Intonational phonology in Bengali and English infant-directed speech

Kristine Yu, Sameer ud Dowla Khan &
Megha Sundara

Under revision: Lab Phonology

Infant-directed speech

- Cross-linguistically, compared to adult directed speech, infant-directed speech (IDS) has
  - slower speech rate
  - shorter utterances

  - higher mean fundamental frequency (F0),
  - higher maximum F0,
  - expanded F0 range,
  - greater F0 variability,

Why f0-related differences in IDS?

- To regulate infant attention and affect
  - IDS ="unstructured exaggeration of the pitch contour"
    - Gradient, language-general phonetic view
  - Possible alternatives
    - IDS = "exaggeration of tonal targets seen in ADS"
      - That is, allophonic differences in IDS
      - Language-specific phonetic view
    - IDS = "different choice of tonal targets compared to ADS"
      - Language-specific phonological view

Previous research

- Evidence for unstructured exaggeration
  - In a lexical tone language, should interfere with transmission of meaning
    - Poor tone identification in Thai in utterance-final not initial position (Kitamura et al., 2002)

- Evidence for allophonic expansion of pitch range
  - F0 expansion, but no compromised transmission of tonal targets
    - Mandarin (Liu, Tsao & Kuhl, 2007); Cantonese (Xu, 2008); Hakka (Cheng and Chang, 2014)
  - F0 expansion, but no compromise of boundary tones
    - Japanese IDS (Igarashi, Nishikawa, Tanaka, and Mazuka, 2013)
      - Expansion restricted to the more variable boundary tones, not the less variable pitch accents

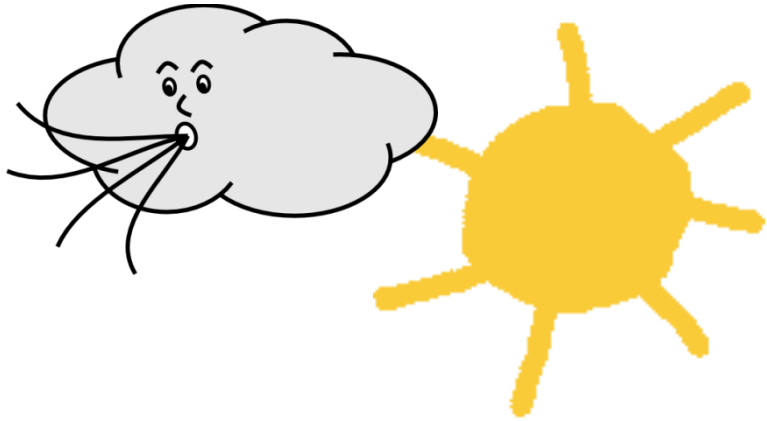
This study

- Compare ADS and IDS in 2 ways
  - Choice of tonal targets
  - Implementation of tonal targets
    - Boundary tones and pitch accent

Methods

- Compared adult- and simulated infant-directed speech
  - North Wind and Sun fable
- Native speakers of American English and Bengali (n=10 each)
  - Neither has tone, or lexical pitch accent
  - Both have intonation
- All recordings in sound booth with a Shure SM10A head-mounted microphone plugged into a laptop computer via a preamplifier
- 2 conditions (IDS, ADS) x 3 repetitions = 6 recordings per speaker
  - 2 English transcribers trained on ADS
  - 1 Bengali transcriber

