INTONATION IN A LINGUISTIC AREA
BUILDING AN INTONATIONAL TYPOLOGY OF SOUTH ASIA

SAMEER UD DOWLA KHAN, REED COLLEGE
UNIVERSITY OF WASHINGTON
FRIDAY 2 FEBRUARY 2024

Goals for this talk

▶ Introduce you to some of the work I'm doing

▶ Describe a method for approaching intonation in an unfamiliar lg

▶ Share findings from applying this approach to a diverse range of South Asian lgs (SALs)

▶ Argue that the intonation of SALs is more complex than previously thought…

▶ …even just at the word-level

Goals for this talk

▶ (Why) is the intonation of SALs interesting?

▶ SALs have little to structure their intonation

  ▶ **No lexical stress** *(*oversimplification!)*

    ▶ Phonetically weak, if at all present

    ▶ Predictable location, if at all present *(*oversimplification!)*

  ▶ **No lexical tone** aside from TB lgs & some IA neighbors

▶ What does this mean for their intonation?

What I do

▶ "I work on all speech sounds other than Cs and Vs"

▶ Voice quality/phonation

  ▶ Breathiness in Gujarati

  ▶ Creak in American English

▶ Intonational phonology

  ▶ Deep view: Bangladeshi Bengali

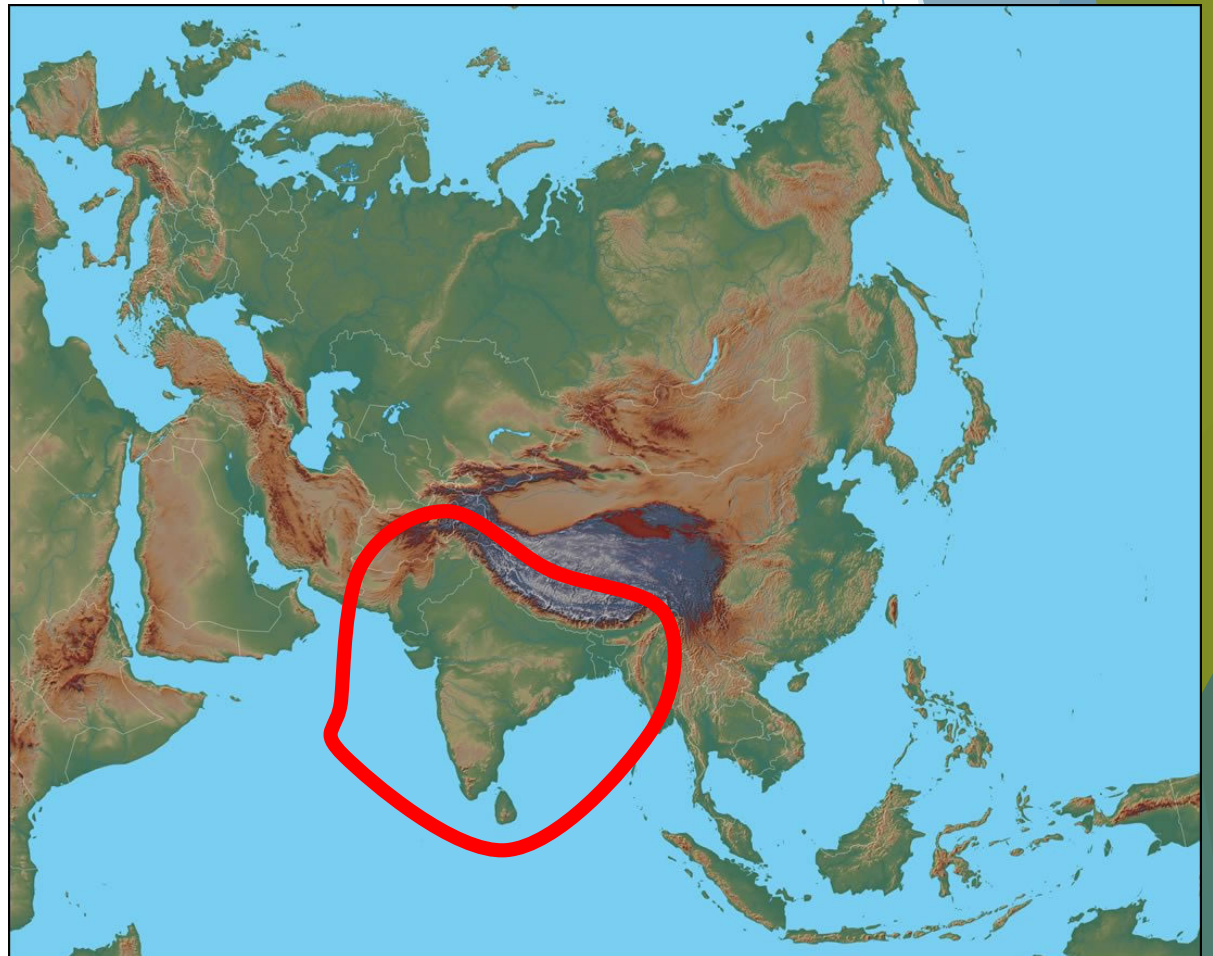 ▶ Broad view: South Asian lgs in general

Sameer (they/he)

Background

▶ South Asia includes:

    ▶ Bangladesh

    ▶ Bhutan

    ▶ India

    ▶ Maldives

    ▶ Nepal

    ▶ Pakistan

    ▶ Sri Lanka

▶ Also sometimes:

    ▶ Tibet

    ▶ Burma

    ▶ Afghanistan

Background

▶ South Asia has four large lg families:

   ▶ **Indo-European**

      ▶ Indo-Aryan (IA)

      ▶ Nuristani

      ▶ Iranian

   ▶ **Dravidian** (Dr)

   ▶ **Austroasiatic** (AA)

      ▶ Munda

      ▶ Khasic

      ▶ Nicobarese

   ▶ **Tibeto-Burman** (TB)

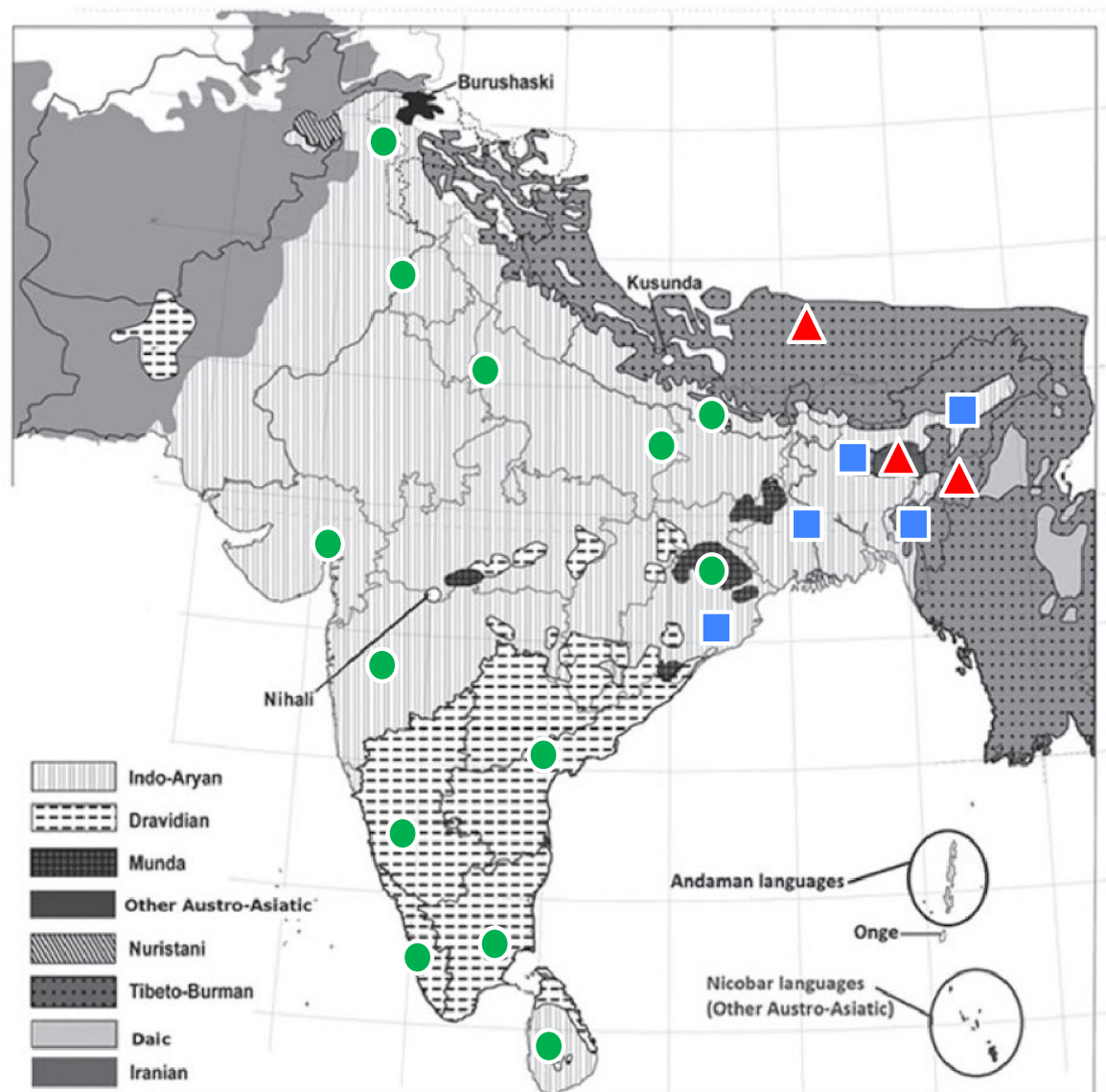South Asian language families (map produced by Suresh Kolichala, 2015)

Background

▶ But it's also identified as a region of convergence: a "**linguistic area**" (Emeneau 1956)

▶ Even across language families, South Asian lgs (SALs) typically have:

  ▶ Retroflexion (ṭ ḍ ṇ ḷ... [ʈ ɖ ɳ ɭ...])

  ▶ Echo reduplication ("doctor-<u>voctor</u>" = "doctors etc.")

  ▶ SOV order ("they rice eat")

  ▶ Non-NOM experiencers ("<u>to me</u> hunger strikes")

  ▶ Vector verbs ("they written-<u>took</u> the address")

Background

▶ But it's also identified as a region of convergence: a "**linguistic area**" (Emeneau 1956)

▶ Even across language families, South Asian lgs (SALs) typically have:

  ▶ Retroflexion **(except Assamese + most TB lgs)**

  ▶ Echo reduplication **(but what pattern? v-, s-, gi-, -u-, Ø...)**

  ▶ SOV order **(except Kashmiri + Khasic)**

  ▶ Non-NOM experiencers **(but which case? *see next slide*)**

  ▶ Vector verbs

Background

- ▶ Experiencers:
- ▶ 'I'm hungry/angry' uses: (Subbārāo 2012)
  - DAT "to me" in Munda lgs, Dr lgs, most IA lgs
  - ▲ NOM "I'm/have" in Khasic, most TB lgs
  - ■ GEN "of me" in East IA lgs; some TB lgs

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Questions

▶ So if there's a typical SAL in all these other areas + exceptions + smaller patterns…

▶ …is there also a **typical SAL intonation**?

   ▶ If so, what are the common characteristics?

   ▶ What are the exceptions and smaller patterns?

▶ What does **having no tone or stress contrasts** (in most lgs) do to the intonation system?

▶ What **new phenomena** do SALs contribute to our knowledge of intonational systems?
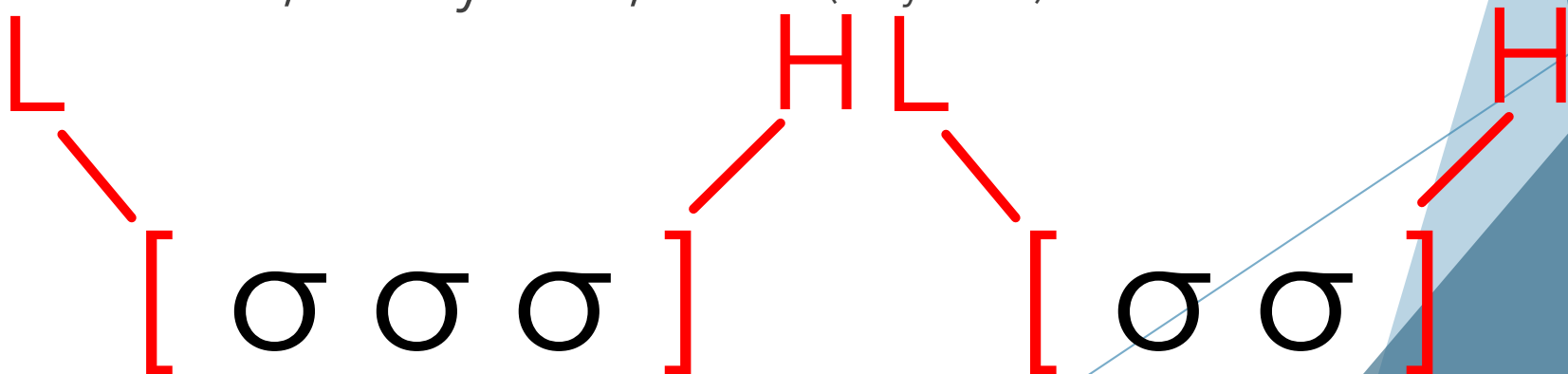
Background

- With no stress or tone contrasts, what does SAL intonation look like?

- Some claim they are **strikingly uniform** (Féry 2010, Hock 2016, Féry & Fanselow 2020)
    - **Identical patterns** across lgs
    - **LH pattern**: 2 boundary tones per word-ish-sized unit

- **"Phrase-language" model** of four SALs: Bengali, Hindi, Malayalam, Tamil (Féry 2010)

$$[\ \sigma\ \sigma\ \sigma\ ] \qquad [\ \sigma\ \sigma\ ]$$

Background

▶ With no stress or tone contrasts, what does SAL intonation look like?

▶ Some claim they are **strikingly uniform** (Féry 2010, Hock 2016, Féry & Fanselow 2020)

  ▶ **Identical patterns** across lgs

  ▶ **LH pattern**: 2 boundary tones per word-ish-sized unit

▶ **"Phrase-language" model** of four SALs: Bengali, Hindi, Malayalam, Tamil (Féry 2010)

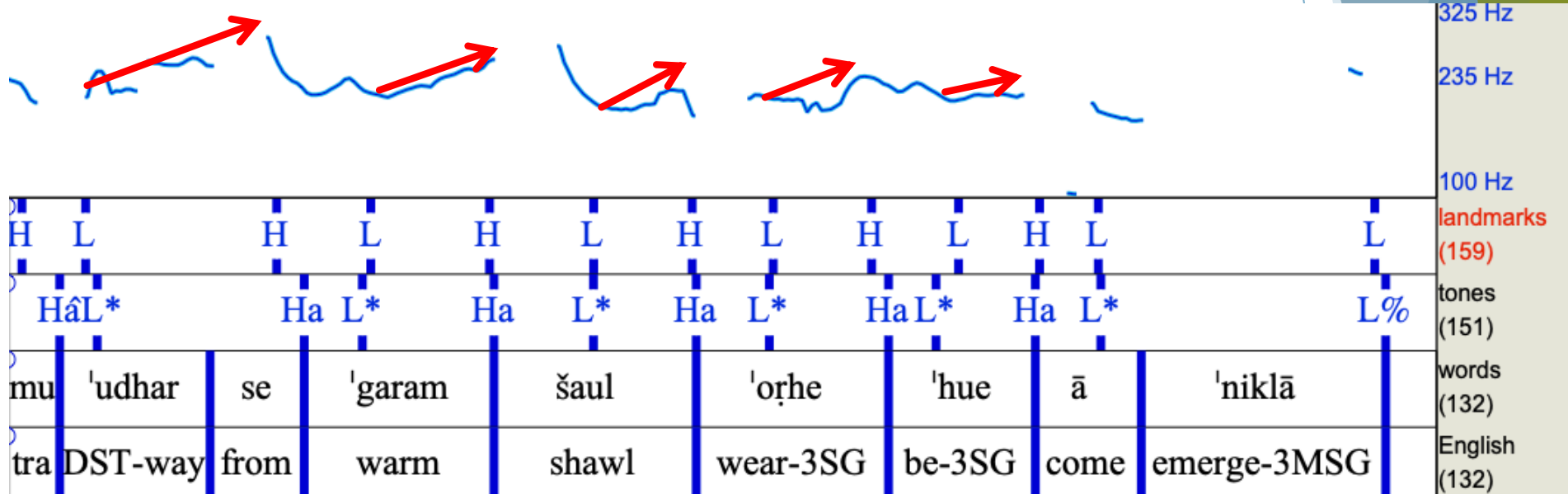$$L\quad [\ \sigma\ \sigma\ \sigma\ ]\quad H\ L\quad [\ \sigma\ \sigma\ ]\quad H$$

Background

- With no stress or tone contrasts, what does SAL intonation look like?

- Some claim they are **strikingly uniform** (Féry 2010, Hock 2016, Féry & Fanselow 2020)

  - **Identical patterns** across lgs

  - **LH pattern**: 2 boundary tones per word-ish-sized unit

- **"Phrase-language" model** of four SALs: Bengali, Hindi, Malayalam, Tamil (Féry 2010)

L   HL   H

[σ σ σ]   [σ σ]
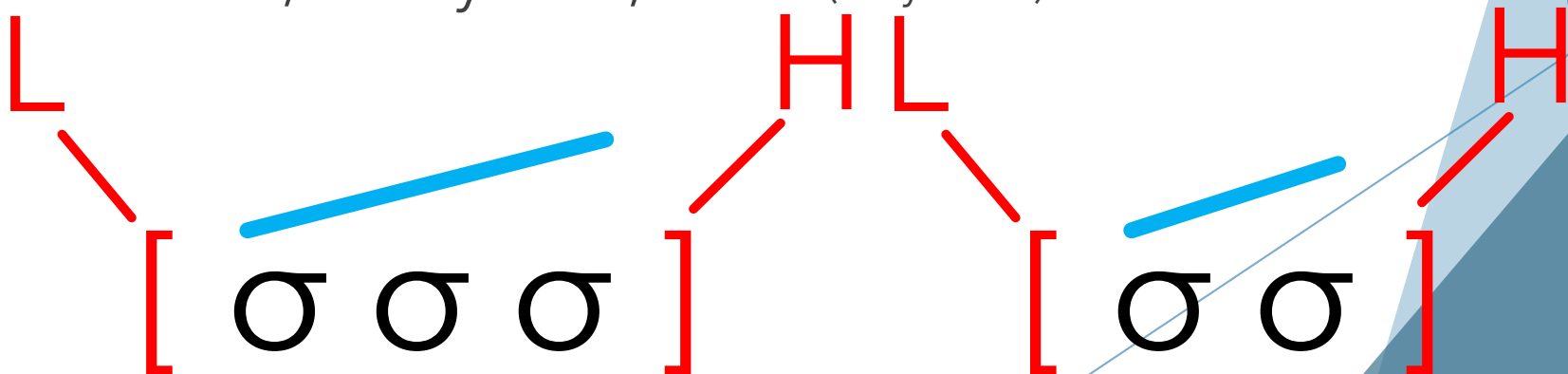
Preview of results

▶ Clip of **Urdu** demonstrating the "typical" SAL pattern described in literature: **LH**'s forever…

	325 Hz
	235 Hz
	100 Hz
H L H L H L H L H L H L L	landmarks (159)
HâL* Ha L* Ha L* Ha L* HaL* Ha L* L%	tones (151)
mu	'udhar	se	'garam	šaul	'oṛhe	'hue	ā	'niklā	words (132)
tra	DST-way	from	warm	shawl	wear-3SG	be-3SG	come	emerge-3MSG	English (132)

▶ '…emerged from there, wearing a warm shawl.'
<NWS-Urdu-R01b>

Preview of results

▶ **Yes**, we'll see there is a typical SAL intonation, but it is **build on many distinct patterns**

▶ My focus: **"word"-level intonation**: the **"LH"**

    ▶ Easier to relate to existing literature

▶ **Preview**: while the LH rise is indeed pervasive…

    ▶ …we need to **interpret exactly what it is**

    ▶ …and **add its many alternatives** to our models

Preview of results

▶ Look at this cute little LH. How might this simple-seeming rise be **analyzed**?

$$[\ '\overline{\sigma}\ \sigma\ ]$$

Preview of results

- Look at this cute little LH. How might this simple-seeming rise be **analyzed**?

  - L on initial, H on final

  - L on stress, H on final

  - L on stress, H follows

  - H on final, L on penult

  - LH anchored to stress

['σ σ]

Preview of results

▶ Look at this cute little LH. How might this simple-seeming rise be **analyzed**?

  ▶ L on initial, H on final

  ▶ L on stress, H on final

  ▶ L on stress, H follows

  ▶ H on final, L on penult

  ▶ LH anchored to stress

$$[\ \acute{\sigma}\ \sigma\ ]$$

▶ What happens when we look at **longer words**, or words with **different stress patterns**?

Preview of results

- In fact, we will see **evidence of all these interpretations** of the rise, across different SALs:

  - L on initial, H on final
    $$[\,\acute{\sigma}\,\sigma\,\sigma\,]\ [\,\sigma\,\acute{\sigma}\,\sigma\,]$$

  - L on stress, H on final
    $$[\,\acute{\sigma}\,\sigma\,\sigma\,]\ [\,\sigma\,\acute{\sigma}\,\sigma\,]$$

  - L on stress, H follows
    $$[\,\acute{\sigma}\,\sigma\,\sigma\,]\ [\,\sigma\,\acute{\sigma}\,\sigma\,]$$

  - H on final, L on penult
    $$[\,\acute{\sigma}\,\sigma\,\sigma\,]\ [\,\sigma\,\acute{\sigma}\,\sigma\,]$$

  - LH anchored to stress
    $$[\,\acute{\sigma}\,\sigma\,\sigma\,]\ [\,\sigma\,\acute{\sigma}\,\sigma\,]$$

Preview of results

▶ The main patterns I plan to cover are:

   ▶ The alignment of **L to the stress σ**

   ▶ The **variable alignment of H**

   ▶ The presence of **more complex contours** than LH

   ▶ The effect of **lexical pitch accent** ("lexical tone")

Approaching a new language

▶ How does one analyze the intonation of a language they have no intuitions in?

