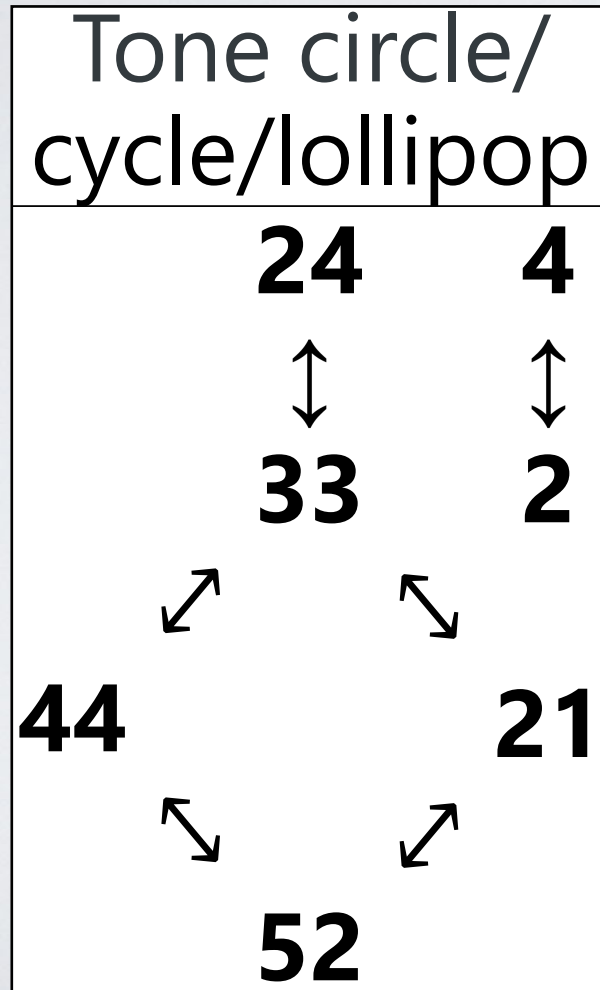
- Georgian (Borise 2023): BTs at **aP**-level vs. **iP**-level



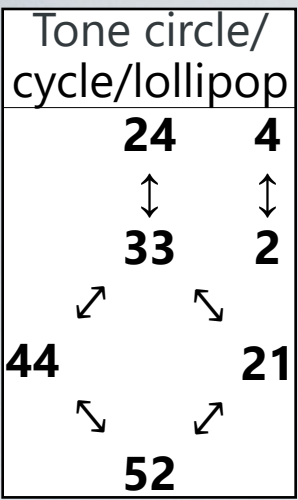
(Giorgi-**s**)_{AP} (mosts'on-**s**)_{AP} (dzalian)_{AP} (lamaz-**i**)_{AP} (gogo)_{AP} (Tbilisi-dan)_{AP}.
 Giorgi-DAT like-PRS.3SG very beautiful-NOM girl.NOM Tbilisi-from
 'Giorgi likes a very beautiful girl from Tbilisi.'

PART V: SUMMARY

SUMMARY OF OUR PROPOSAL FOR SMC TONE SANDHI

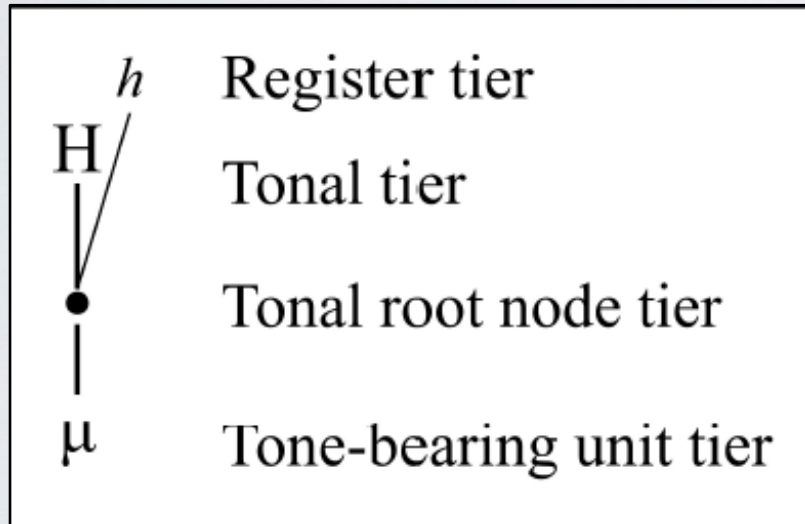


- SMC Tone Sandhi: “**Not derivable**” (Moreton 1999); “Paradigmatic **replacement**” (Schuh 1978); “Arbitrary **substitution**” (Chen 2000)
- “The ... alternations are **suppletive**” (Tsay & Myers 1996)



SUMMARY OF OUR PROPOSAL FOR SMCTONE SANDHI

- Geometric representation of tone features



Our proposal: (i) XT tones decomposed into features, with (ii) tone sandhi due to a phrasal polar register boundary tone **R%**