Two styles

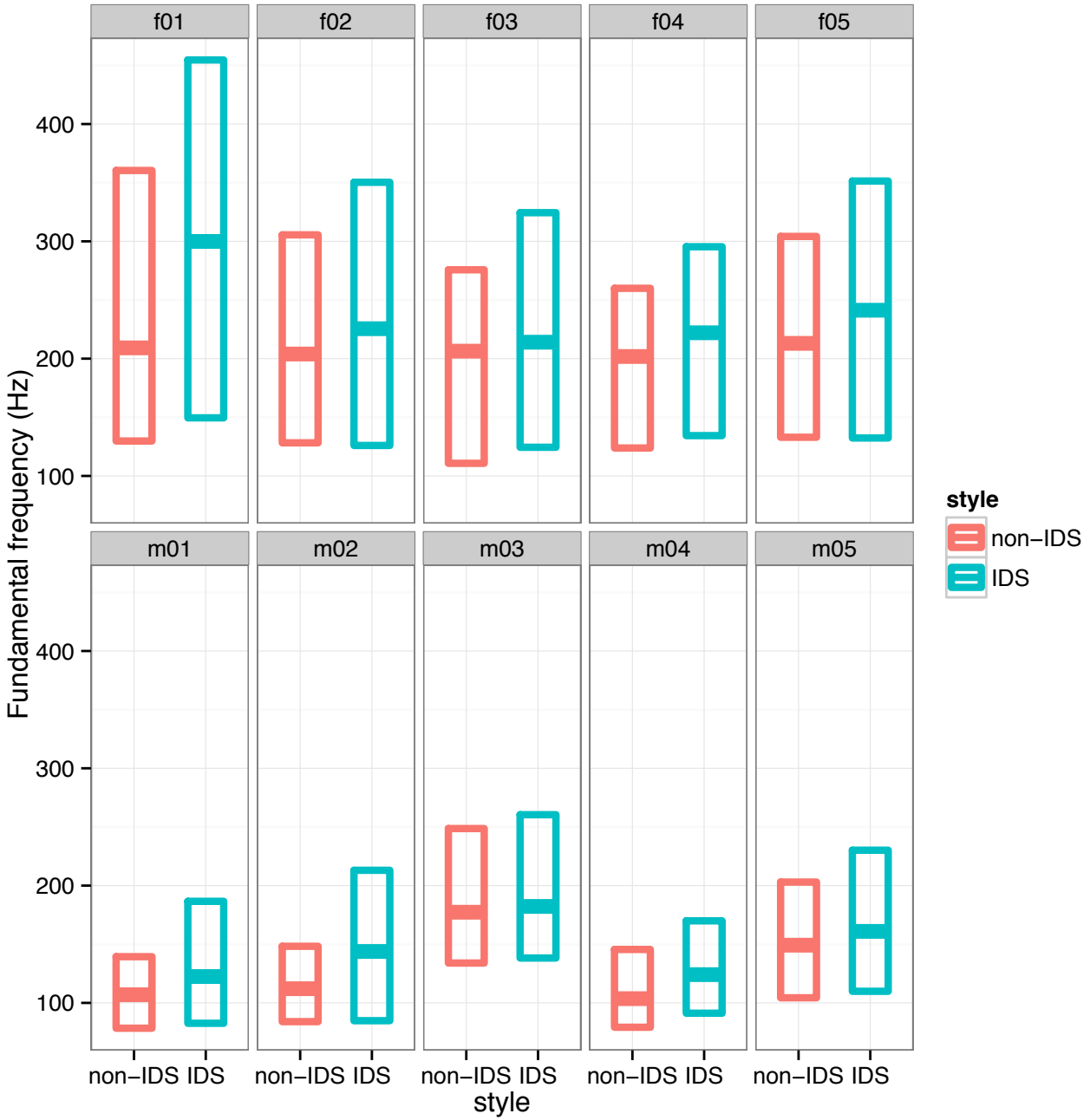


- **Default reading** (non-IDS): “Read at a comfortable pace.”
- **Simulated infant-directed reading** (IDS): “Read as if speaking to your 4-mo-old child.”