▶ **Speech type**

   ▶ Avoid fully naturalistic speech (too much variability & reduction, hard to parse)

   ▶ Avoid formal reading (too slow, speakers aren't really "feeling" what they're saying)

   ▶ Best compromise: **reading a brief story**

Approaching a new language

- ▶ How does one analyze the intonation of a language they have no intuitions in?

- ▶ **Scope of study**
  - ▶ IP boundaries often rely on native judgment about turn-taking, speaker confidence
  - ▶ Question-answer pairs often rely on native judgment of given and focused items
  - ▶ Best compromise: **non-phrase-final word-level prosody in storytelling**

Approaching a new language

▶ How does one analyze the intonation of a language they have no intuitions in?

▶ **Data collection**

  ▶ Rely less on diasporic speakers whose intonation may reflect multiple lgs at once

  ▶ Avoid recording people in a lab while surrounded by speakers of other lgs

  ▶ Pay (use Wise, WesternUnion) bilingual native speakers in situ to translate stories for you

  ▶ Have them find (effective-)monolinguals to self-record without fancy gadgets (use WhatsApp)

Approaching a new language

▶ How does one analyze the intonation of a language they have no intuitions in?

▶ **Data analysis**

  ▶ For **crosslinguistic consistency**, examples are transcribed using a single model: **InTraSAL**

    ▶ Intonational Transcription of South Asian lgs

    ▶ A substantial elaboration of B-ToBI (Khan 2008/2014)

  ▶ Autosegmental-Metrical (AM) Theory (Beckman & Pierrehumbert 1986, Ladd 1996…) provides structure, labels

  ▶ *Intonation in Romance* (Frota & Prieto 2015) provides a guide to build an abstract analysis across lgs

Data sources

▶ Data come from **corpus in progress** (Khan 2020-):

  ▶ *North Wind and Sun* fable

    ▶ ***JIPA***   IA: Assamese, Bengali, Hindi, Kalasha, Khowar, Kholosi, Nepali, Punjabi, Saraiki, Sindhi
             Dr: Malayalam, Telugu, Tamil
             TB: Sumi

    ▶ **UCLA Archive**   IA: Hindi, Sindhi, Urdu
                     Dr: Kannada, Telugu, Tamil

  ▶ *North Wind and Sun + Frog, Where Are You?* (Mayer 1969)

    ▶ **Reed College LoL**   IA: Hindi, Nepali, Urdu
                     Dr: Kannada, Malayalam

    ▶ **Socially-distant self-recording** — 50 languages (IA, Ir, Dr, TB)
            Most recordings both made and sent using WhatsApp

PANDEMIC-FRIENDLY!

Data sources

- Lgs in the corpus:
  - ● IA: 41 lgs
  - ● Ir: 3 lgs
  - ■ Dr: 5 lgs
  - ▲ TB: 2 lgs
  - isol.: 1 lg (Burushaski)
  - **55 lgs total**
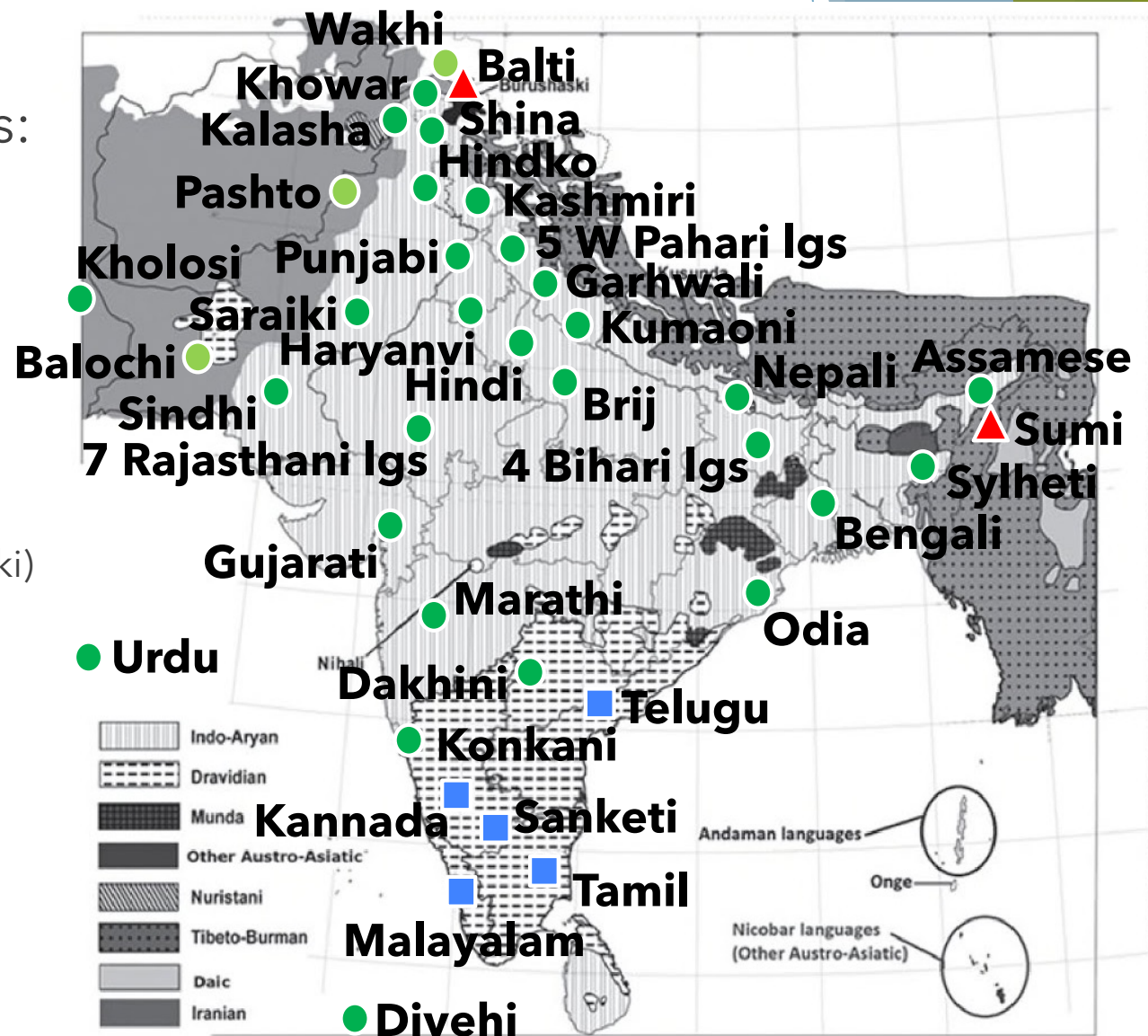
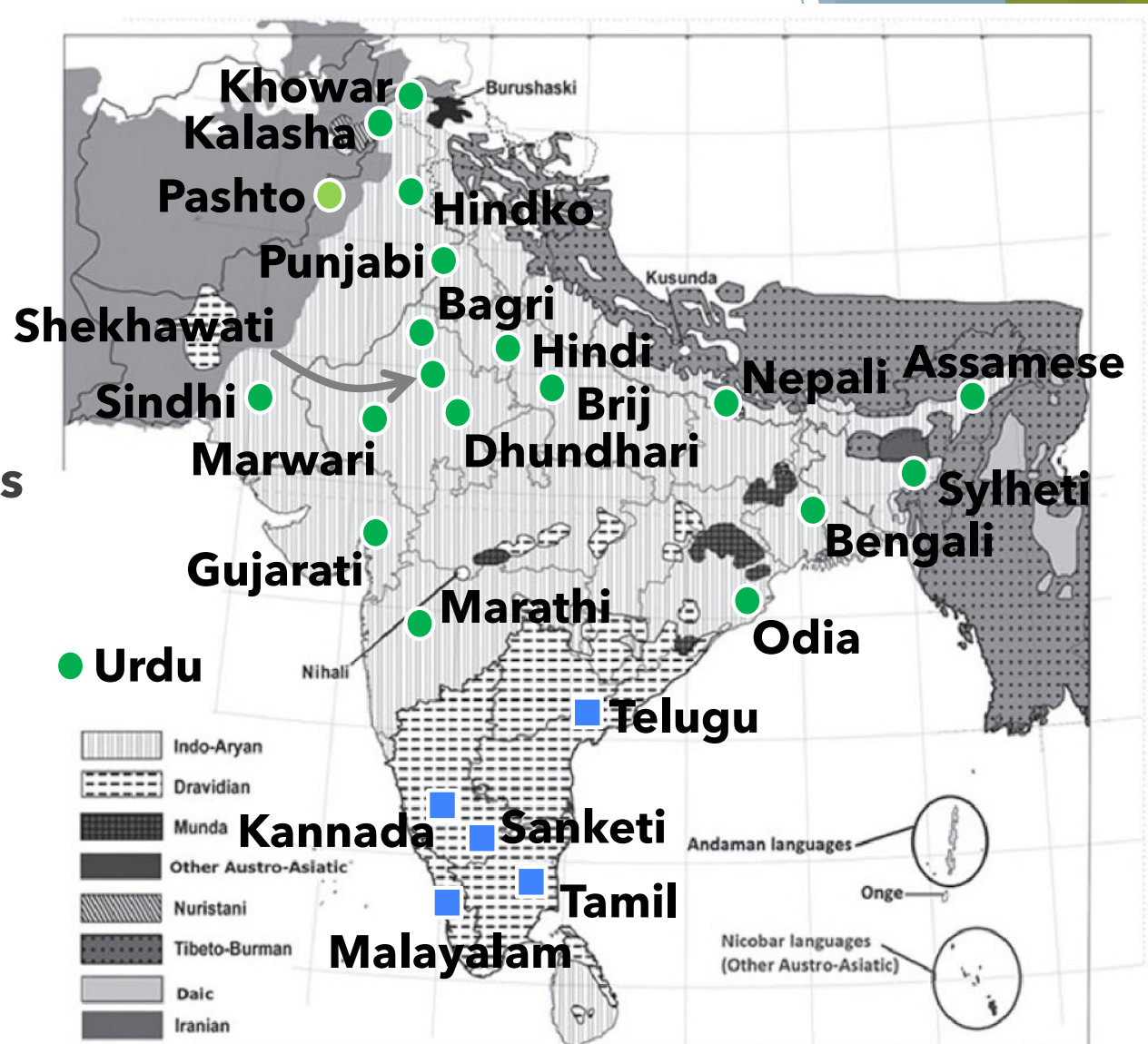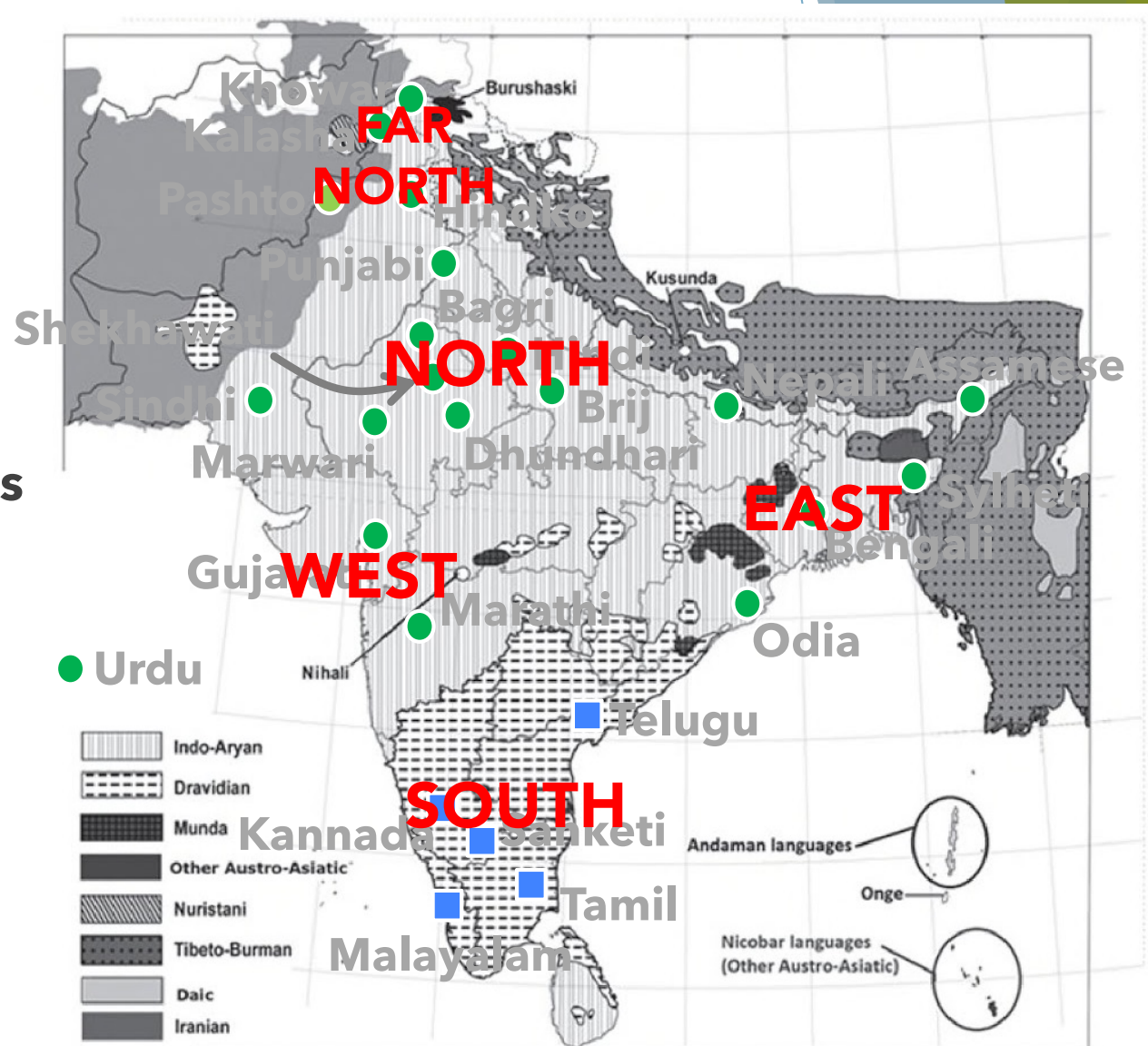South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Data sources

▶ Lgs studied here:

- 🟢 IA: 19 lgs
- 🟢 Ir: 1 lg (Pashto)
- 🟦 Dr: 5 lgs

25 lgs of 3+ spkrs in the corpus

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Data sources

▶ Lgs studied here:
- ● IA: 19 lgs
- ● Ir: 1 lg (Pashto)
- ■ Dr: 5 lgs

25 lgs of 3+ spkrs in the corpus

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Data transcription: InTraSAL

▶ What does InTraSAL look like?

  ▶ Here's a **spectrogram** and **pitch track** of Sindhi

 ▶ First the transcriber just listens to the **pitch movements** and their **timing**

300 Hz

100 Hz
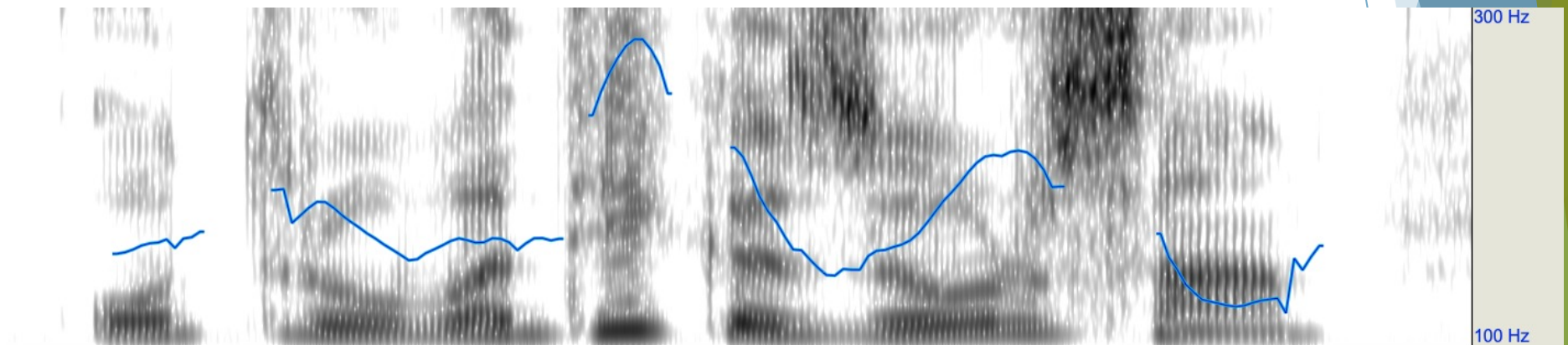
▶ '…with that much more force…' <NWS-Sindhi-J>

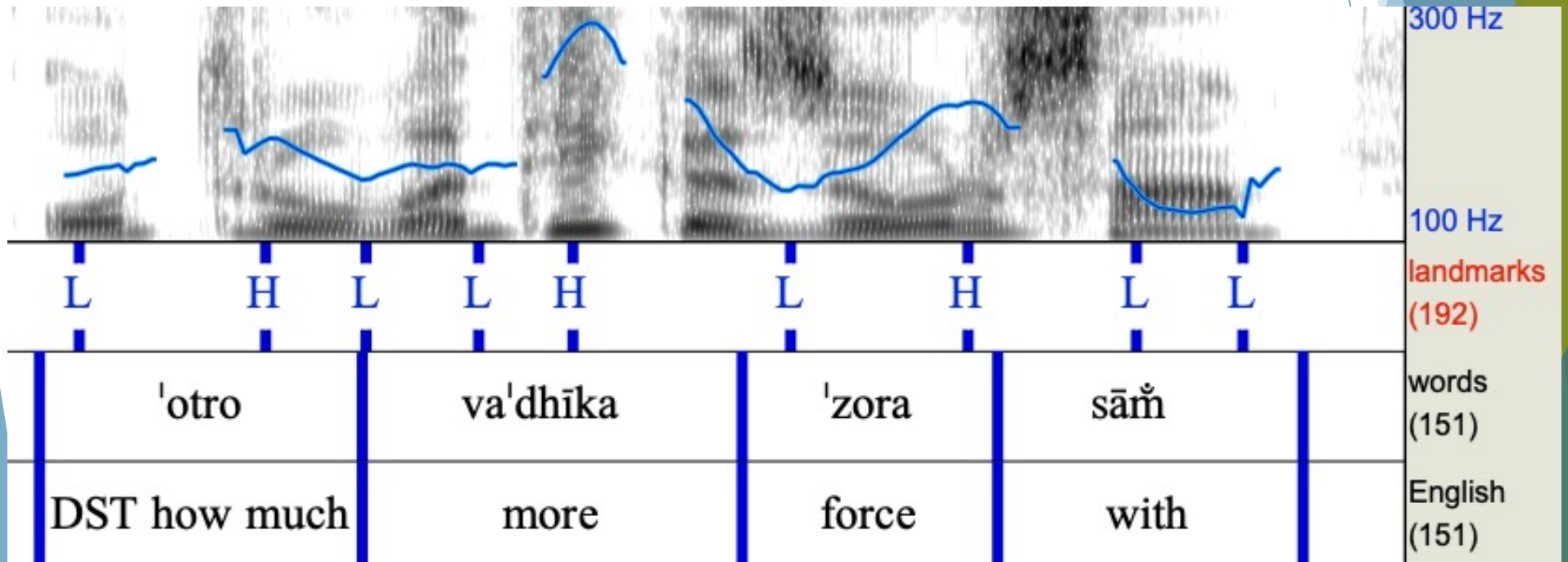# Data transcription: InTraSAL

- Label the **Landmarks Tier**: label for each target
  - Just label; don't think too hard about "why"
  - But listen carefully for whether a visual "target" is audible vs. a tracking error / microprosody
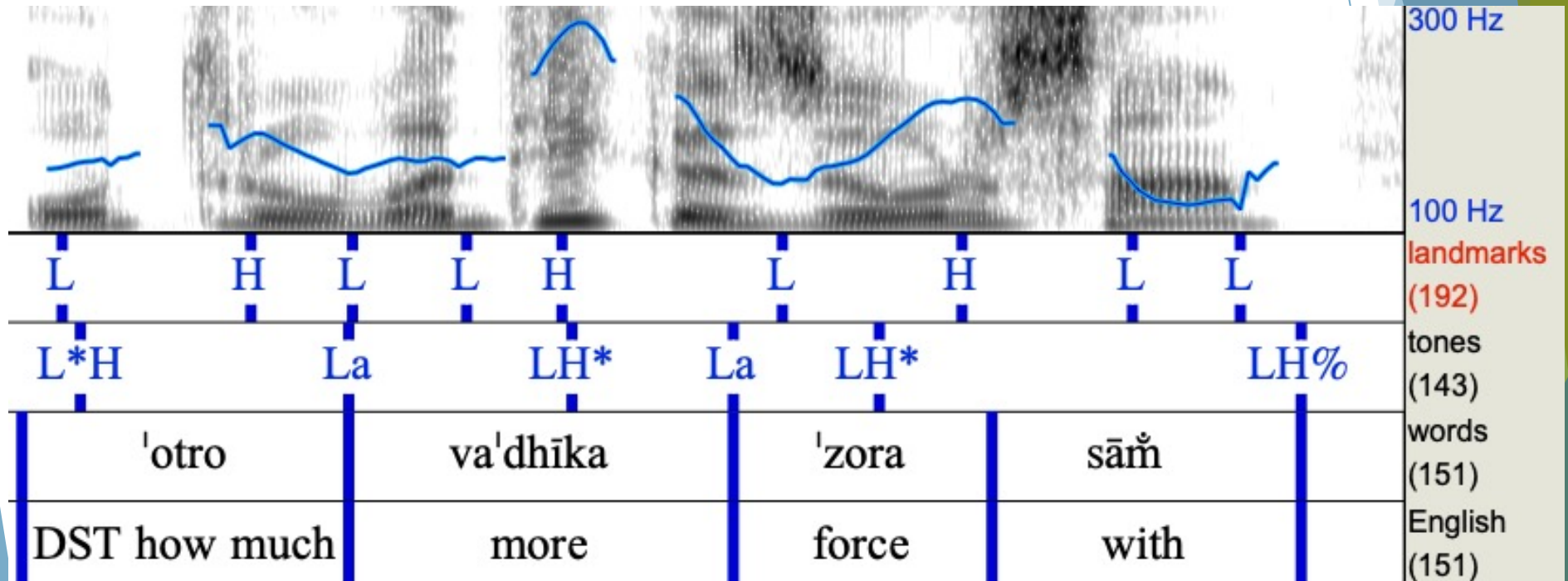 - Mitigates unconscious shoehorning of a new lg into tunes we find familiar
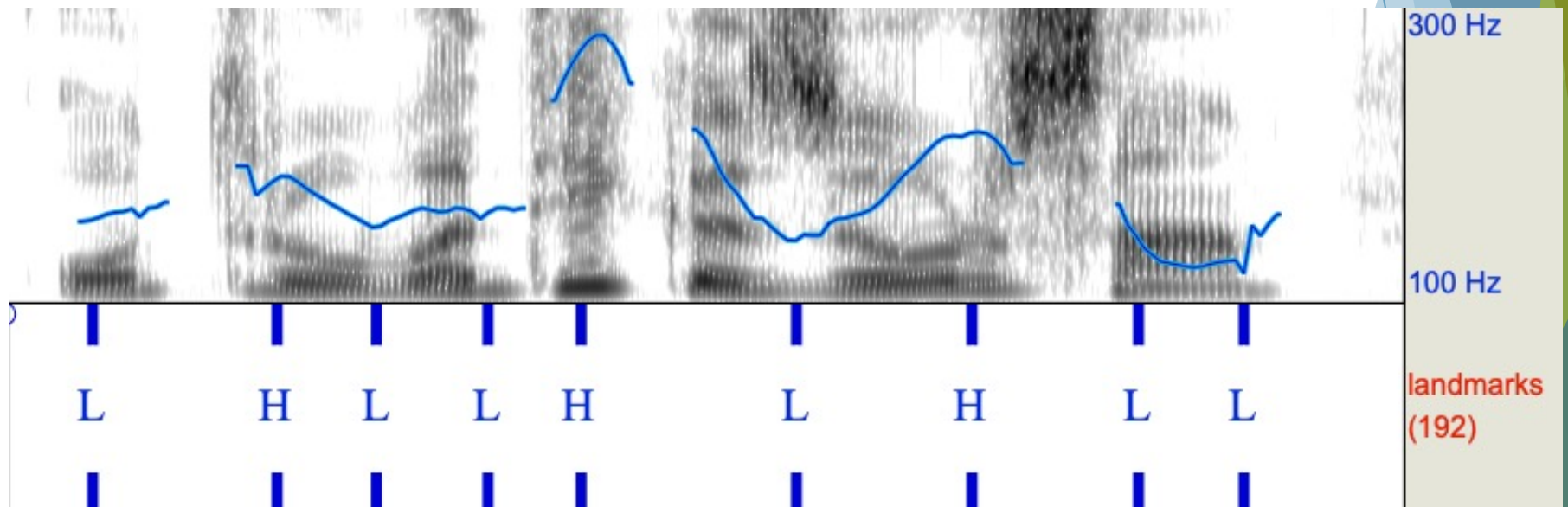
Data transcription: InTraSAL