F0 Range

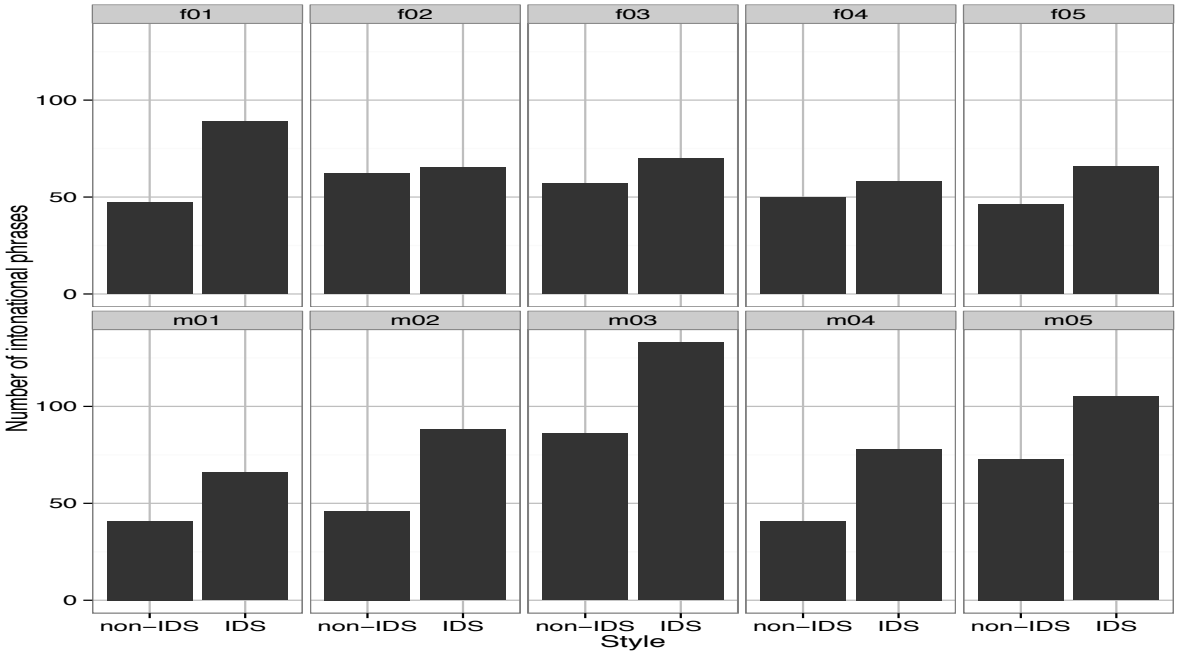
Bengali

The middle line = mean F0 averaged across IPs for a given speaker; bottom & top bounds = the minimum & maximum F0 averaged across IPs.