▶ Then the Words and English tiers

    ▶ Words tier: consonant+vowel transcription

    ▶ Split orthographic words into stems if possible

    ▶ English tier: technical translation

Data transcription: InTraSAL

▶ Last, the Tones tier

  ▶ This is when you're ready to interpret the Landmarks based on your (evolving) model

  ▶ Use InTraSAL labels, from AM Theory *(next slide)*

L	H L L H	L H	L L
L*H	La	LH* La LH*	LH%
'otro	va'dhīka	'zora	sām̌
DST how much	more	force	with

300 Hz

100 Hz

landmarks (192)

tones (143)

words (151)

English (151)

Data transcription: InTraSAL

▶ In AM theory (Pierrehumbert 1980, see Ladd 1996), intonational tones are classified in two ways

▶ **Target**: relative pitch

   ▶ Low (L)

   ▶ High (H)

   ▶ Combinations of the two, e.g. LH, HL

▶ **Type**: structure it marks / is associated to

   ▶ **Pitch accent** (T*): marks **accented σ** *(see next slide)*

   ▶ **Boundary tone** (T%, T-, Ta, T_I, T_P): marks **phrase edge**

Data transcription: InTraSAL

- Accent: σ with **privileged phonological status**
  - Resists neutralization (host more contrasts)
  - Resists reduction
  - Attracts intonational tones (**pitch accents**, T*)
- Can be **marked phonetically** with:
  - **Stress** ("stress accent")
  - **Pitch contour** ("lexical pitch accent")
  - **Both stress and pitch contour**
  - **Neither** ("unmarked accent")

Patterns to cover

▶ The main patterns I plan to cover are:

   ▶ The alignment of **L to the stressed σ**

   ▶ The **variable alignment of H**

   ▶ The presence of **more complex contours** than LH

   ▶ The effect of **lexical pitch accent** ("lexical tone")

L tone alignment

- ▶ Let's first describe the initial **L** of the rise

- ▶ Some call it a boundary tone, aiming for the left edge of a phrase (aL) (Féry 2010)

- ▶ Others call it a pitch accent (L*), aiming for a stressed σ (Hayes & Lahiri 1991, Genzel 2007, Patil et al. 2008, Khan 2014, Féry & Fanselow 2020…)
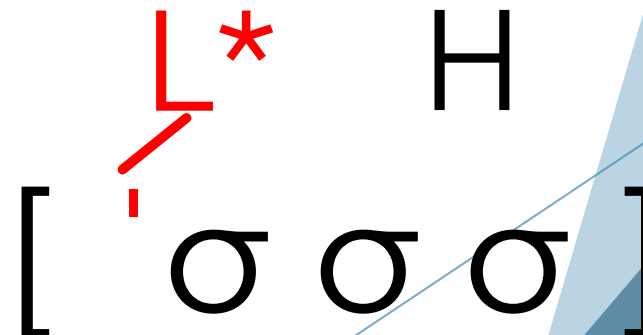
aL   H   L*   H

[σ σ σ]   [ˈσ σ σ]

L tone alignment

▶ Stress assignment
(Hock 2016:398-400):

- 🟢 initial
- 🟣 initial+weight
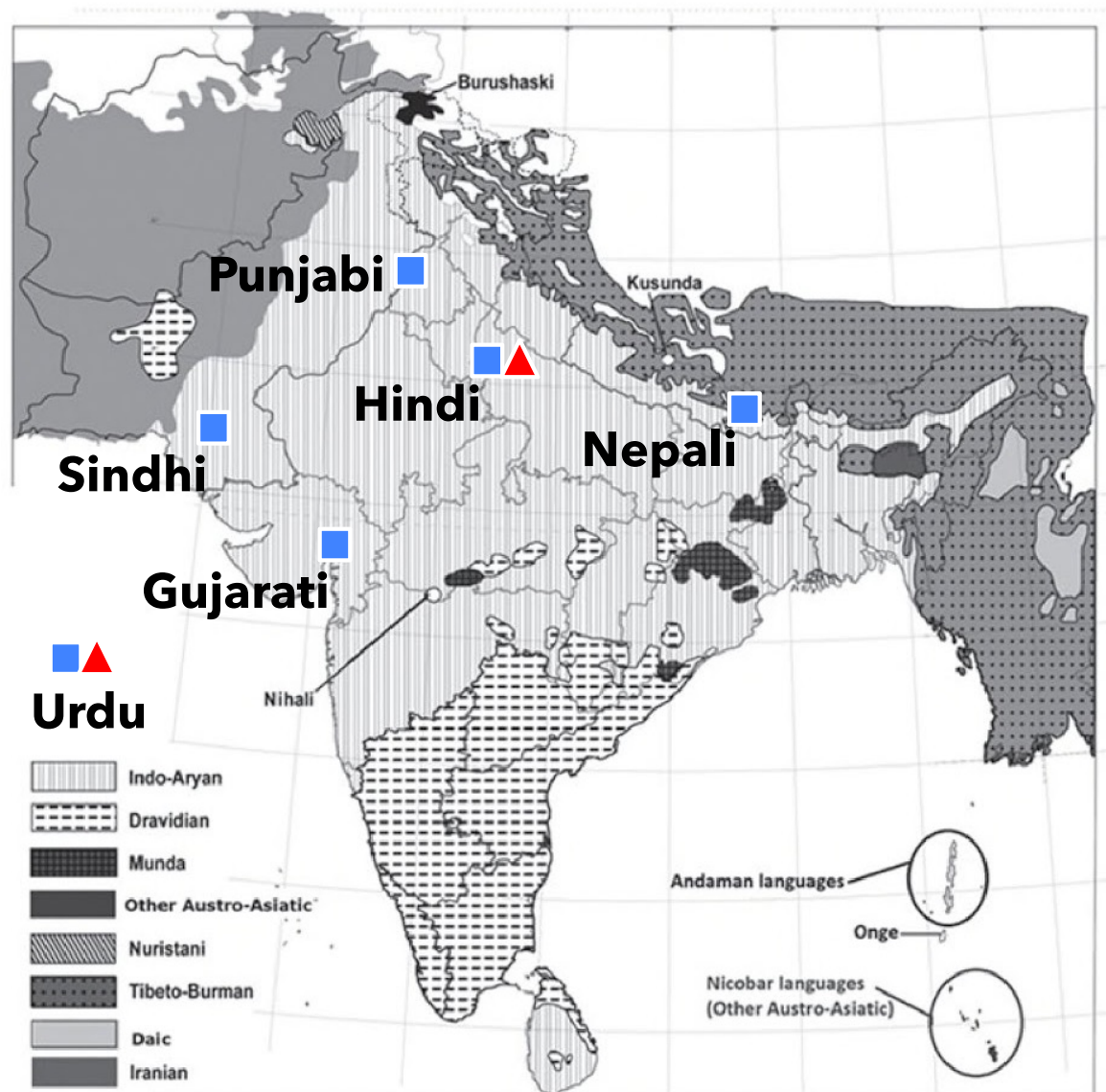- 🟦 penult+weight
- 🔺 contrastive
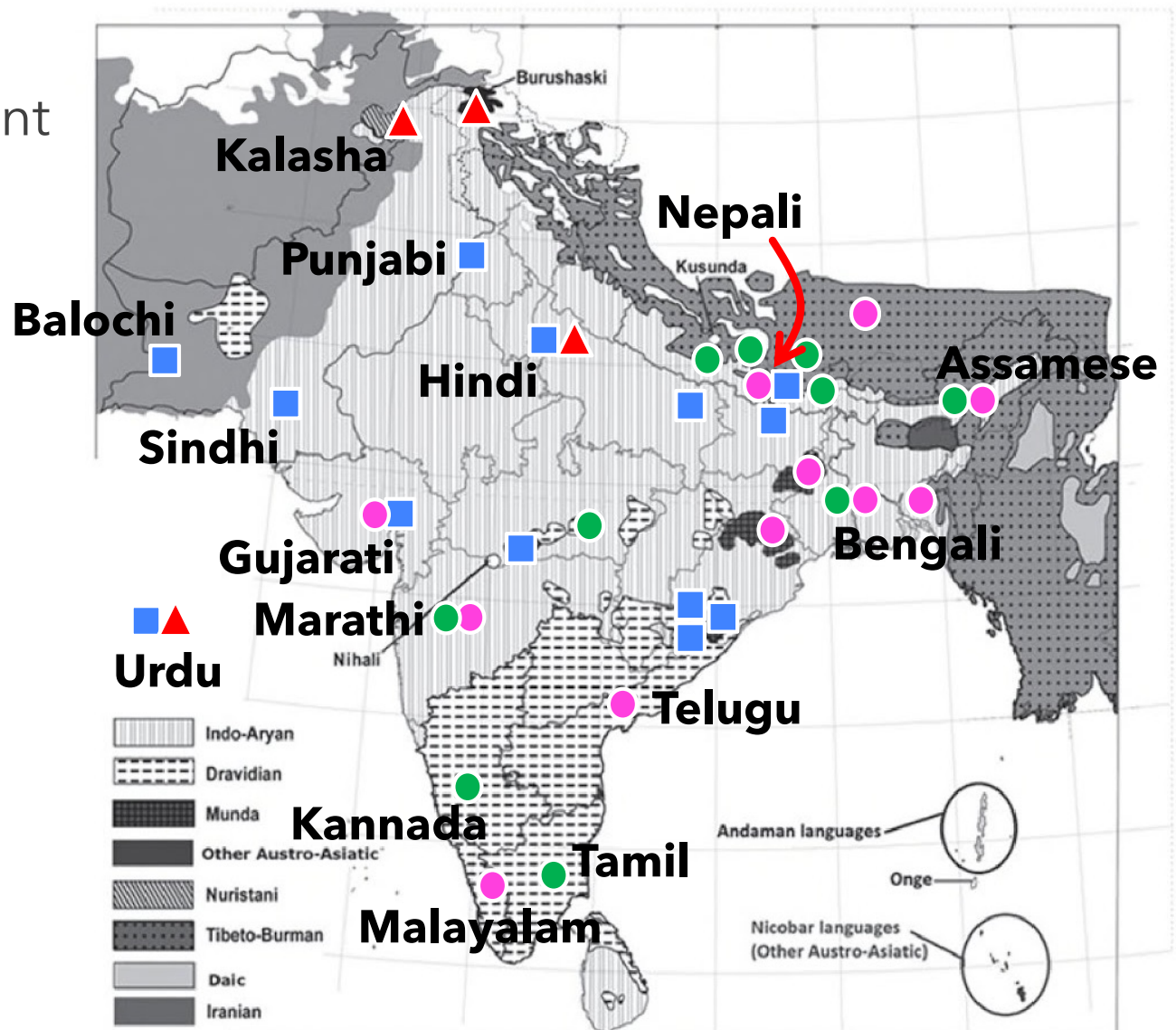
South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

L tone alignment

▶ Stress assignment

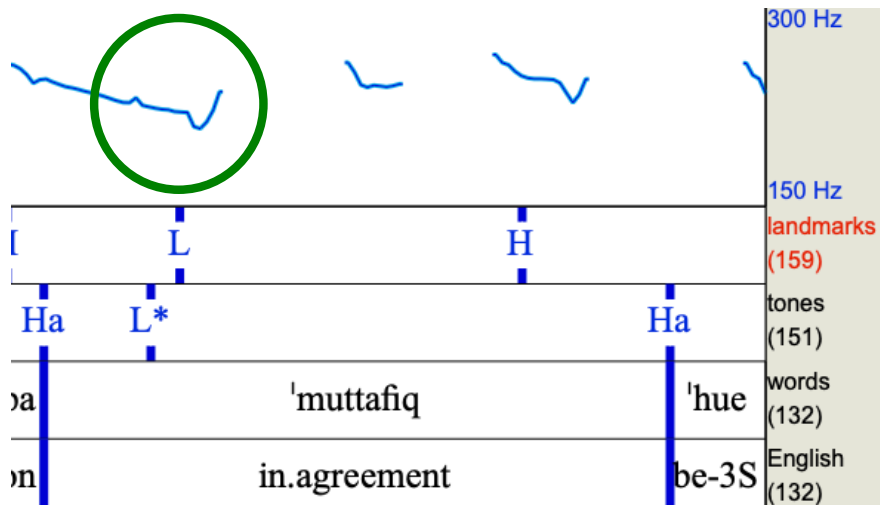(Hock 2016:398-400):

■ penult+weight

▲ contrastive

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

L tone alignment [ˈσ σ σ] [σ ˈσ σ]

▶ L* in **Urdu**: on **heavy σ**, not strictly 1ˢᵗ σ/left edge

▶ '…in agreement…'

▶ '…the North Wind…'

<NWS-Urdu-R01b>

L tone alignment

▶ L in northern SALs aligns with stressed σ, even when not on the left edge: L*

▶ This stress pattern is determined by weight and/or lexical marking

L* H

[σ σ σ]

ˈmut ta fiq

L* H

[σ σ σ]

ši ˈmā lī

L tone alignment

▶ Stress assignment
(Hock 2016:398-400):

● initial

● initial+weight

South Asian language families (map produced by Suresh Kolichala, 2015)
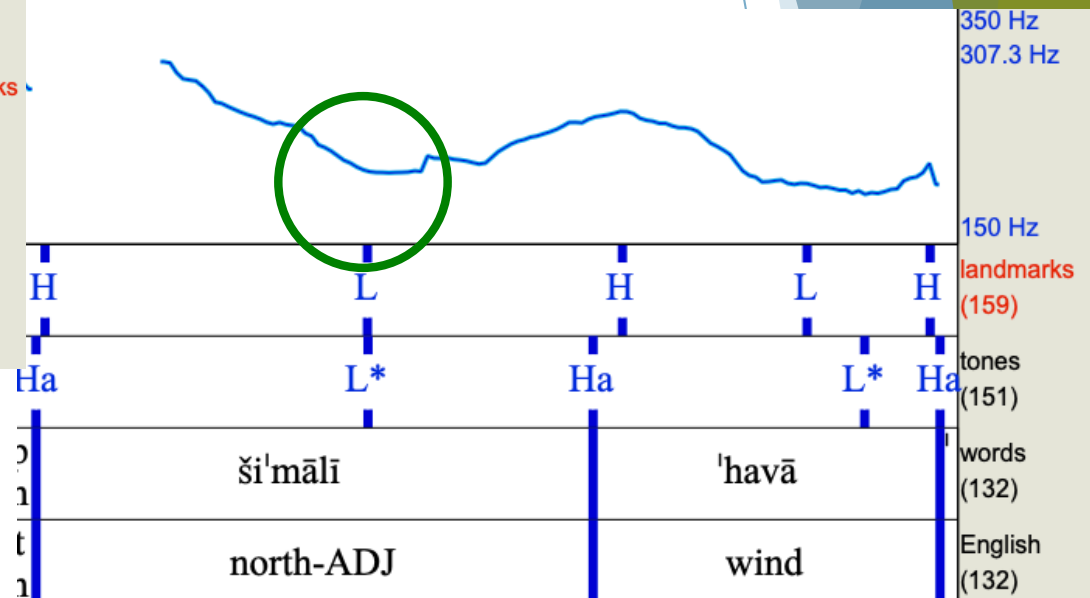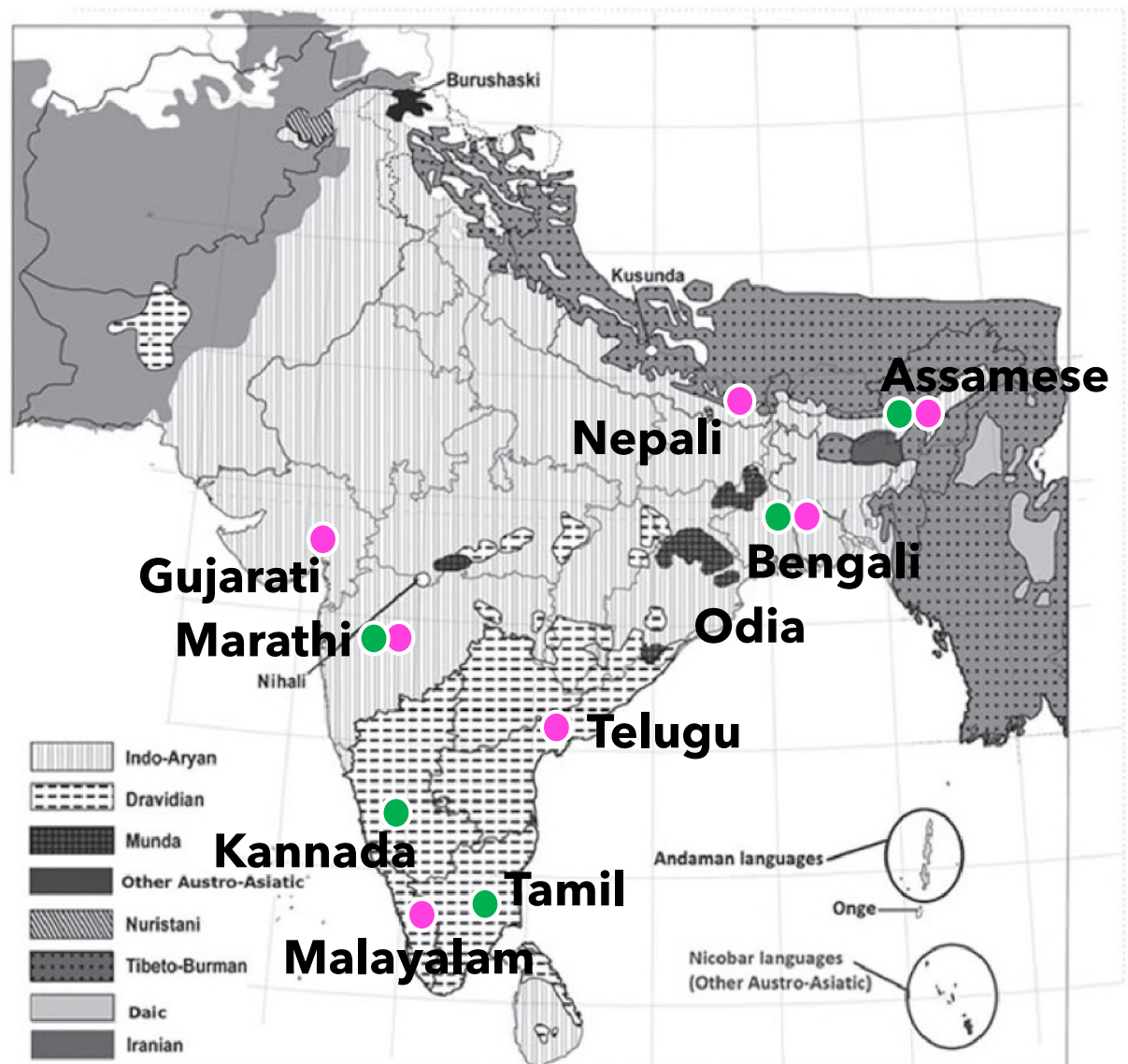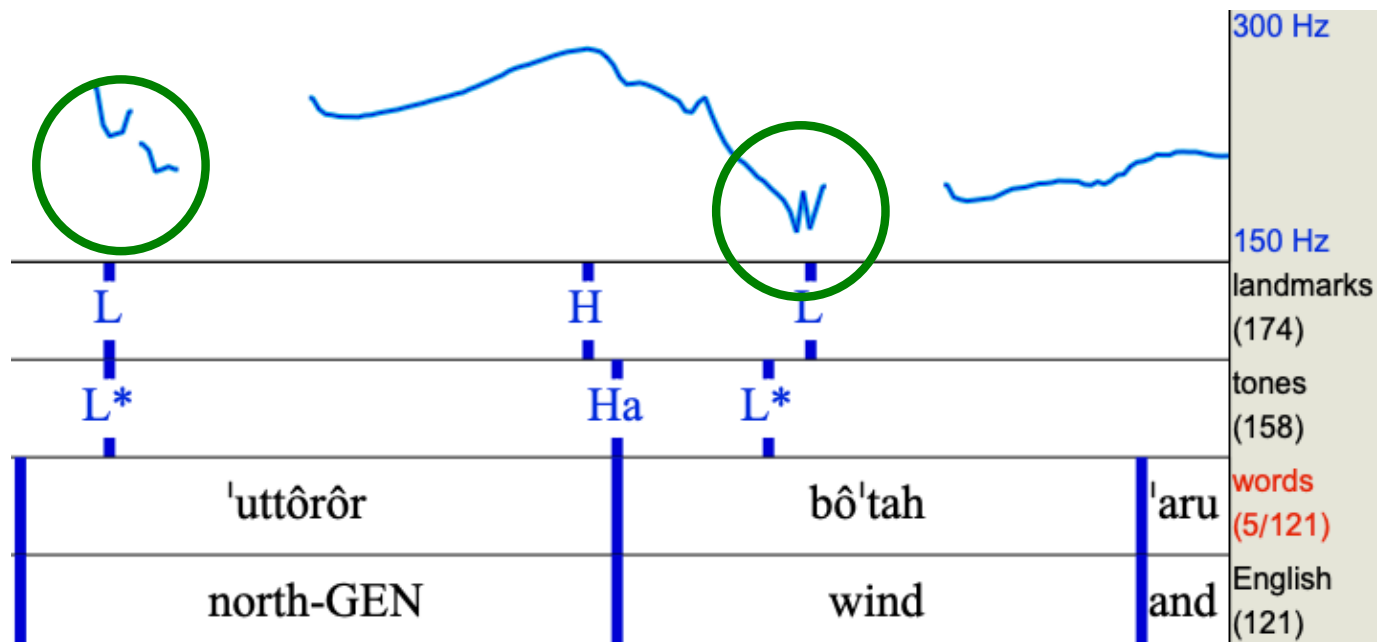Base map taken from Hock (2016, p.7)