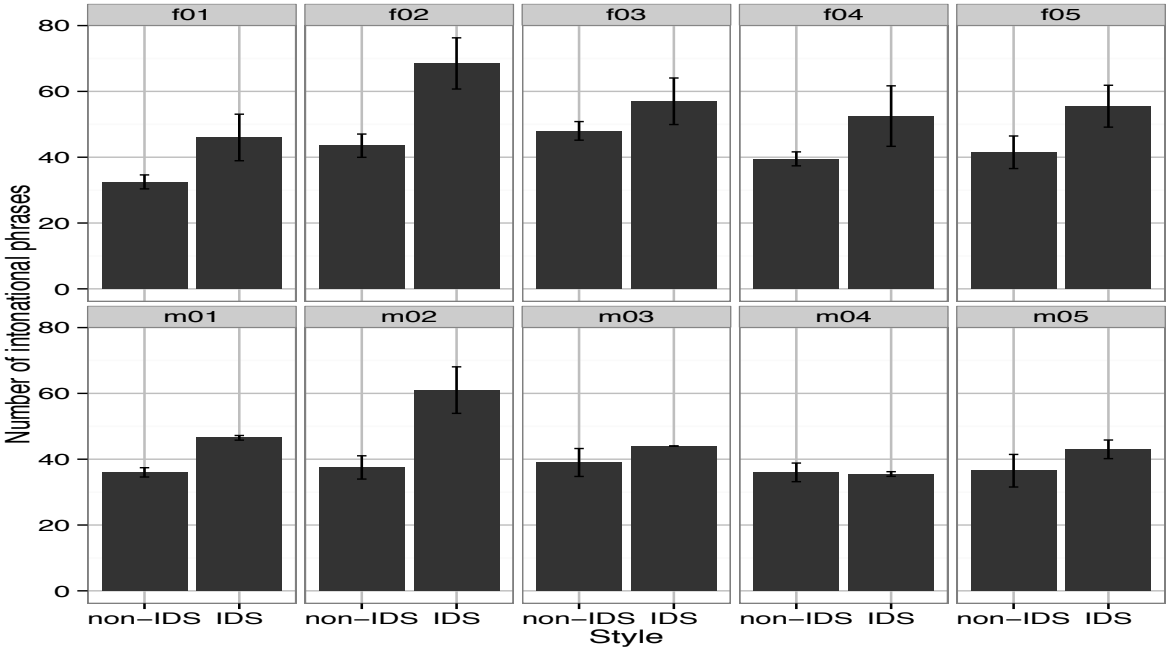


IPs

Bengali

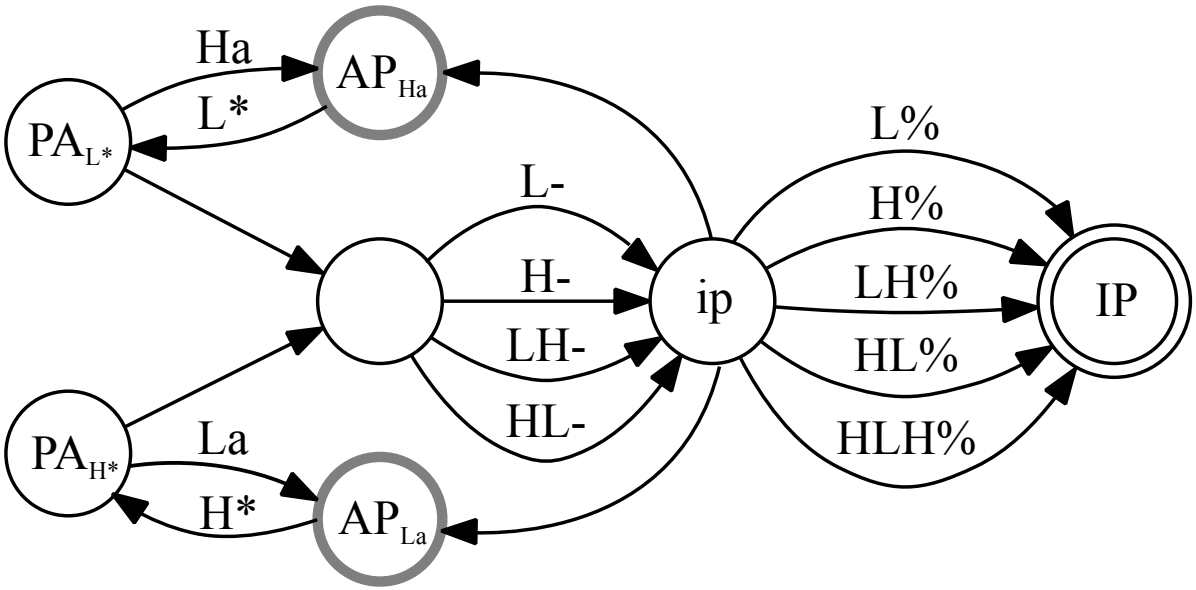