L tone alignment [ˈσ σ σ] [σ ˈσ σ]

▶ L* in **Assamese**: ignores weight, always initial

▶ 'The North Wind…' <NWS-Assamese-J>

L tone alignment [ˈσ σ σ] [σ ˈσ σ]

▶ L* in **Malayalam**: ignores weight, always initial

	L		H		H	L
Ha	L*		Ha	L*		Ha	L
	bhūmiyil		oru		saṉyāsi		mēl
	earth-LOC		one		monk		on

150 Hz
100 Hz
landmarks (164)
tones (114)
words (19/120)
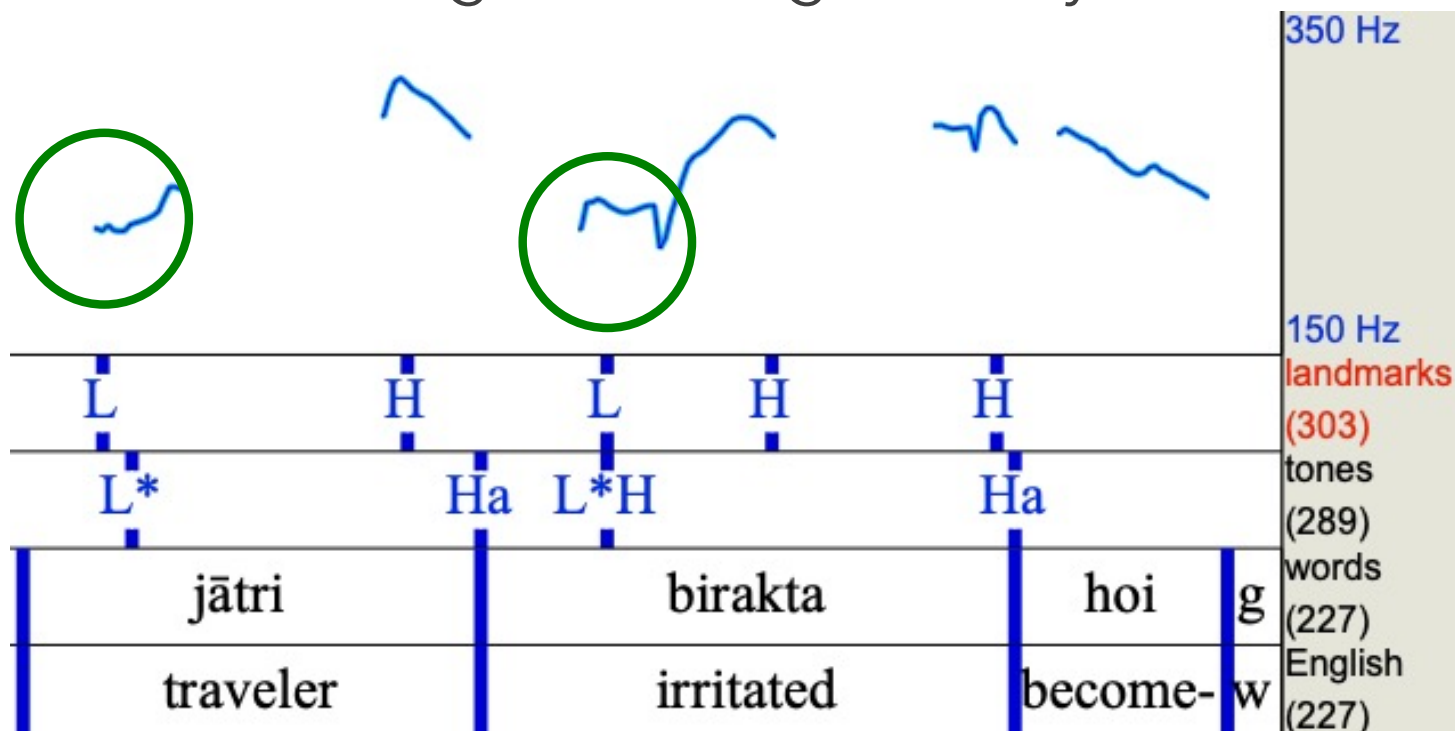English (120)
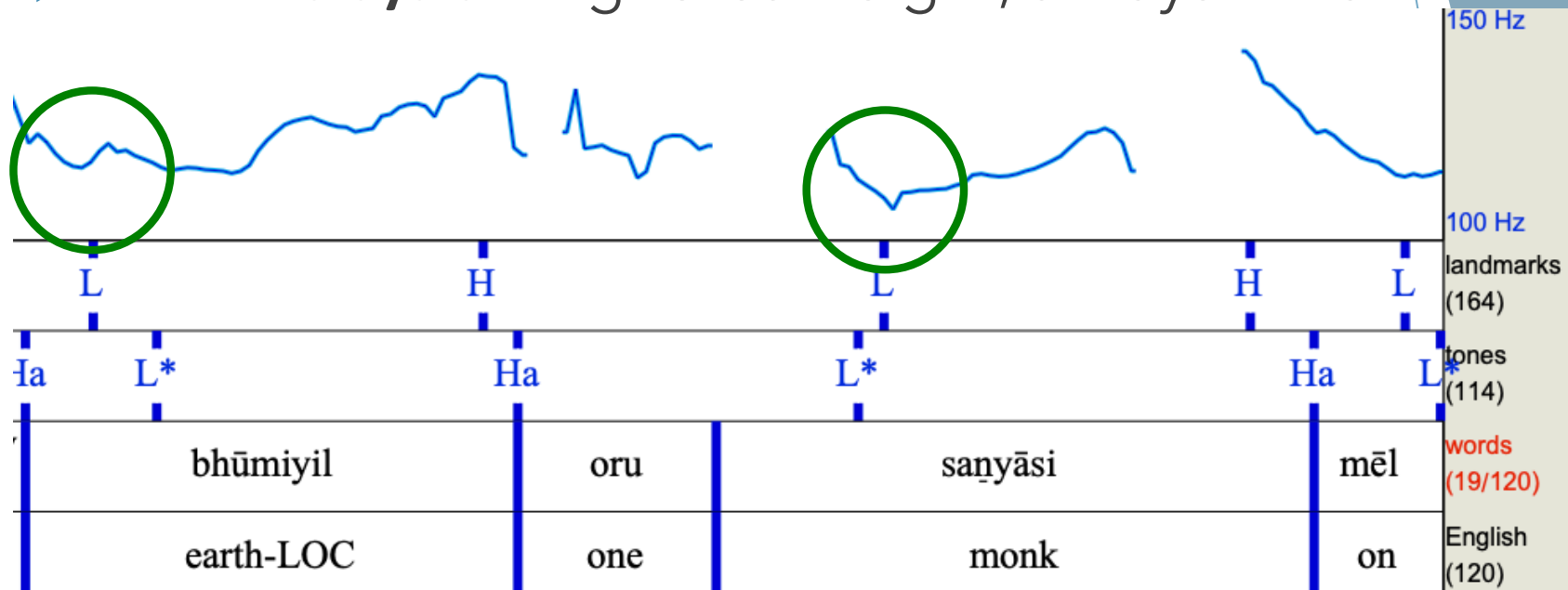
▶ 'On earth, a monk…' <NWS-Malayalam-R01c>

L tone alignment [ˈσ σ σ] [σ ˈσ σ]

▶ L* in **Odia**: ignores weight, always initial

▶ 'Having gotten irritated, the traveler…'
<NWS-Odia-D01b>

L tone alignment [ˈσ σ σ] [σˈσ σ]

▶ In Dr and most IA lgs, it seems the L is strictly on the 1ˢᵗ σ

▶ Ignores σ weight, no word-specific behavior

▶ Two interpretations: aL or L*?

L	H

[σ σ σ]
bhū mi yil

L	H

[σ σ σ]
sa nyā si

L tone alignment [ˈσ σ σ] [σˈσ σ]

▶ **aL** interpretation: L on word's left edge

▶ Would mean that L is attracted to left edges, and ignores stress

▶ This would mean that stress has no role in the intonation of such lgs

aL   H

[ˈσ σ σ]

ˈbhū mi yil

aL   H

[σˈσ σ]

sa ˈnyā si

L tone alignment [ˈσ σ σ]

- **L*** interpretation: L on stress

- Would mean that in these lgs, stress is weight-insensitive and always initial

- Indeed, evidence from Bengali & Tamil data show the initial σ is phonologically "special"

L*　H

[ˈσ σ σ]

ˈbhū mi yil

L*　H

[ˈσ σ σ]

ˈsa nyā si

L tone alignment [ˈσ σ σ]

- ▶ Considering all this, we can adopt the L* interpretation, with reservations (Keane 2014)
  - ▶ Allows for cross-SAL transparency
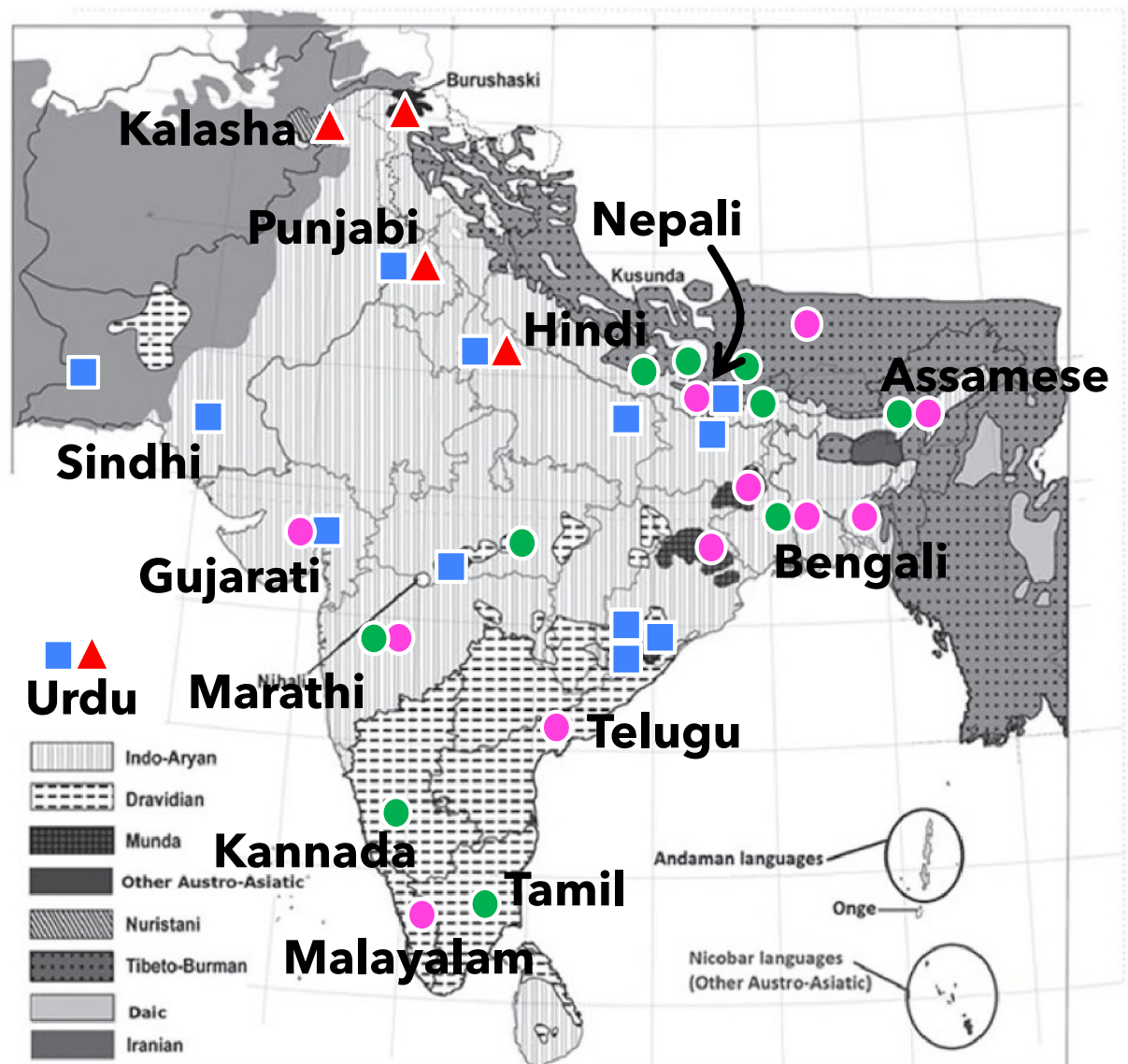 - ▶ Takes into account phonological status

L* H

[ˈσ σ σ]

ˈbhū mi yil

L* H

[ˈσ σ σ]

ˈsa nyā si

L tone alignment

- Stress assignment
  (Hock 2016:398-400):
  - 🟢 initial
  - 🔴 initial+weight
  - 🟦 penult+weight
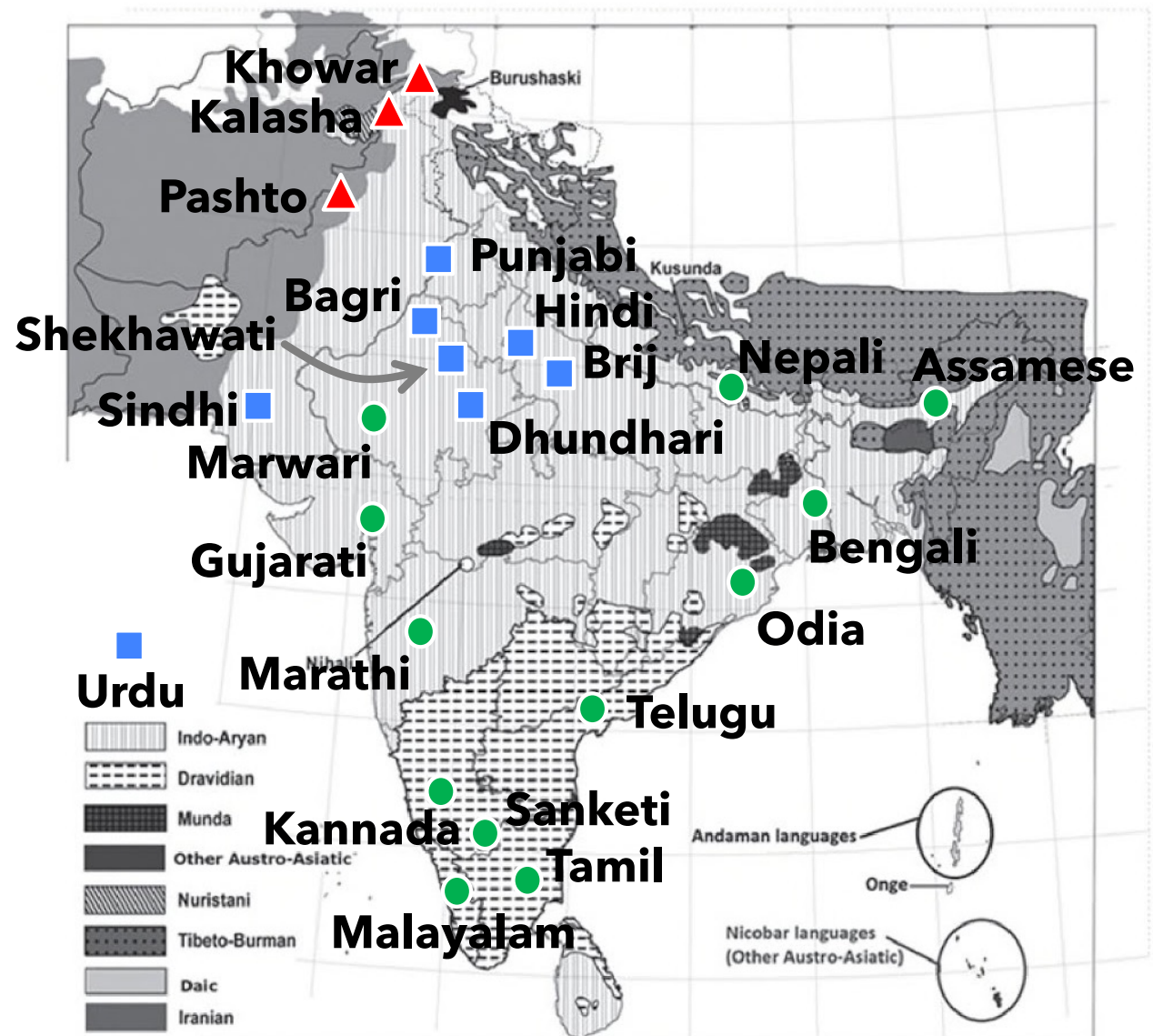 - 🔺 contrastive

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

L tone alignment

- L* location:
  - ● initial
  - ■ penult+weight
  - ▲ contrastive

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Patterns to cover

▶ The main patterns I plan to cover are:

  ▶ The alignment of **L to the stressed σ**

  ▶ The **variable alignment of H**

  ▶ The presence of **more complex contours** than LH

  ▶ The effect of **lexical pitch accent** ("lexical tone")

H tone alignment

▶ Most accounts of SALs describe strict **right-alignment** of the rise's H tone: **Ha** (Genzel 2007)

▶ Ha location is useful for describing:

  ▶ Domains of various morphophonological processes (Hayes & Lahiri 1991, Twaha & Mahanta 2016)

  ▶ Disambiguation of syntactic structure (Lahiri & Fitzpatrick-Cole 1999, Féry & Fanselow 2020)

H tone alignment [ˈσ σ σ]

▶ Two examples of **Ha** in **Assamese**

	'mani	'loboloi	'baiddhyô	hol	zê
H	L		H	L		H		L H
LHa L*			Ha	L*		Ha	L* H%
ɔo	accept-P	take-INF-D	obliged	become	C

275 Hz
179.1 Hz
150 Hz
landmarks (174)
tones (158)
words (121)
English (121)
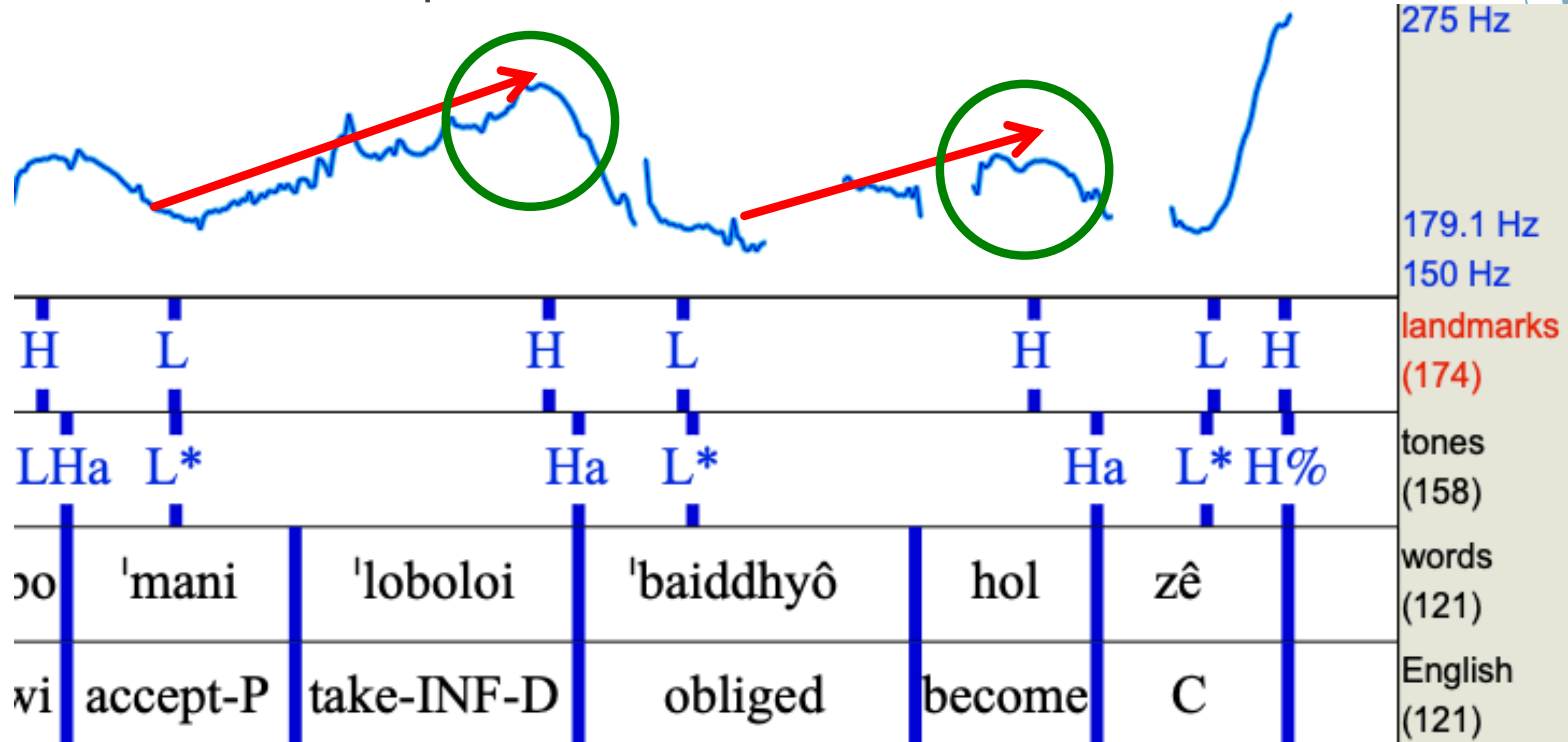
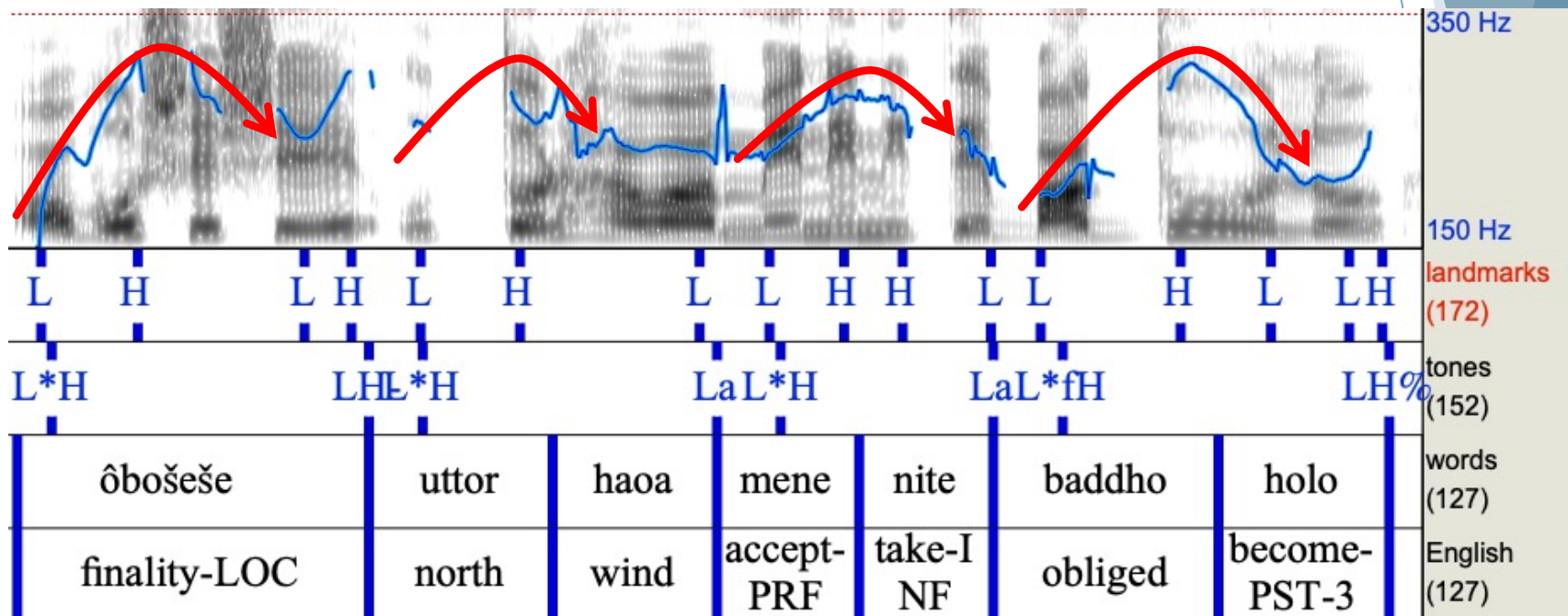▶ '…was obliged to accept…' <NWS-Assamese-J>

H tone alignment

- ▶ But occasionally, researchers mention a mysteriously **early appearance of this H**:
- ▶ **H immediately after the σ with L***, regardless of distance to word edge, in:
  - ▶ Bengali (Khan 2014)
  - ▶ Tamil (Keane 2014)
- ▶ Complex pitch accent **L*H**?

H tone alignment

▶ This suggests "the" LH rise in a conflation of at least two different structures

  ▶ **L*...Ha:** rise peaks at AP edge   [ˈσ σ σ]

  ▶ **L*H:** rise peaks just after L*   [ˈσ σ σ]

▶ Most SAL research looks mostly/only at words of two syllables, which are highly frequent

▶ This means L*...Ha and L*H could not be adequately distinguished

▶ We need to look at **longer stems**

H tone alignment [ˈσ σ σ]

▶ Four examples of **L*H** in **Bengali**

▶ Peak at end of 2ⁿᵈ σ, can persist into 3ʳᵈ σ

L	H		L H	L	H		L	L H H	L L		H	L	LH

L*H		LHL*H		LaL*H	LaL*fH		LH%

ôbošeše	uttor	haoa	mene	nite	baddho	holo
finality-LOC	north	wind	accept-PRF	take-I NF	obliged	become-PST-3

landmarks (172)
tones (152)
words (127)
English (127)

350 Hz
150 Hz
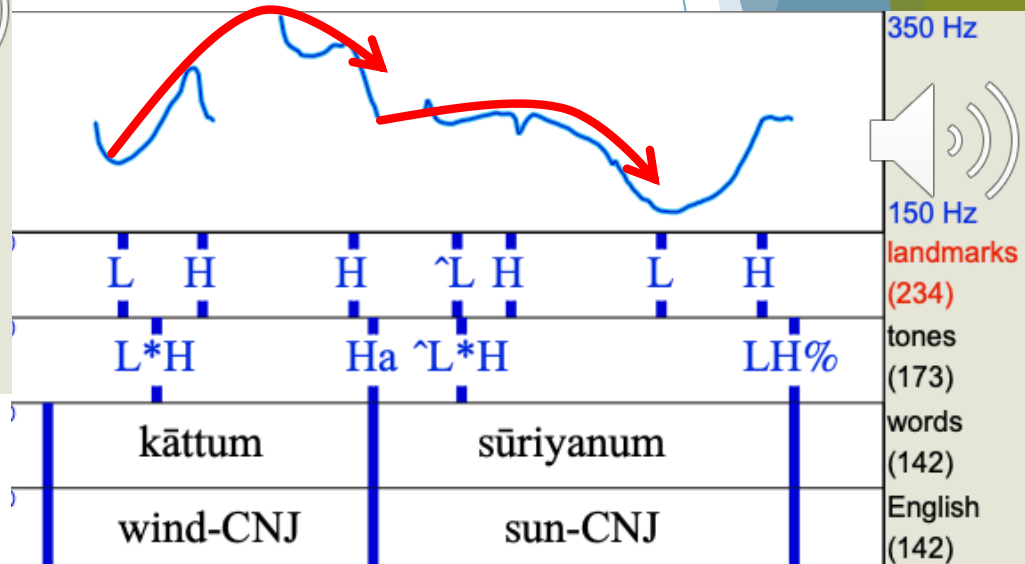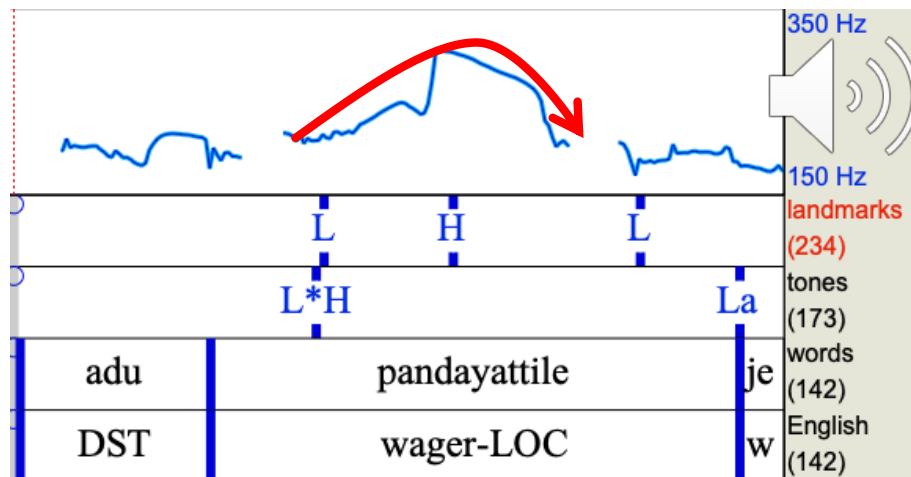
▶ 'Finally, the North Wind was obliged to admit…'
<NWS-Bengali-J>

H tone alignment [ˈσ σ σ]

- Examples of **L*H** in **Tamil**
- Peak reached on 1st or 2nd σ (based on weight)

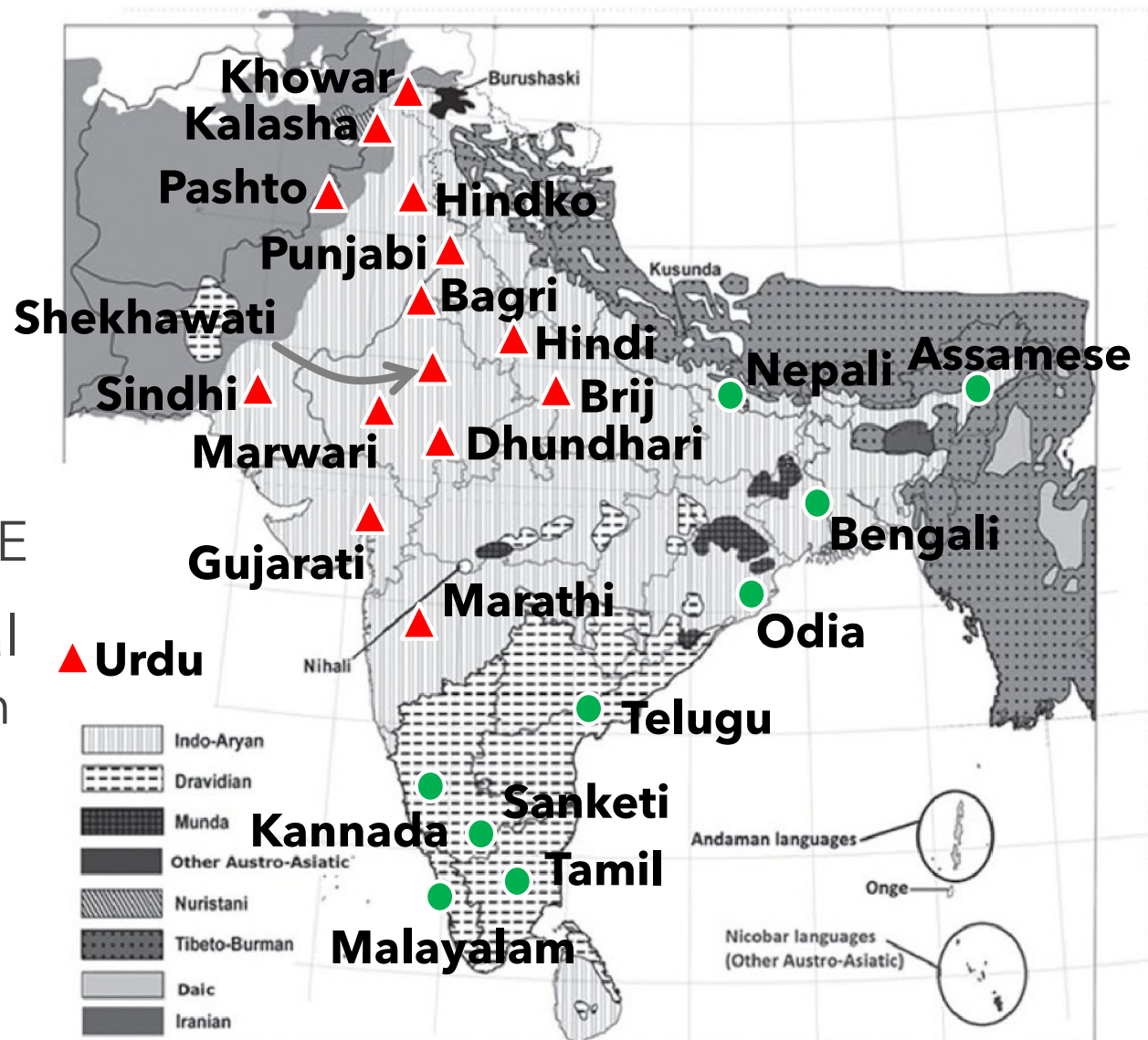

- '...in that wager...'   '...the Wind and Sun...'

<NWS-Tamil-J>

Complex pitch accents

- SALs with L*H clearly distinct
  - ● L*H well-attested
  - ▲ L*H not attested

- Restricted to S & E

- Overlaps w/ initial stress distribution (S, E, W)

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Patterns to cover

▶ The main patterns I plan to cover are:

  ▶ The alignment of **L to the stressed σ**

  ▶ The **variable alignment of H**

  ▶ The presence of **more complex contours** than LH

  ▶ The effect of **lexical pitch accent** ("lexical tone")

More complex contours

- So far, we've assumed that even with some variation, SALs use LH contours on all words

- But exploring this corpus reveals more complex rises, involving medial tones:
    - LLH (late rise):   L*…**L**Ha   [ˈσ σ σ]
    - LHH (early rise):   L***H**…Ha   [ˈσ σ σ]

- We also see contours that involve falls:
    - LHL (rise-fall):   L*H…**L**a   [ˈσ σ σ]
    - HLH (fall-rise):   **H**L*…Ha   [σ ˈσ σ]
    - LHLH (double rise):   L***H**…**L**Ha   [ˈσ σ σ]

More complex contours

▶ So far, we've assumed that even with some variation, SALs use LH contours on all words

▶ But exploring this corpus reveals more complex rises, involving medial tones:
  ▶ LLH (late rise):   L*…**L**Ha $[\,'\sigma\,\sigma\,\sigma\,]$
  ▶ LHH (early rise):   L***H**…Ha $[\,'\sigma\,\sigma\,\sigma\,]$

▶ We also see contours that involve falls:
  ▶ LHL (rise-fall):   L*H…**L**a $[\,'\sigma\,\sigma\,\sigma\,]$
  ▶ HLH (fall-rise):   **H**L*…Ha $[\,\sigma\,'\sigma\,\sigma\,]$
  ▶ LHLH (double rise):   L***H**…**L**Ha $[\,'\sigma\,\sigma\,\sigma\,]$

Late rises: L*…LHa [ˈσ σ σ]

▶ Two examples of **L*…LHa** in **Nepali***
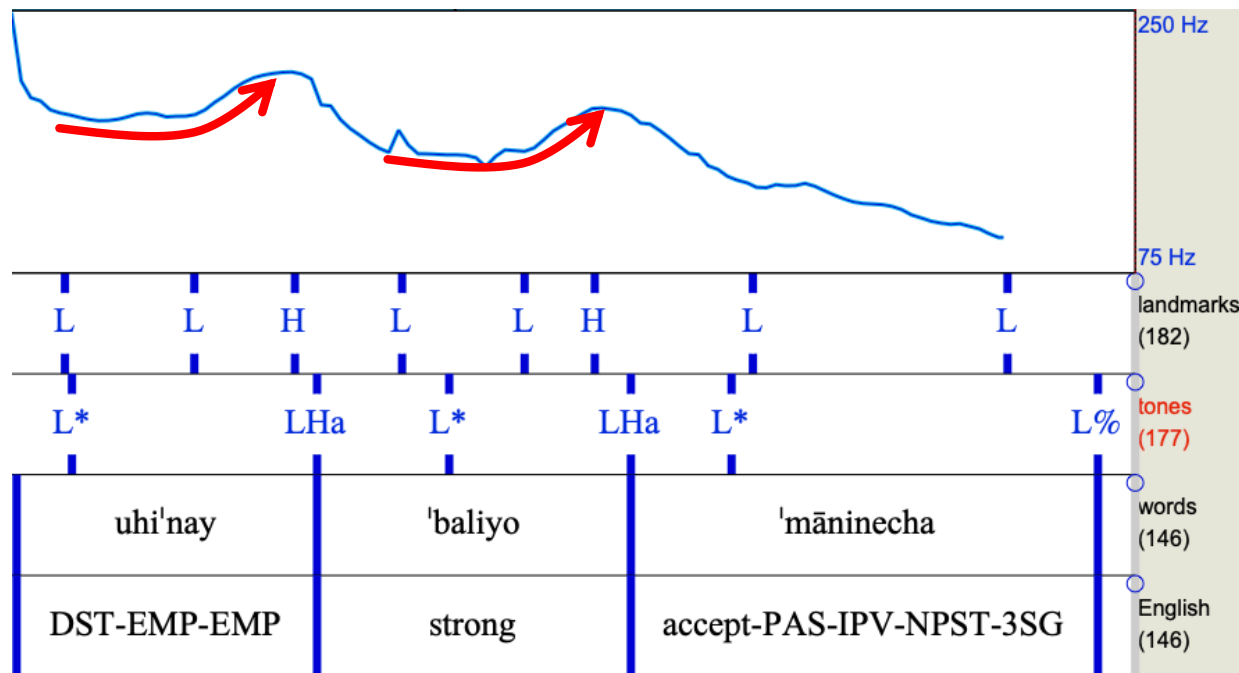
▶ 'That one will be accepted as the stronger.'
<NWS-Nepali-J>

stress marks reflect the weight-sensitive stress pattern in Riccardi (2003), but this does not reliably reflect L alignment in the current study

Late rises: L*…LHa [ˈσ σ σ]

- ▶ Three examples of **L*…LHa** in **Assamese***
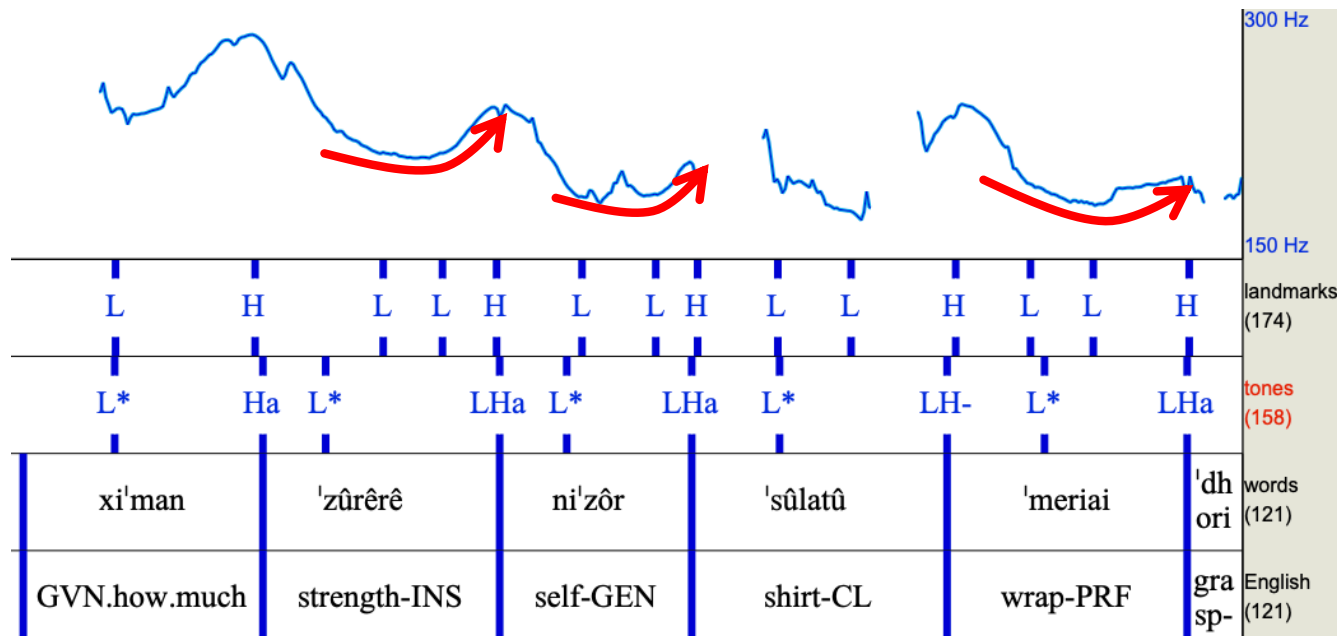
- ▶ 'That much more tightly did (he hold) his shirt.'
  <NWS-Assamese-J>

stress marks reflect the weight-sensitive stress pattern from Mahanta (2012), but this does not reliably reflect L alignment in the current study

Late rises: L*...LHa [ˈσ σ σ]

▶ Two examples of **L*...LHa** in **Malayalam**

								325 Hz
								125 Hz
L	L	H	L	L	H	L	H	landmarks (169)
L*		LHa	L*		LHa	L*	Ha	tones (149)
mēlŭ			podappŭ			cutti	naḍ	words (110)
on			cloak-NMZ			wrap-PST	wal	English (110)
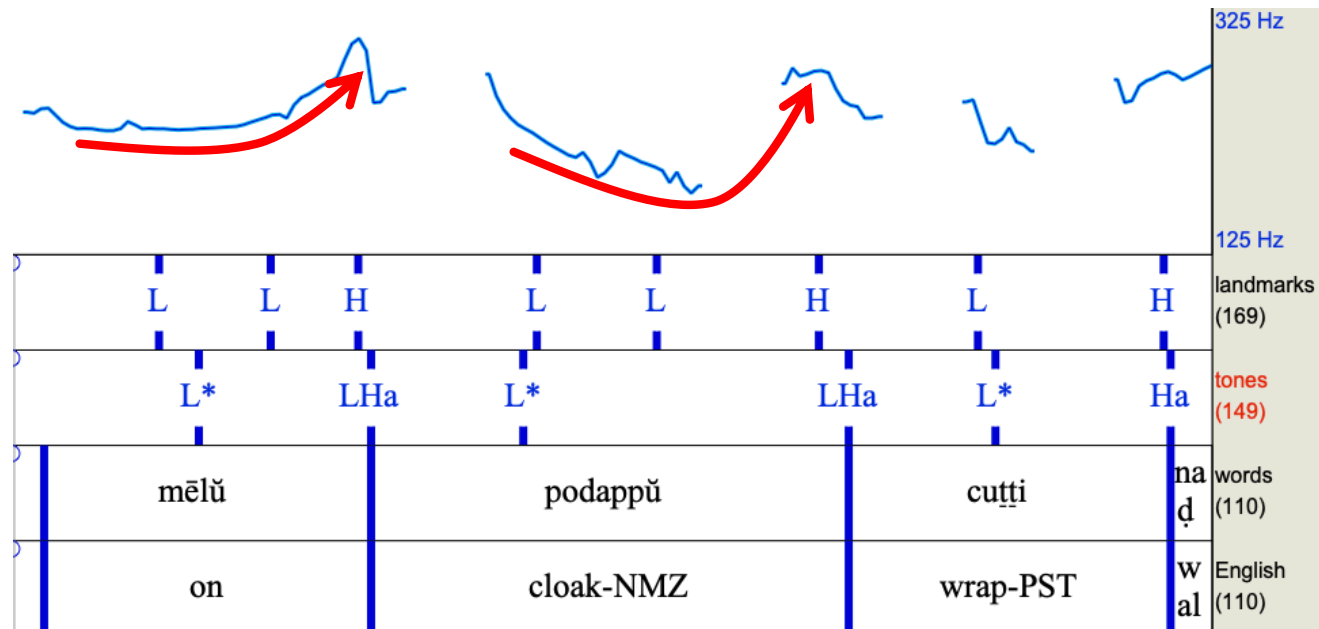
▶ '...wrapped in a cloak on (him)...'
<NWS-Malayalam-J>