English



BENGALI

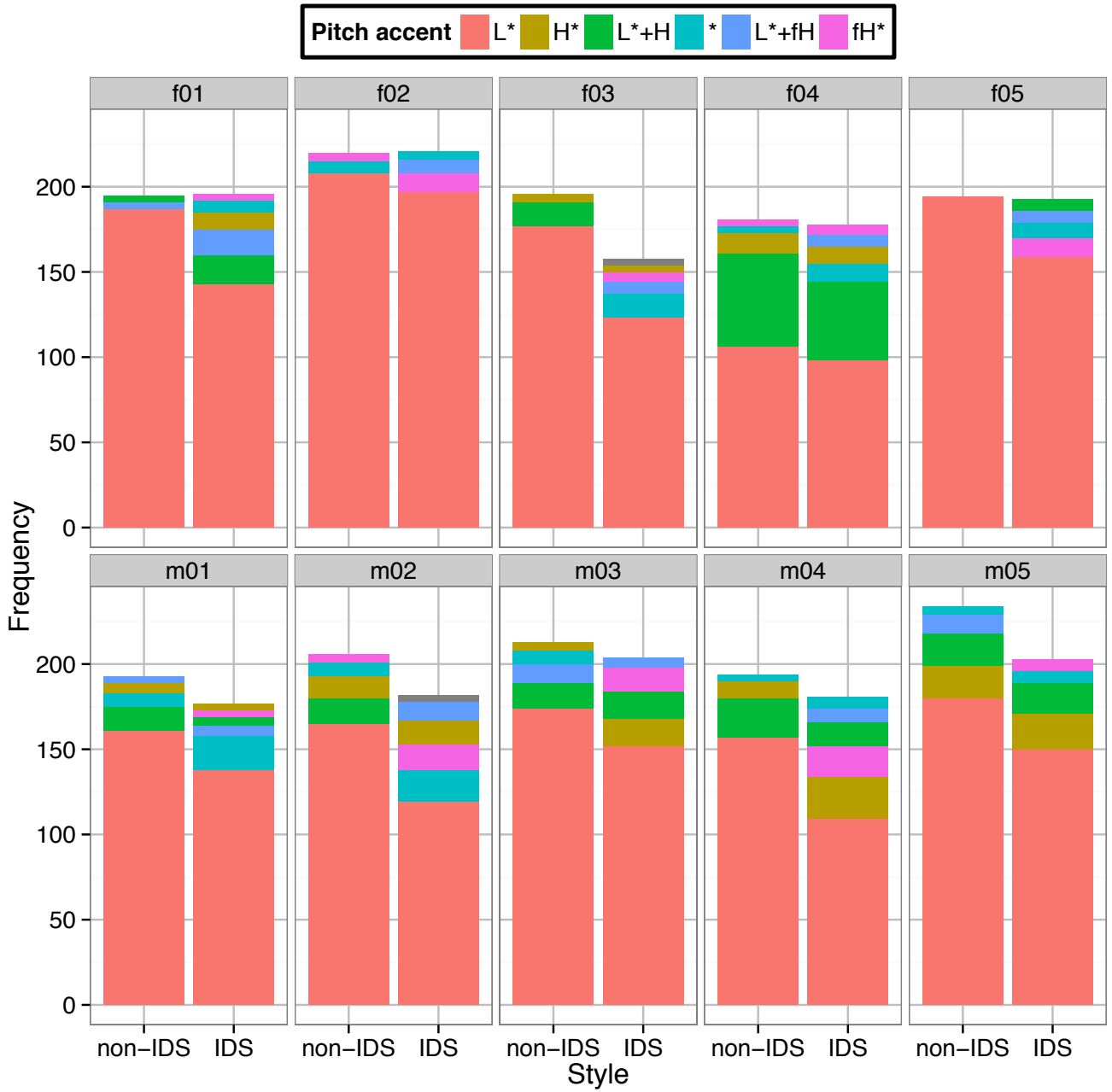
B-ToBI: (Khan, 2008; Khan, 2014)