Late rises: L*…LHa ['σ σ σ]

▶ This late rise (L*…LHa) might just be a variant of the direct rise (L*…Ha)

▶ Variation within speaker, within a sentence

▶ "Free" variation, w/ lg-specific preferences?
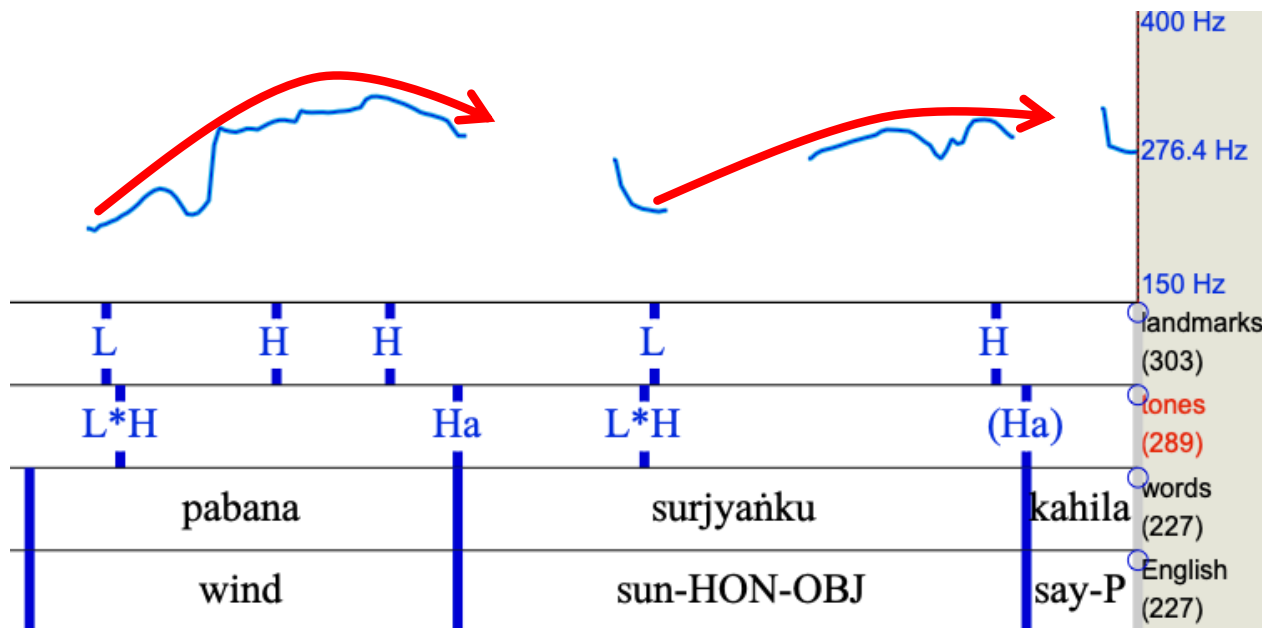
▶ Fortunately, for those using the H as a sign of phrasing: **H is still at the end of the unit**

Early rises: L*H…Ha [ˈσ σ σ]

► Two examples of early rise **L*H…Ha** in **Odia**

L	H	H	L		H
L*H		Ha	L*H		(Ha)
pabana			surjyaṅku		kahila
wind			sun-HON-OBJ		say-P

(landmarks (303), tones (289), words (227), English (227); 400 Hz, 276.4 Hz, 150 Hz)
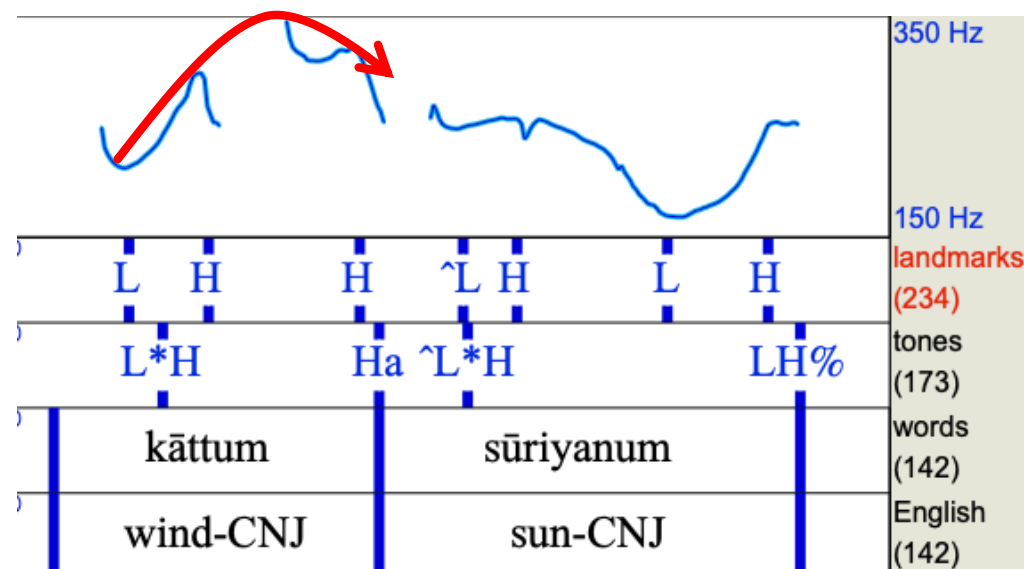
► 'The Wind (told) the Sun…' <NWS-Odia-D01b>

► Notice the slight fall after L*H peak

Early rises: L*H…Ha [ˈσ σ σ]

▶ Example of early rise **L*H…Ha** in **Tamil**

▶ '…the Wind and the Sun…' <NWS-Tamil-J>

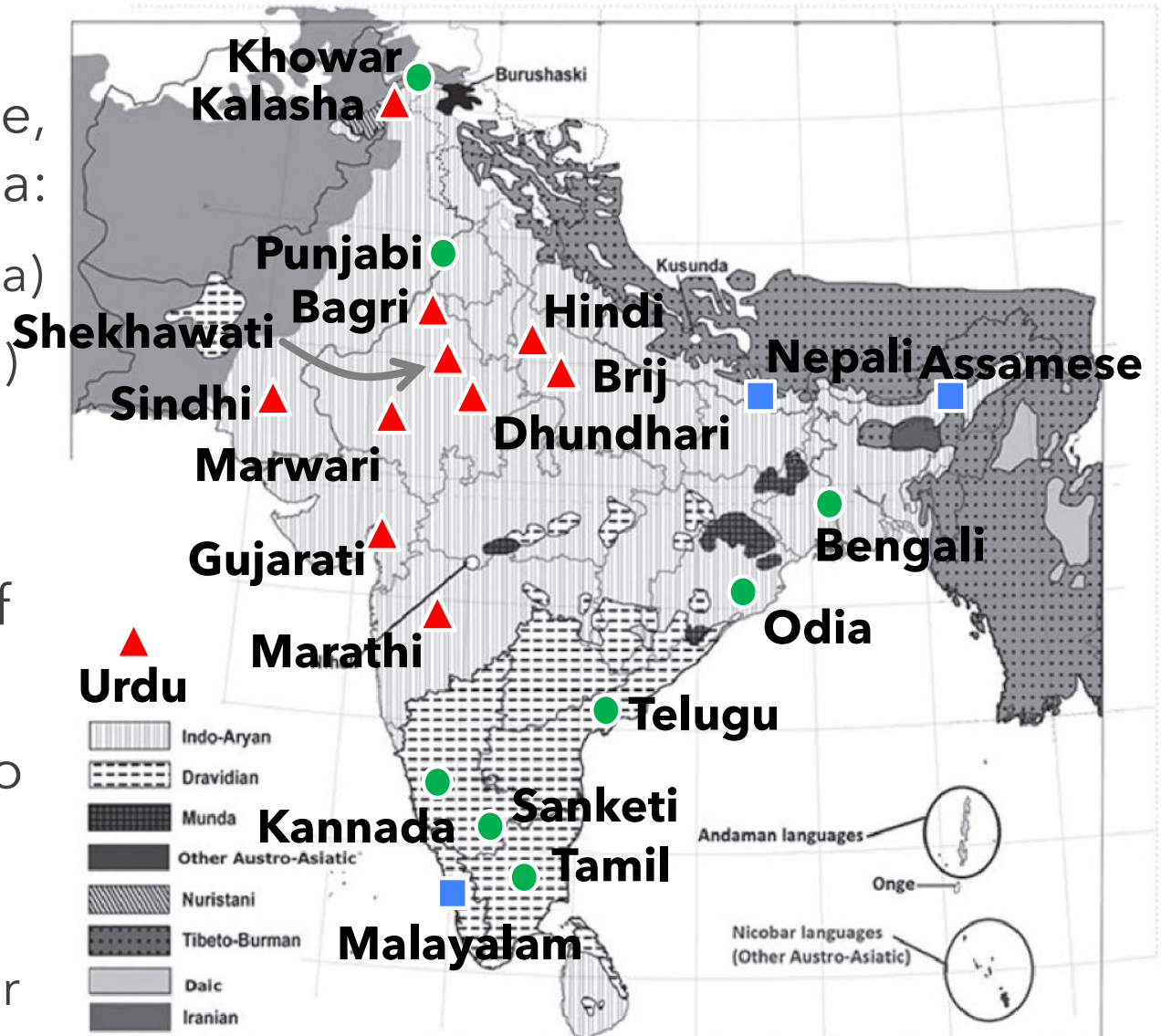# Early rises: L*H…Ha [ˈσ̑ σ σ]

- ▶ This early rise (L*H…Ha) might also be a variant of the direct rise (L*…Ha)

- ▶ Variation within speaker, within a sentence

- ▶ "Free" variation, w/ lg-specific preferences?

- ▶ However, for those using the H as a sign of phrasing: **peak is generally not on the edge**

- ▶ Overlooking this could affect analysis

Early and late rises [´σ σ σ] [´σ σ σ]

- ► Most frequent rise, aside from L*…Ha:
  - ● Early rise (L*H…Ha)
  - ■ Late rise (L*…LHa)
  - ▲ neither found
- ► E & S lgs make substantial use of alternative rises
- ► Other lgs have no alternative rise
  - ► We'll return to Punjabi, Khowar

South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

More complex contours

▶ So far, we've assumed that even with some variation, SALs use LH contours on all words

▶ But exploring this corpus reveals more complex rises, involving medial tones:

  ▶ LLH (late rise):   L*…**L**Ha   [ˈσ σ σ]

  ▶ LHH (early rise):   L***H**…Ha   [ˈσ σ σ]

▶ We also see contours that involve falls:

  ▶ LHL (rise-fall):   L*H…**L**a   [ˈσ σ σ]

  ▶ HLH (fall-rise):   **H**L*…Ha   [σ ˈσ σ]

  ▶ LHLH (double rise):   L***H**…**L**Ha   [ˈσ σ σ]

Rise-falls: L*H...La [ˈσ͡σ͡σ]

▶ Example of **L*H...La** in **Tamil**

350 Hz	
150 Hz	
landmarks (234)	L H L H L L H
tones (173)	L*H LaL* HaL* LH%
words (142)	kambaḷiye eḍukka vaikkirāṅgaḷō
English (142)	blanket-ACC take.off-V put-PRS-3PL-DUB

▶ '...can make (the traveler) take off the shawl...'

Alt. rise-falls: LH*...La [σˈσ σ]

▶ **Sindhi** has another rise-fall, with H on stress:

	L		H	L		H	L	H
		LH*	La		LH*	La		LH*	La
ta	ˈjeko			muˈsāfira		jo		ˈkoṭu	la
C	whoever			traveler		of		coat	?

300 Hz
100 Hz
landmarks (192)
tones (143)
words (151)
English (151)

▶ 'That whoever...the traveler's coat...' <NWS-Sindhi-J>

Fall-rises: HL*…Ha [σ'σ̃σ]

▶ As fall-rises involve a pitch accent **preceded** by H, they only occur in SALs with non-initial stress

▶ Two examples of **HL*…Ha** in **Urdu**

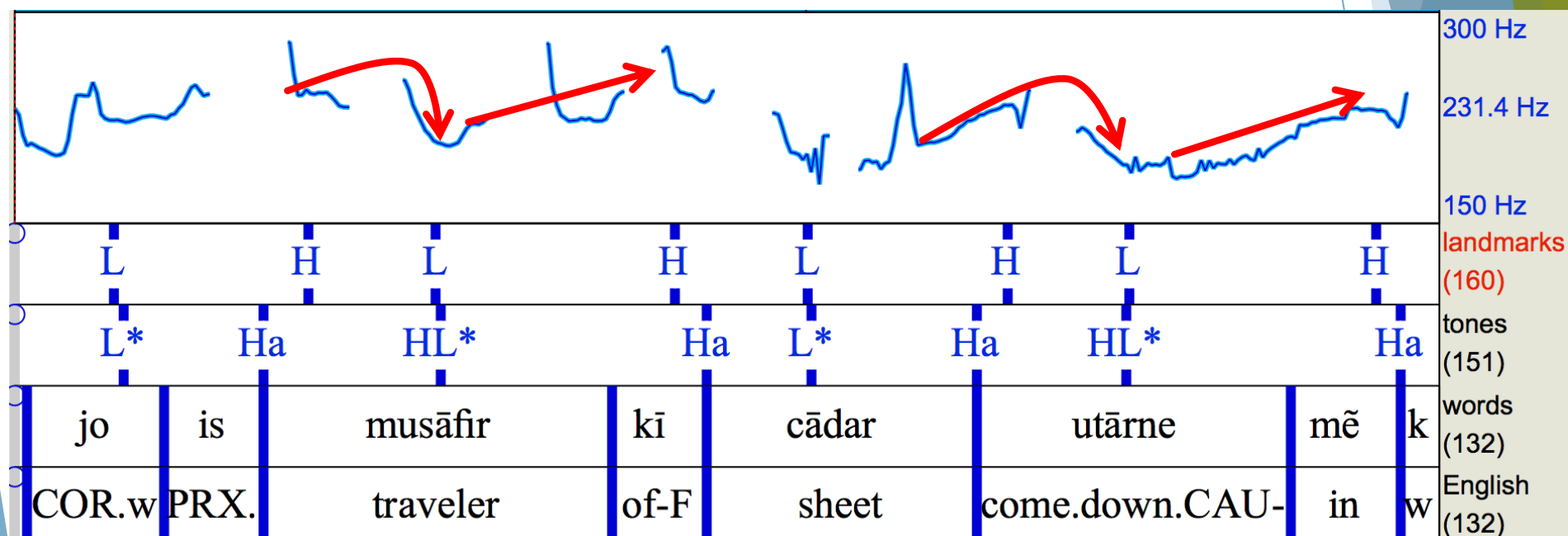|										
landmarks (160)	L		H	L		H	L			H
tones (151)	L*	Ha		HL*		Ha	L*	Ha	HL*	Ha
words (132)	jo	is		musāfir	kī		cādar		utārne	mẽ	k
English (132)	COR.w	PRX.		traveler	of-F		sheet		come.down.CAU-	in	w

▶ 'The one who…in taking off the traveler's sheet…'

<NWS-Urdu-R01b>

Fall-rises: HL*…Ha [σ́ σ̃ σ]

▶ **HL*…Ha** in **Urdu** is clearest when in the initial AP

L	H	H	L	H	L	H
HL*	Ha		HL*	Ha	L*	Ha
musāfir	ne	parešān			hokar	fau
traveler	ERG	uncomfortable			be-CNJ	im

300 Hz
150 Hz
landmarks (159)
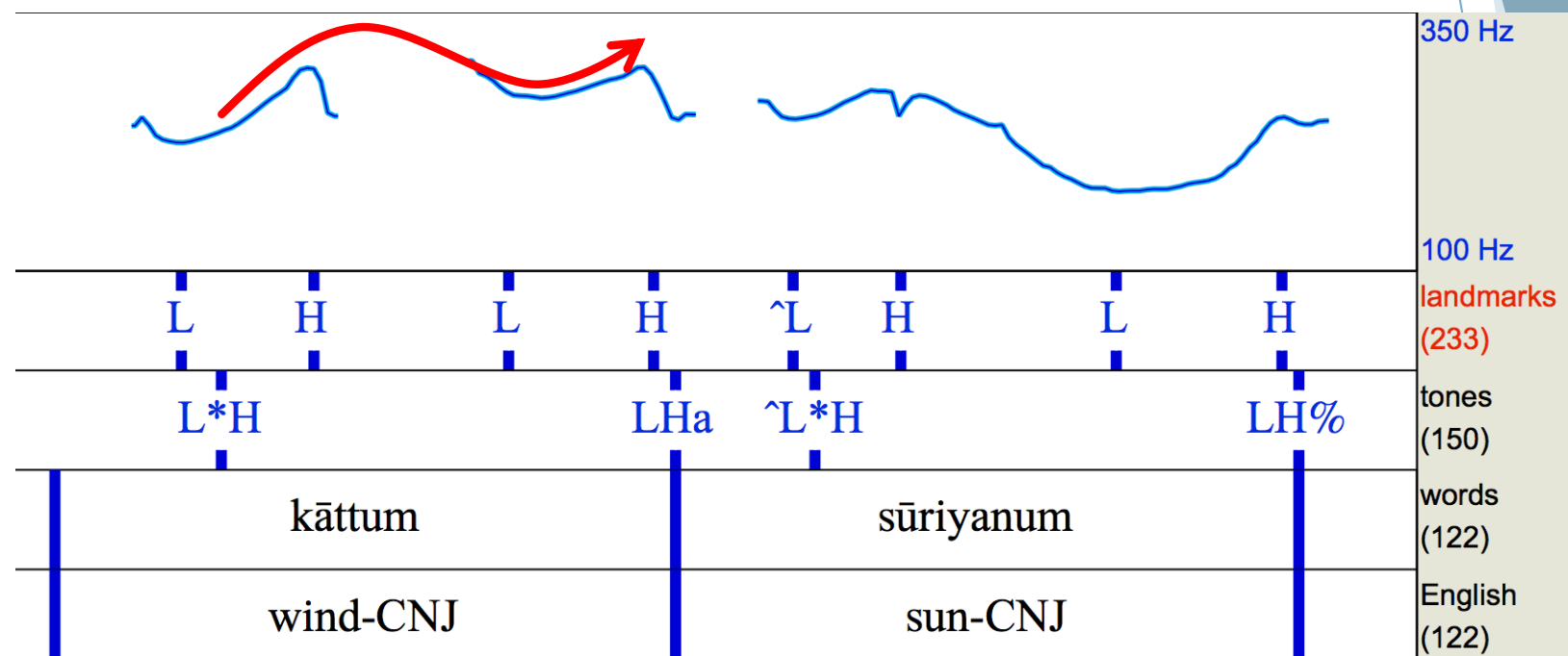tones (151)
words (132)
English (132)
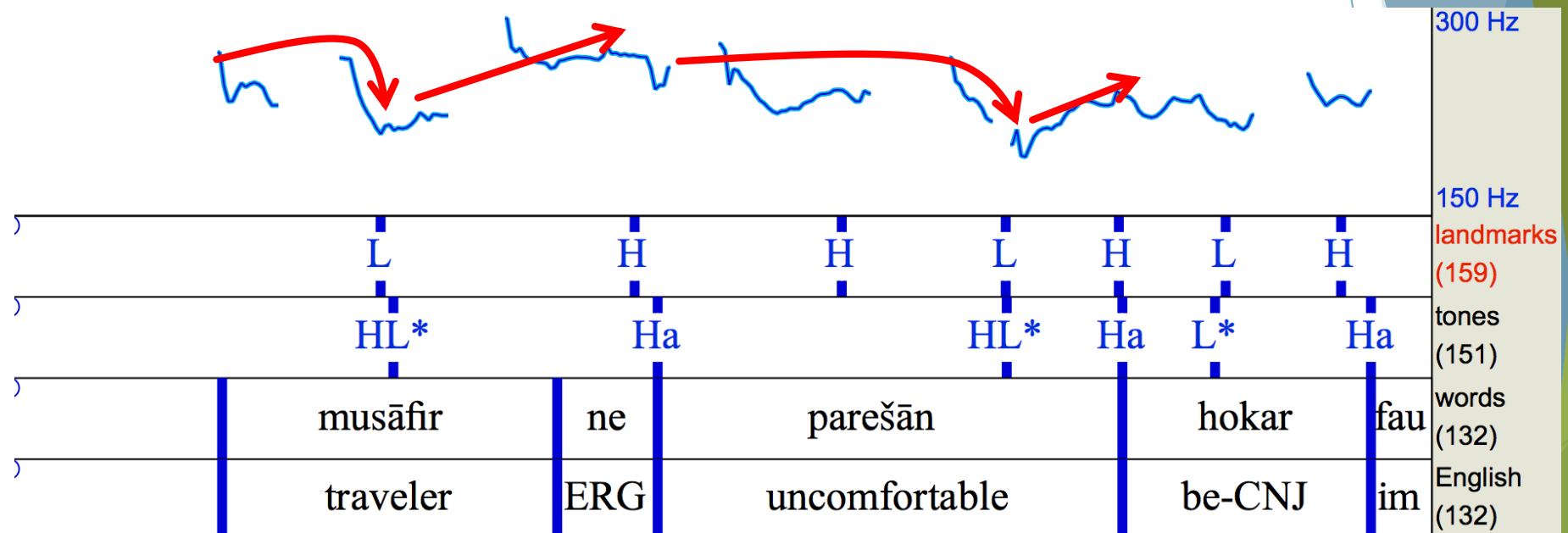
▶ 'Having gotten uncomfortable, the traveler…'
<NWS-Urdu-R01b>