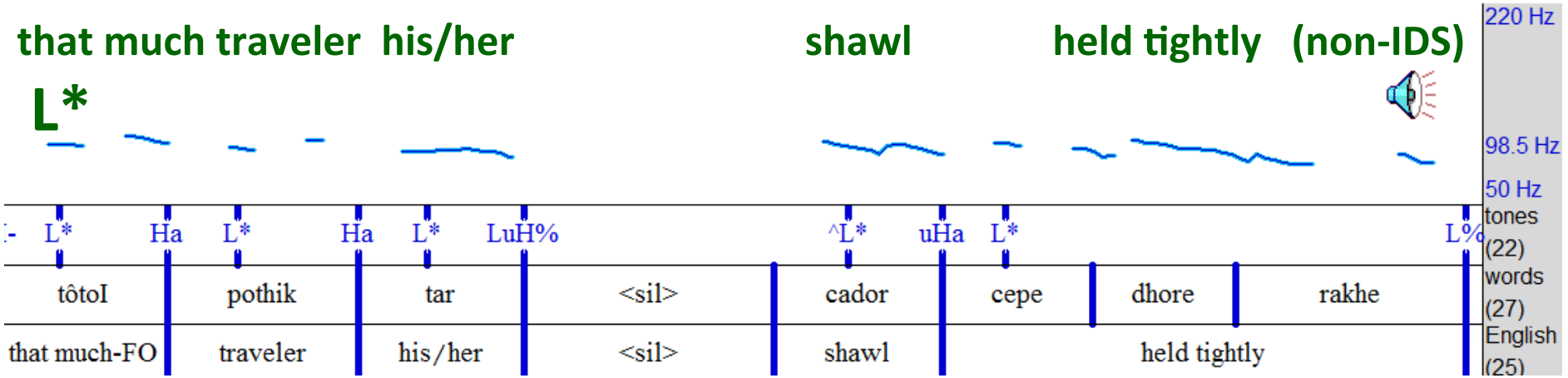
No stacking; only overrides at ip and IP boundaries

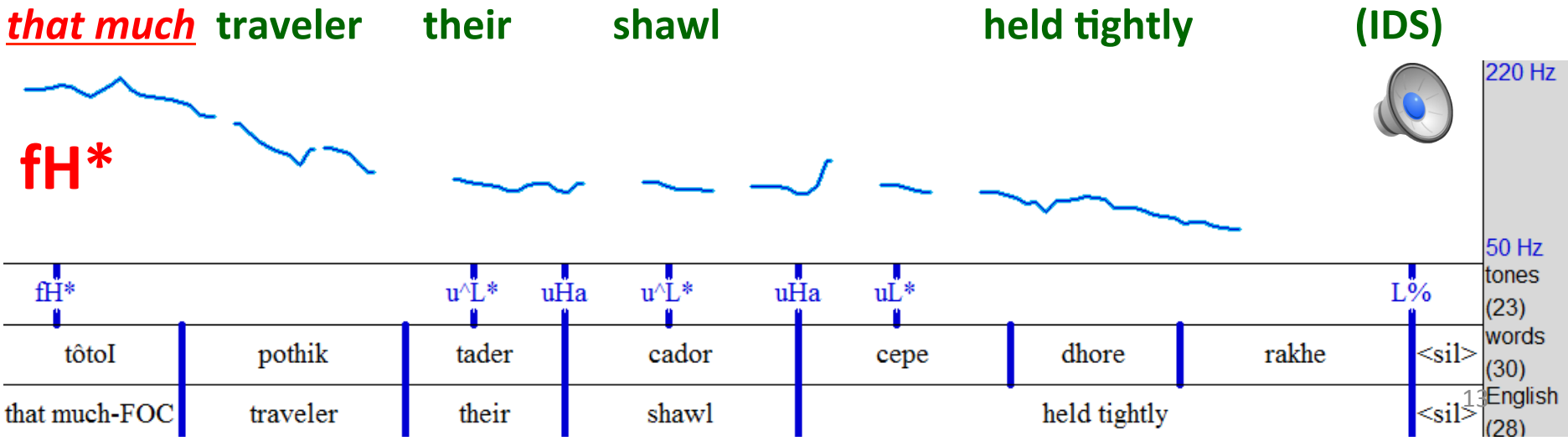
Grammar of licit sequences of tonal elements proposed in the B-ToBI model of Bengali intonation, ignoring the f-marked tones associated with focus as well as the effects of concurrent boundary tone overriding. Sequences can only begin at states with a thickly outlined circle ("AP_{Ha}" and "AP_{La}") and terminate at states with a double outlined circle ("IP").

Pitch Accents

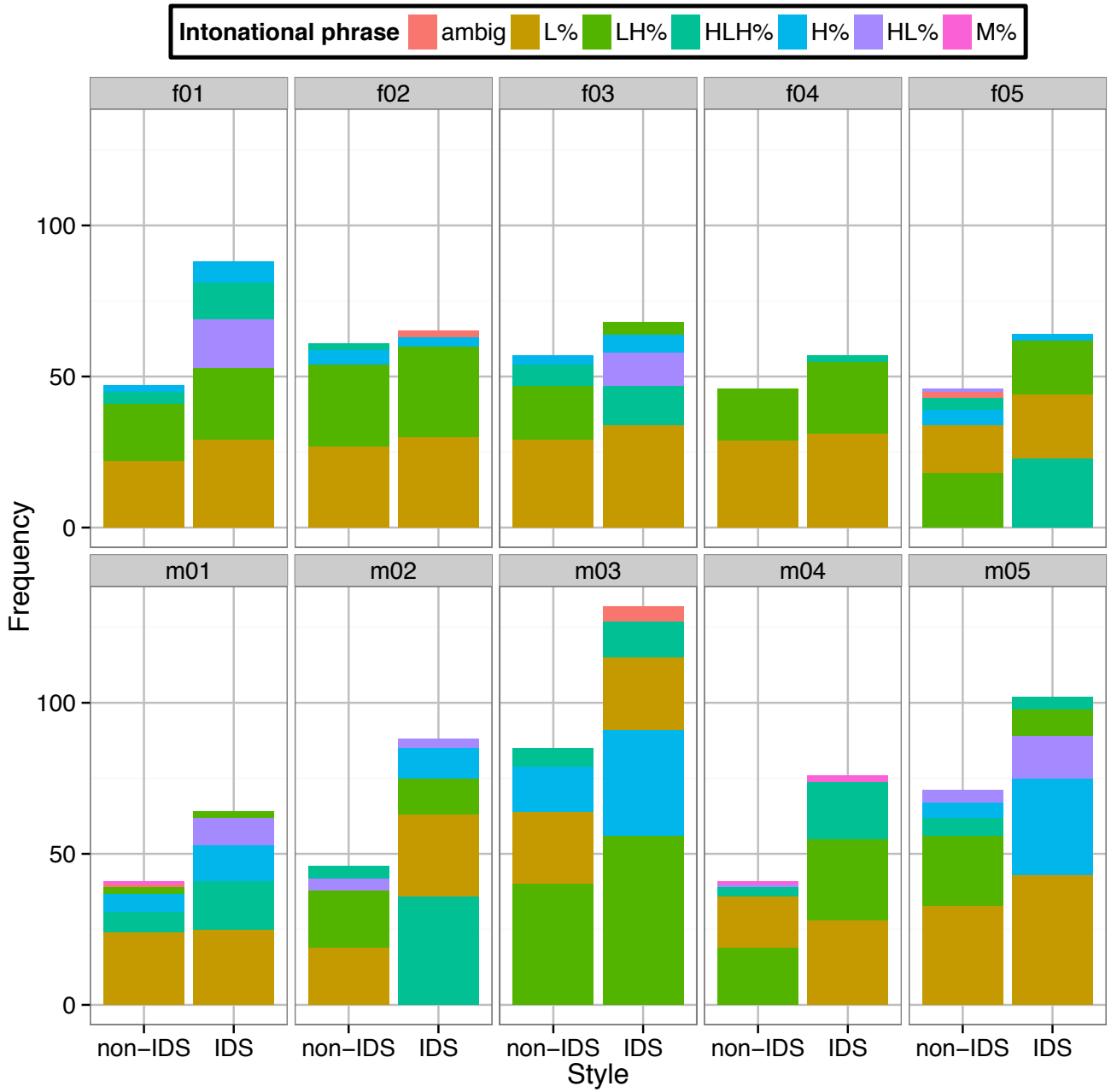