Double-rises: L*H...LHa [ˈσ̃σ̃σ̃]

▶ Double rises are characteristic in Dravidian lgs e.g. **Tamil**:

	L	H	L	H	ˆL	H	L	H
	L*H			LHa	ˆL*H			LH%
	kāttum				sūriyanum			
	wind-CNJ				sun-CNJ			

350 Hz

100 Hz

landmarks (233)

tones (150)

words (122)

English (122)

▶ 'The North Wind and the Sun...' <NWS-Tamil-J>

Double-rises: L*H...LHa [ˈσ̃σ̃σ]

▶ Three double rises in **Kannada**:

L	H	L	H	L	H	L	H	L	H	L	H	L	H	L

L*H	LHa-	L*H	LHa-	^L*H	LHa- L*!H	L%

konege	gāḷiyu	sōtu	summan	āyitu
end-INS	wind-NOM	lose-PST-3NSG	relent	become-PST-3NSG

400 Hz
100 Hz
landmarks (232)
tones (133)
words (105)
English (105)
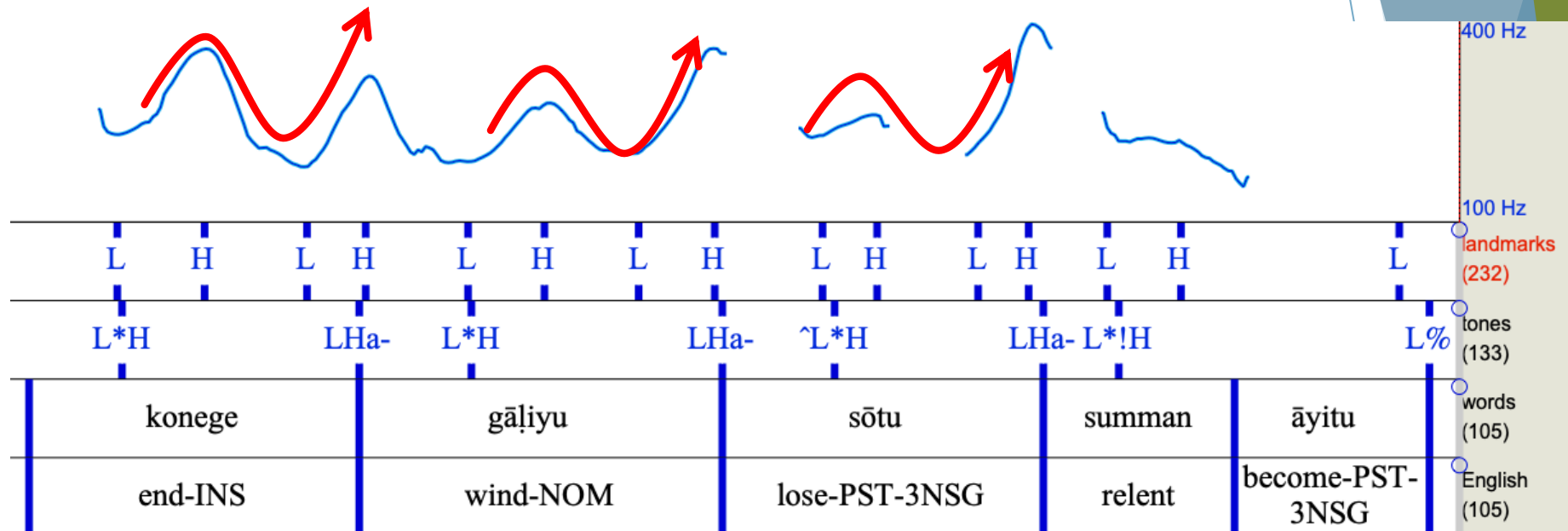
▶ 'In the end the Wind admittedly (?) lost...'
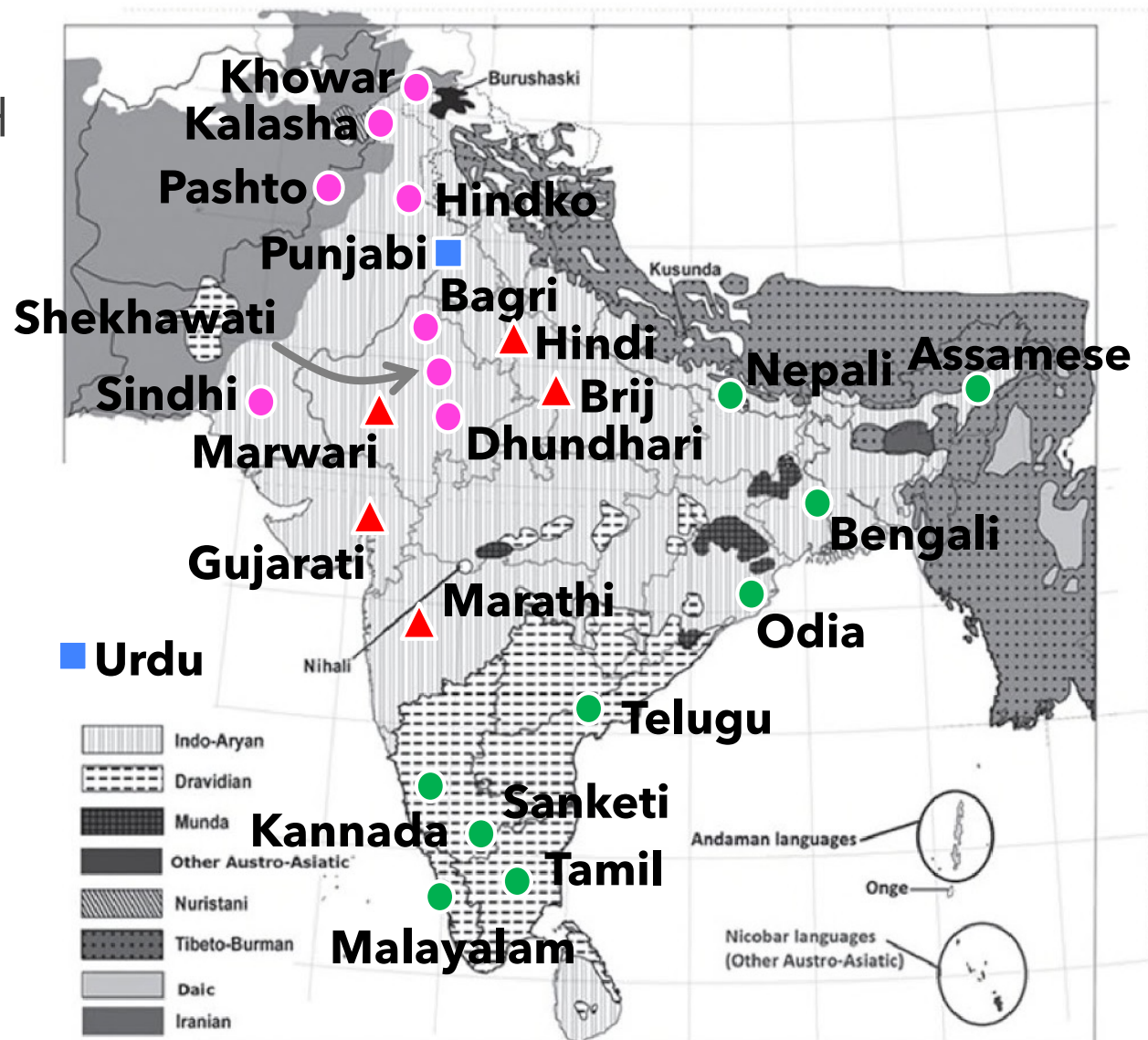<NWS-Kannada-R01d>

Complex pitch accents

▶ We've now seen an expanded inventory of pitch accents across SALs

▶ We can distinguish them based on the **position of the H vis-à-vis the stressed σ**

Complex pitch accents

- Stress-attracted H tone typically:
  - ● Trails stress (L***H**)
  - ● Is on stress (L**H***)
  - ■ Leads stress (**H**L*)
  - ▲ None attested

- N & far N lgs prefer H **on or before** the stress
- S & E lgs prefer H **after the stress**

South Asian language families (map produced by Suresh Kolichala, 2015)
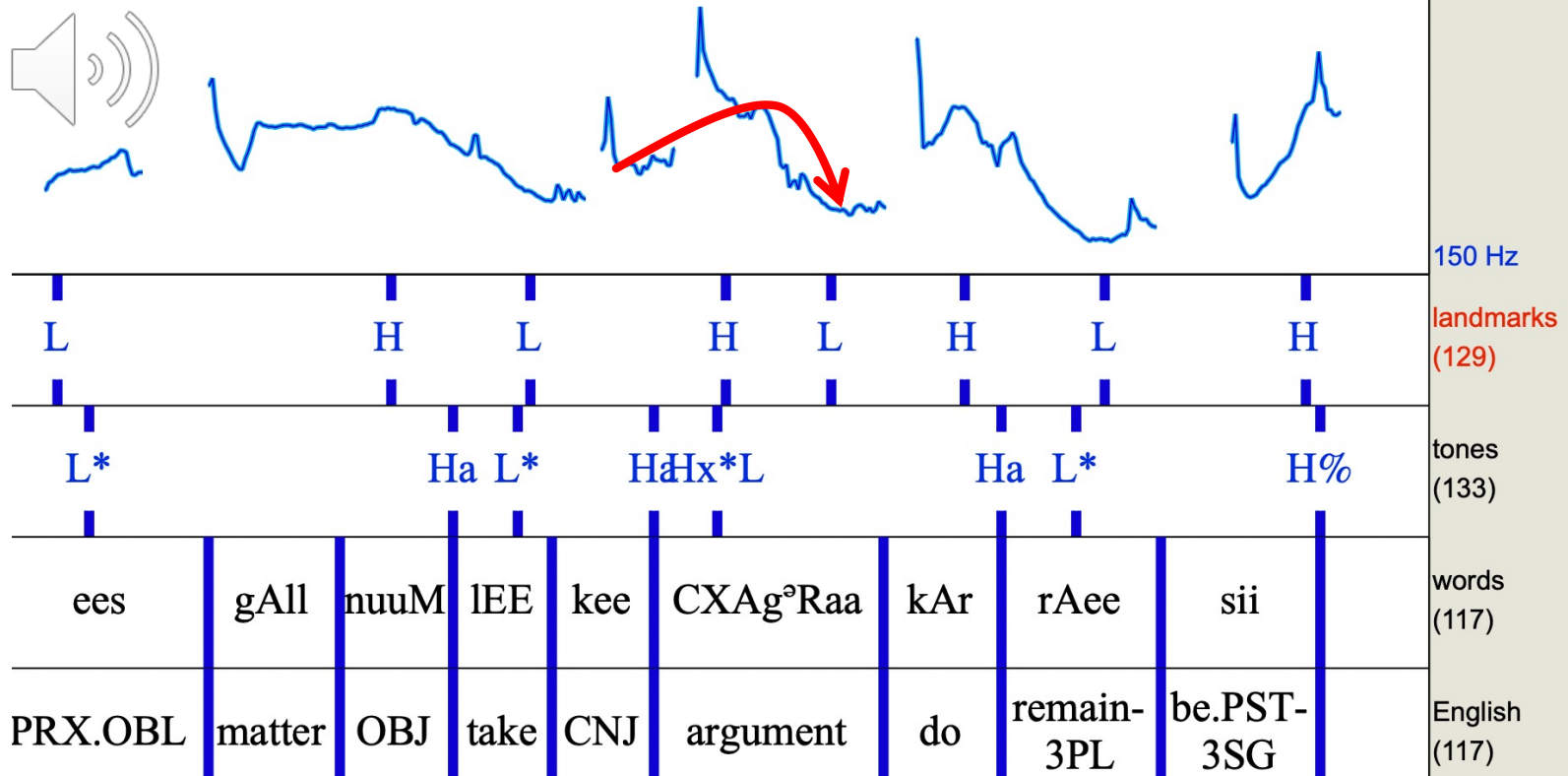Base map taken from Hock (2016, p.7)

Complex pitch accents

- As we move north and west, the SALs…
  - …are **weight-sensitive**
  - …have **right-leaning stress** placement
  - …have **earlier H** in complex pitch accents
  - …resemble Arabic, English, Swedish
- As we move south and east, the SALs…
  - …have a **fixed, left-leaning stress** pattern
  - …have a **trailing H** in complex pitch accents
  - …**rely less** on "default" Ha
  - …resemble Korean, Mongolian, French

Patterns to cover

▶ The main patterns I plan to cover are:

  ▶ The alignment of **L to the stressed σ**

  ▶ The **variable alignment of H**

  ▶ The presence of **more complex contours** than LH

  ▶ The effect of **lexical pitch accent** ("lexical tone")
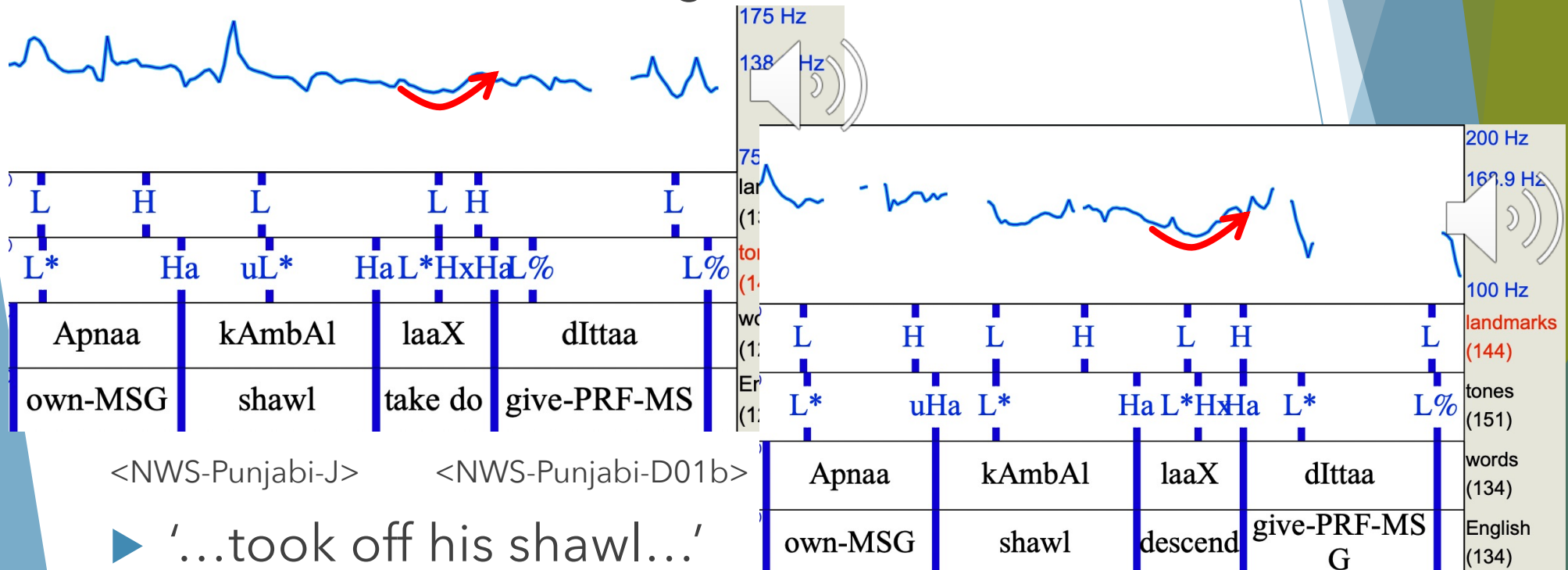
Lexical pitch accent [ˈσ σ σ]

▶ Stressed σ's with historical breathy onsets in Punjabi bear **falling lexical accent Hx*L**, not L*

L	H	L	H	L	H	L	H
L*		Ha	L*	HaHx*L		Ha	L*	H%

ees	gAll	nuuM	lEE	kee	CXAgᵊRaa	kAr	rAee	sii
PRX.OBL	matter	OBJ	take	CNJ	argument	do	remain-3PL	be.PST-3SG

350 Hz

150 Hz

landmarks (129)

tones (133)

words (117)

English (117)

▶ '…were arguing on this matter…' <NWS-Punjabi-D02b>

Lexical pitch accent [ˈσ σ σ]

▶ Stressed σ's followed by historical breathy C's in Punjabi bear **rising lexical accent L*Hx**

▶ Can be hard to distinguish from L*…Ha rise

| L	H	L	L H	L
L*	Ha	uL*	Ha L*HxHa L%	L%
Apnaa	kAmbAl	laaX	dIttaa	
own-MSG	shawl	take do	give-PRF-MS	

<NWS-Punjabi-J>   <NWS-Punjabi-D01b>

L	H	L	H	L H	L
L*	uHa L*		Ha L*HxHa L*		L%
Apnaa	kAmbAl		laaX	dIttaa	
own-MSG	shawl		descend	give-PRF-MS G	

175 Hz
138 Hz
75
200 Hz
169.9 Hz
100 Hz
landmarks (144)
tones (151)
words (134)
English (134)

▶ '…took off his shawl…'

Lexical pitch accent

- The Punjabi pattern is:

  - Historically **breathy onset → falling** accent **Hx*L**:

    - ˈdʒʱəgɽɑ → ˈtʃ�âgɽɑ 'argument'

    - ˈgʱʊ̂ʈ̣ː → ˈkʊ̂ʈ̣ː 'tight'

  - Historically **breathy offset → rising** accent **L*Hx**:

    - ˈɪɦ → ˈě 'this (OBL)' (cf. Hindi ɪs)

    - ˈpæɦli → ˈpěli 'first'

Lexical pitch accent

▶ In **Shekhawati**, stressed σ's with historical breathy onsets bear a **rising lexical accent Lx*H**

  ▶ NB: Punjabi has falling HxL* in this position

▶ Can be hard to distinguish from L*…Ha
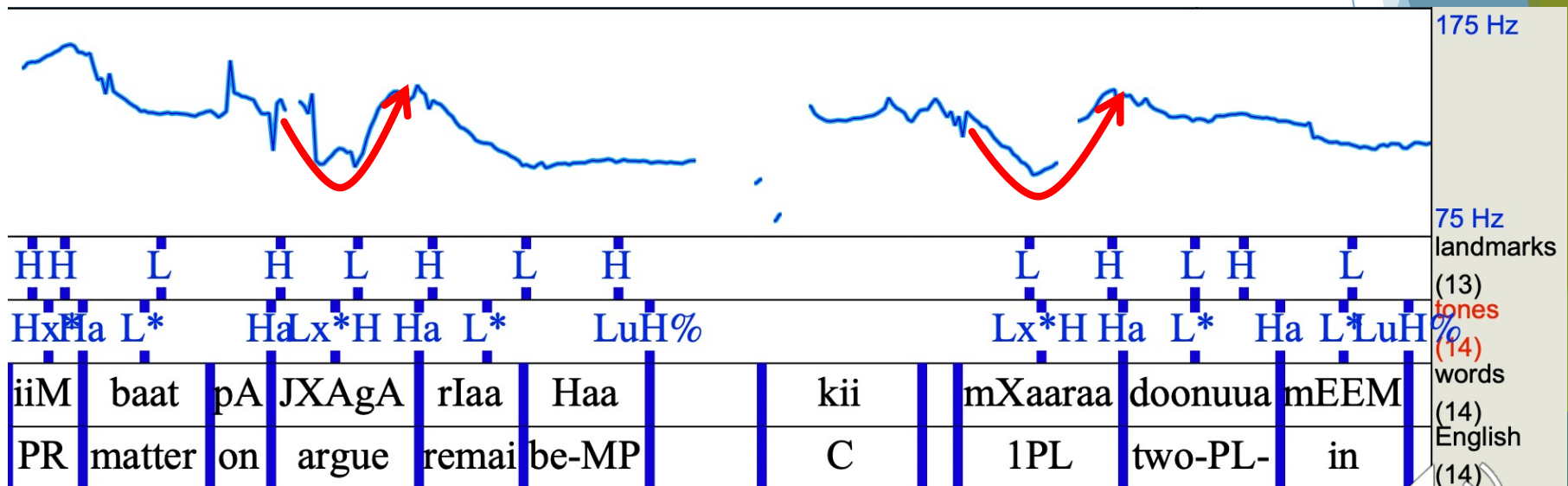
▶ '…were arguing on this matter of (who) among the two of us…' <NWS-Rajasthani-Shekhawati-D01a>

Lexical pitch accent

▶ In **Shekhawati**, and (to a lesser extent) **Dhundhari**, **Merwari**, and **Bagdi** (Gusain 1999), stressed σ's followed by historical breathy C's bear **high lexical accent Hx***
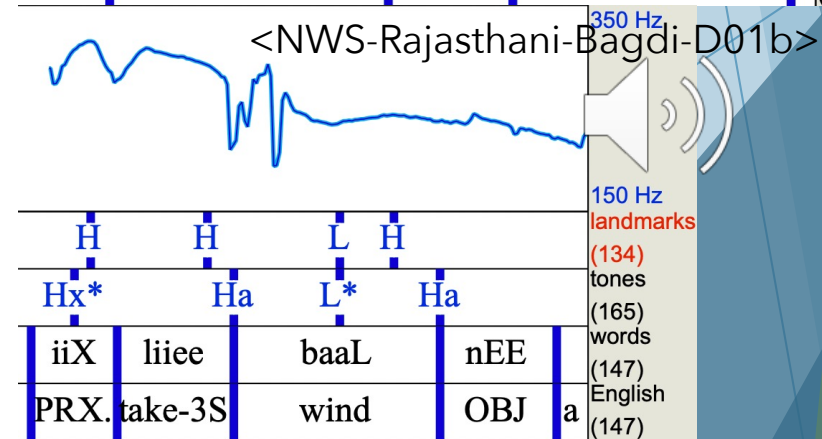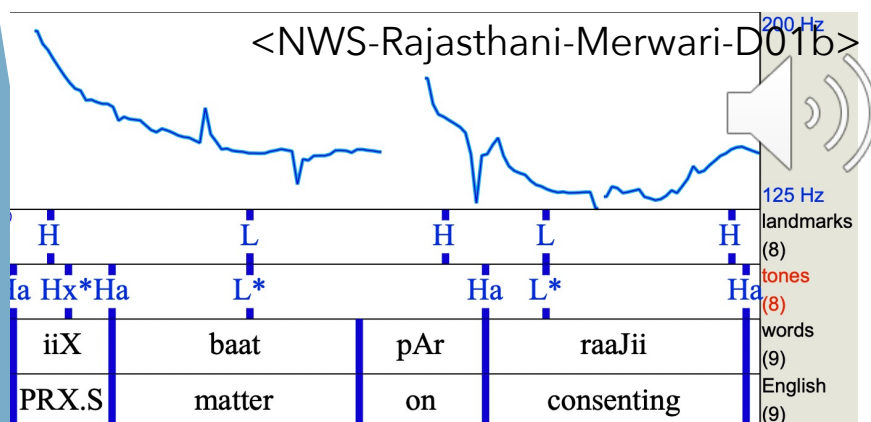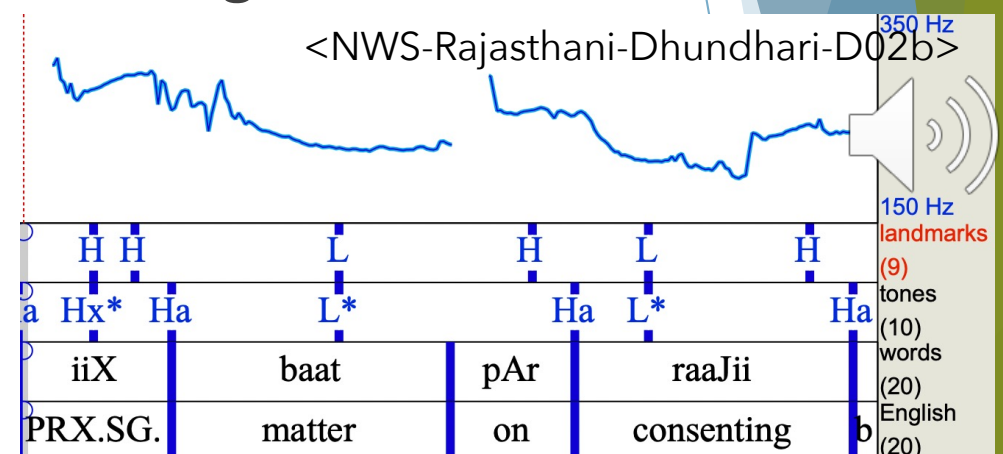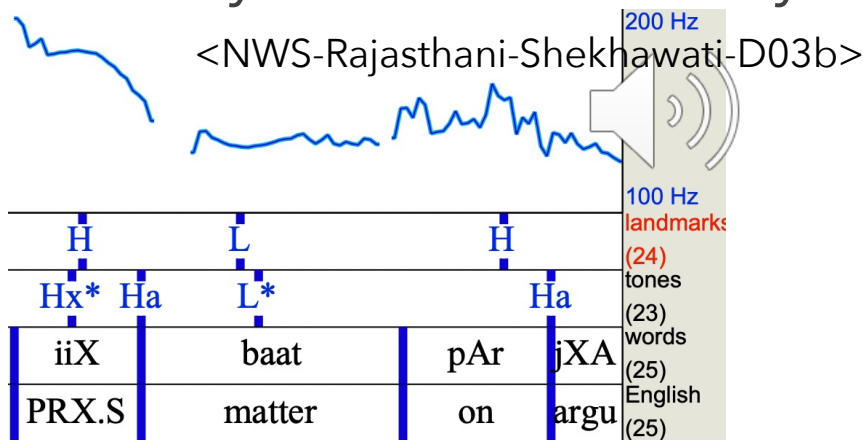
Lexical pitch accent

- The more robust Shekhawati pattern is:
  - Historically **breathy onset → rising** accent **Lx*H**:
    - ˈdʒʱəgəɽ → ˈdʒə̌gəɽ 'argument'
    - ˈm̤ʱɑɾɑ → ˈmɑ̌ɾɑ 'us'
    - bəˈɽ̥ʱaʈːi → bəˈɽɑ̌ʈːi 'increasing (FSG)'
  - Historically **breathy offset → high** accent **Hx***:
    - ˈɪɦ → ˈí 'this (OBL)' (cf. Hindi ɪs)
    - ˈpɛɦli → ˈpɛ́li 'first'
    - ˈsəmdʒʱjo → ˈsə́mdʒjo 'understood (MSG)'
- Similar to the pattern described for Bagdi (Gusain 1999)

Lexical pitch accent

- Lexical input in pitch accent?
  - ■ Yes ("tonal"): 7
  - ▲ No: 18
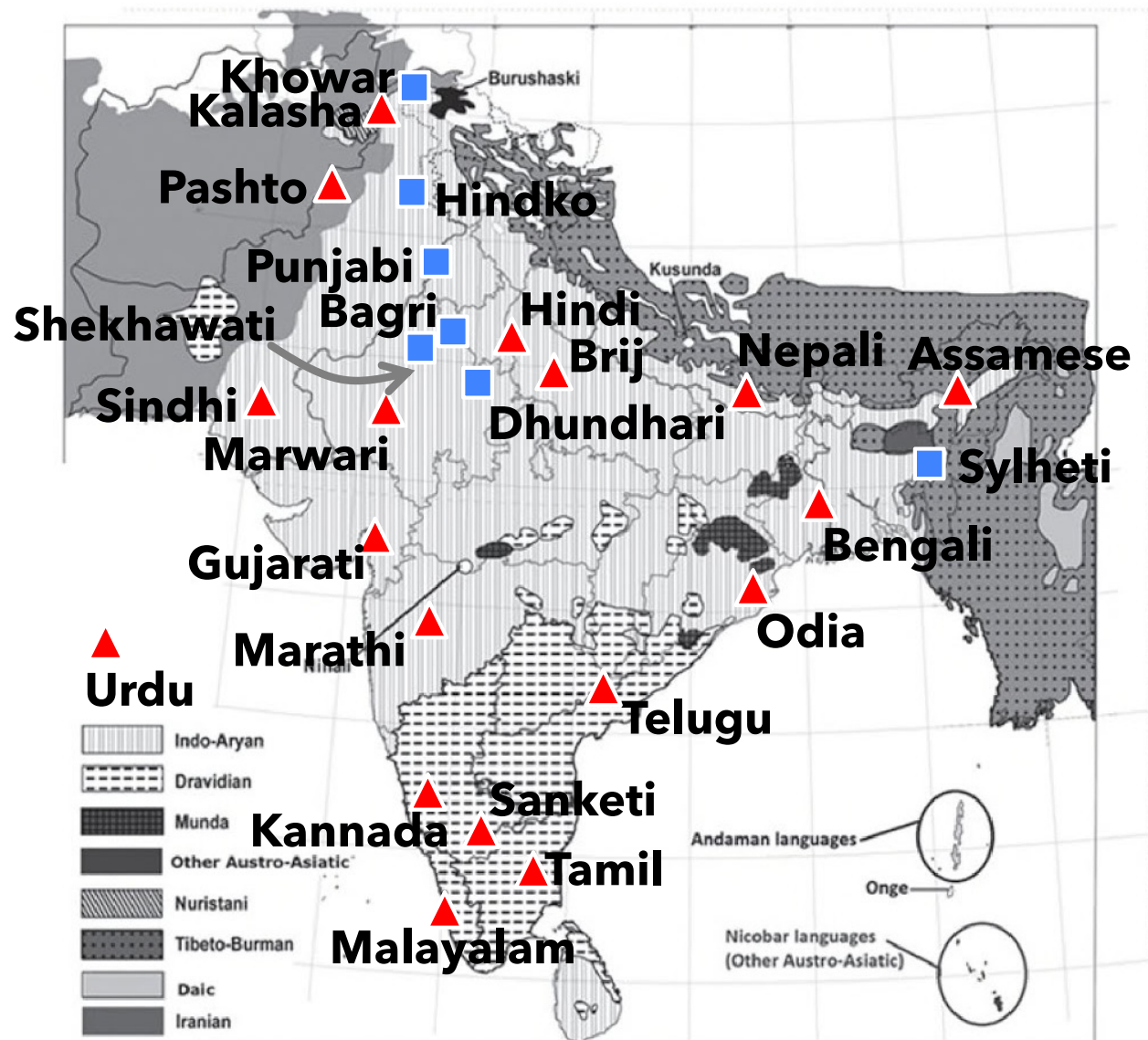
South Asian language families (map produced by Suresh Kolichala, 2015)
Base map taken from Hock (2016, p.7)

Returning to our questions

▶ Is there a **typical SAL intonation**?

▶ What does **having no contrastive tone or stress** do to the intonation system?

▶ What **new phenomena** do SALs contribute to our knowledge of intonational systems?

Returning to our questions

▶ Is there a **typical SAL intonation**?

▶ In some very basic ways, **yes**:

  ▶ Stress placement is generally **not contrastive**

  ▶ Tone is generally **not contrastive**

  ▶ Words generally bear **rising contours**

▶ We already knew these things (Ladd 1996, Féry 2010)

Returning to our questions

▶ Is there a **typical SAL intonation**?

▶ But we see substantial **differences within/across lgs**

 ▶ **L of rise** might mark **L-edge** and/or **stressed σ**

 ▶ **H of rise** can mark **R-edge** and/or **stressed σ**

 ▶ Lots of **alternative patterns**: medial tones, falls

Returning to our questions

▶ Is there a **typical SAL intonation**?

▶ Our inventory is much larger than any previous model of SALs

  ▶ Pitch accents: L*

  ▶ Boundary tones: Ha

Returning to our questions

▶ Is there a **typical SAL intonation**?

▶ Our inventory is much larger than any previous model of SALs

   ▶ Pitch accents: L*, **L*H, LH*, H*, HL***

   ▶ Boundary tones: Ha, **La, LHa, HLa**

Returning to our questions

▶ Is there a **typical SAL intonation**?

▶ Our inventory is much larger than any previous model of SALs

  ▶ Pitch accents: L*, **L*H, LH*, H*, HL***

  ▶ Boundary tones: Ha, **La, LHa, HLa**

  ▶ ip/IP boundary tones: those at the ends of larger units e.g. sentences

Returning to our questions

▶ What does **having no contrastive tone or stress** do to the intonation system?

▶ **Wide variability** within and across SALs in tonal inventory and frequency

▶ The conflation of stressedness and initialness also allows for **ambiguity in analysis**

▶ Even in lgs where the two are not conflated, **speakers vary** as to whether stressedness or initialness attracts the L tone of the rise

Returning to our questions

▶ What does **having no contrastive tone or stress** do to the intonation system?

▶ These observations suggest that an intonation system without contrastive tone or stress will **concentrate its tones at phrase edges**

  ▶ cf. Féry's (2010) "phrase languages"

▶ Initial stress arguably can be an "edge"

▶ Only in a subset of IA lgs (in N/FN) is the stress "strong" enough to draw a pitch accent away from a phrase edge

Returning to our questions

▶ What **new phenomena** do SALs contribute to our knowledge of intonational systems?

▶ AM Theory may need to incorporate **tones with variable association**

  ▶ The L of a rise can behave like a pitch accent (L*) or a boundary tone (aL)

  ▶ The H of a rise a boundary tone (Ha) or part of the pitch accent (L*H)

  ▶ If L*H, is the H a sort of **boundary tone for a foot**?

Remaining questions

▶ What about **other SAL families**?

▶ Rises described in at least **two TB lgs** (Féry & Fanselow 2020)

  ▶ Meithei

  ▶ Boro

▶ AA?

South Asian language families (map produced by Suresh Kolichala, 2015)

Remaining questions

▶ How else can **lexical properties** interact with the LH rise?

  ▶ Data from **Punjabi, Shekhawati, etc.** suggest different tonal pattern entirely (Kanwal & Ritchart 2015)

  ▶ Data from **Sylheti** suggest scaling variation on same tonal pattern (Gope & Mahanta 2014)

  ▶ Data from **Boro** suggest either lack of rise (Das & Mahanta 2019), or no effect of lexical tone (Féry & Fanselow 2020)

Remaining questions

- How much **within-language variation** is there?
- L*H in **Kannada** (Dr) is weight-sensitive for speaker at top-left, but not for other speaker
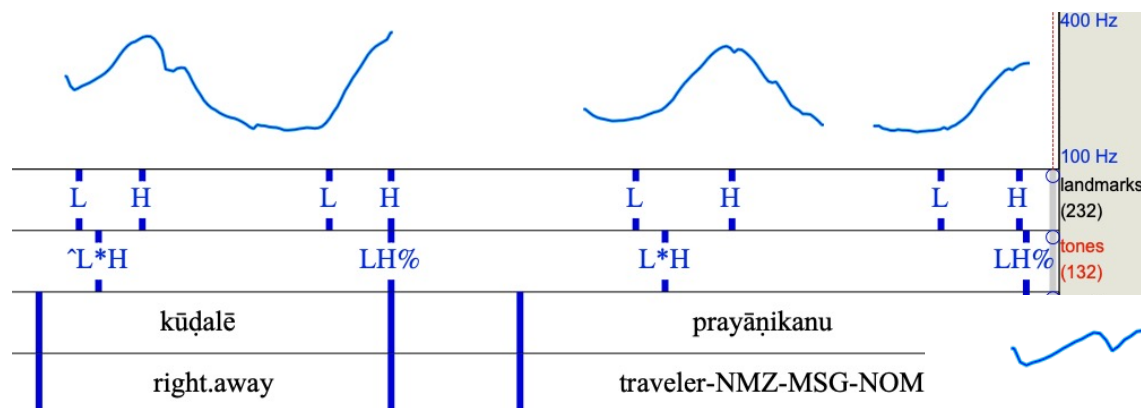
- Influence of Hindi for speaker at top-left?
- Multilingualism
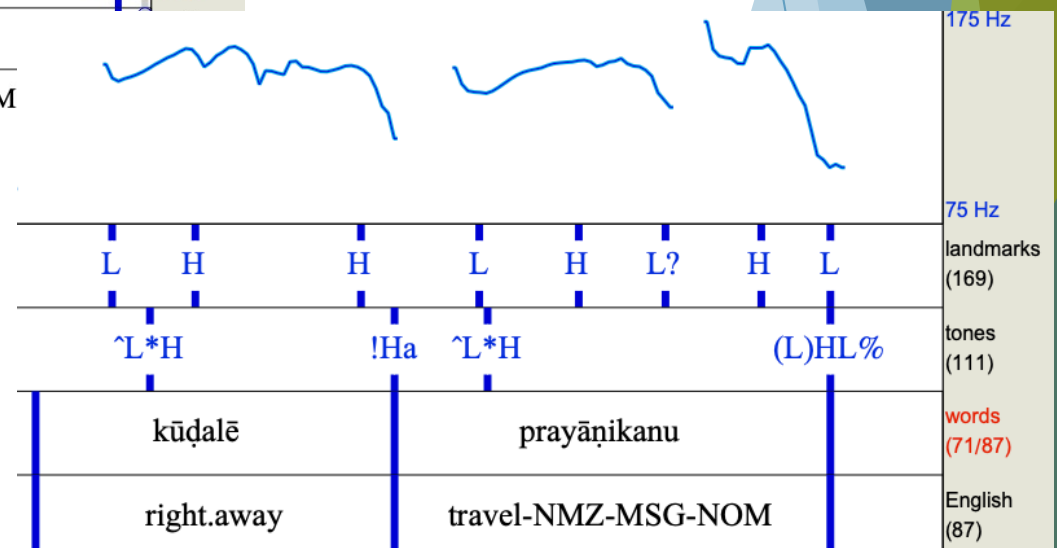
Many thanks

- Andrew Hedding and the UW Dept of Linguistics for inviting me

- Everyone in the audience here

- The authors who shared their recordings with me

- The consultants who provided translations and morphological analysis

- The speakers who contributed new recordings to build the corpus

L*H   **L%**
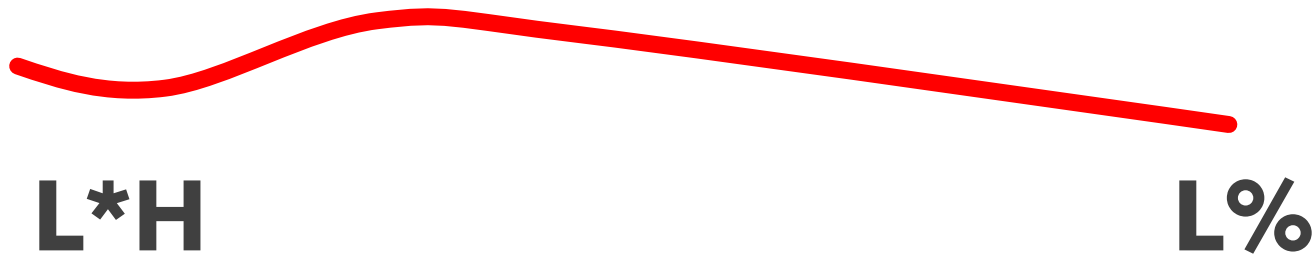
ôšoṅkho dhonnobad

অসংখ্য ধন্যবাদ!

References

Asher, Ron E.; Keane, Elinor. 2005. Diphthongs in colloquial Tamil. In Hardcastle, W.; Beck, J. (eds.), *A figure of speech: A festschrift for John Laver*. Hillsdale, NJ: Lawrence Erlbaum,147-171.

Choudhury, Arunima; Kaiser, Elsi. 2012. Prosodic focus in Bangla: a psycholinguistic investigation of production and perception. Talk, LSA.

Christdas, P. 1988. The phonology and morphology of Tamil. PhD thesis, Cornell University.

Das, Kalyan; Mahanta, Shakuntala. 2015. Variation in prosodic phrasing in Bodo. Workshop on ProVar.

Dasgupta, Probal. 2003. Bangla. In: Cardona, George; Jain, Dhanesh (eds.) *The Indo-Aryan Languages*. Routledge.

Dyrud, Lars O. 2001. Hindi-Urdu: stress accent or non-stress accent? Master's thesis, University of North Dakota.

Emeneau, M. B. 1956. India as a linguistic area. *Language* 32 (1), 3–16.

Féry, Caroline. 2010. Indian languages as intonational 'phrase languages'. In: Hasnain, Imtiaz; Chaudhury, Shreesh (eds). *Problematizing Language Studies. Festschrift for Rama Agnihotri*. Delhi: Aakar Books, 288-312.

Firth, J. R. 1957. Phonetic observations on Gujarati. *Bulletin of SOAS*, 20(1), 231–241.

Genzel, Susanne. 2007. Phonetische Realisierung von Fokus auf Adjektiven im Hindi. MA thesis, Potsdam Univ.

Gill, Harjeet Singh; Gleason, Jr., Henry A. 1962. *A Reference Grammar of Panjabi*. Hartford Studies in Linguistics.

Gope. Amalesh; Mahanta, Shakuntala. 2014. Lexical tones in Sylheti. Presented at TAL-2014.

Harnsberger, James D. 1999. The role of metrical structure in Hindi intonation. Presented at SALA Roundtable.

Harnsberger, James D. 1996. Towards an intonational phonology of Hindi. Presented at Laboratory Phonology V.

Hayes, Bruce; Lahiri, Aditi. 1991. Bengali intonational phonology. *NLLT* 9, 47–96.

References, cont.

Hussain, Sarmad. 1997. Phonetic correlates of lexical stress in Urdu. Northwestern University.

Kanwal, Jasmeen; Ritchart, Amanda. 2015. An experimental investigation of tonogenesis in Punjabi. *Proceedings of ICPhS 18*.

Keane, Elinor. 2003. Word-level prominence in Tamil. *Proceedings of the 15th International Congress of Phonetic Sciences* 2, 1257–1260.

Keane, Elinor. 2004. Tamil. *JIPA* 34(1), 111–116.

Keane, Elinor. 2014. The intonational phonology of Bangladeshi Standard Bengali. In Sun-Ah Jun (ed.). *Prosodic Typology II: the Phonology of Intonation and Phrasing*. Oxford University Press.

Khan, Sameer ud Dowla. 2008. Intonational phonology and focus prosody of Bengali. PhD dissertation, UCLA.

Khan, Sameer ud Dowla. 2010. Bengali (Bangladeshi standard). *JIPA* 40(2), 221-225.

Khan, Sameer ud Dowla. 2014. The intonational phonology of Bangladeshi Standard Bengali. In Sun-Ah Jun (ed.). *Prosodic Typology II: the Phonology of Intonation and Phrasing*. Oxford University Press.

Khan, Sameer ud Dowla. 2016. Rising contours across South Asian languages. Work in progress.

Khatiwada, Rajesh. 2009. Nepali. *JIPA* 39(3), 373-380.

Kingston, John. 2011. Tonogenesis. In van Oostendorp, M.; Ewen, C. J.; Hume, E.; Rice, K. (eds.) *Blackwell Companion to Phonology, v. 4*, (pp. 2304-2334). Oxford: Blackwell Publishing.

References, cont.

Ladd, D. Robert. 1996. *Intonational Phonology.* Cambridge University Press.

Lahiri, Aditi; Fitzpatrick-Cole. 1999. Emphatic clitics and focus intonation in Bengali. In: Kager, René; Zonneveld, Wim (eds.), *Phrasal phonology*, 119–144. Nijmegen: University of Nijmegen Press.

Mahanta, Shakuntala. 2001. Some aspects of prominence in Assamese and Assamese English. MPhil diss.

Mahanta, Shakuntala. 2012. Assamese. *JIPA* 42(2), 217–224.

Mahanta, Shakuntala; Das, Kalyan. 2012. Prosody of focus in Bodo. 18th Himalayan Linguistics Symposium.

Mahanta, Shakuntala; Raychoudhury, Priti. 2016. Focus and intonation in Dimasa. Work in progress.

Mahesh, M. 2016. Phonology of intonation in Malayalam. Work in progress.

Malik, Amar Nath. 1995. *The phonology and morphology of Panjabi*. New Delhi: Munshiram Manoharlal Pub.

Nihalani, Paroo. 1999. Sindhi. *Handbook of the IPA.* Cambridge University Press, 131–134.

Ohala, Manjari. 1999. Hindi. *Handbook of the IPA*. Cambridge University Press, 100–103.

Patil, Umesh; Kentner, Gerrit; Gollrad, Anja; Kügler, Frank; Féry, Caroline; Vasishth, Shravan. 2008. Focus, word order, and intonation in Hindi. *JSAL* 1(1).

Subbārāo, Kārumūri. V. 2012. *South Asian languages: a syntactic typology*. Cambridge University Press.

Twaha, Asimul Islam; Mahanta, Shakuntala. 2016. Intonational and prosodic structure of Standard Colloquial Assamese. Talk, Workshop on Tone and Intonation-3.

Ullah, Rashad. 2016. Epistemic indefinites, negative polarity, and free choice in Bengali. Poster, FASAL-6.