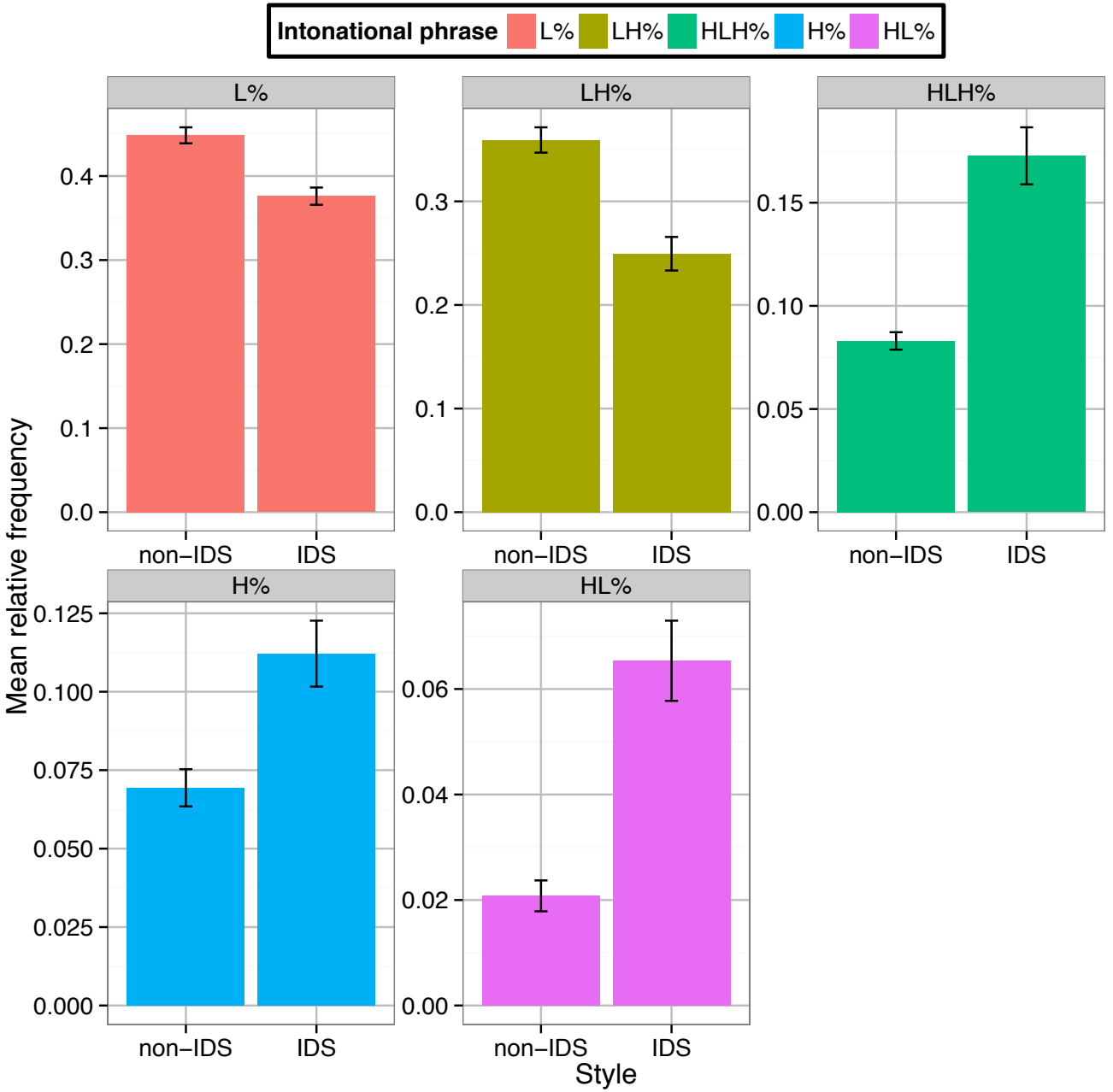
Bengali pitch accents



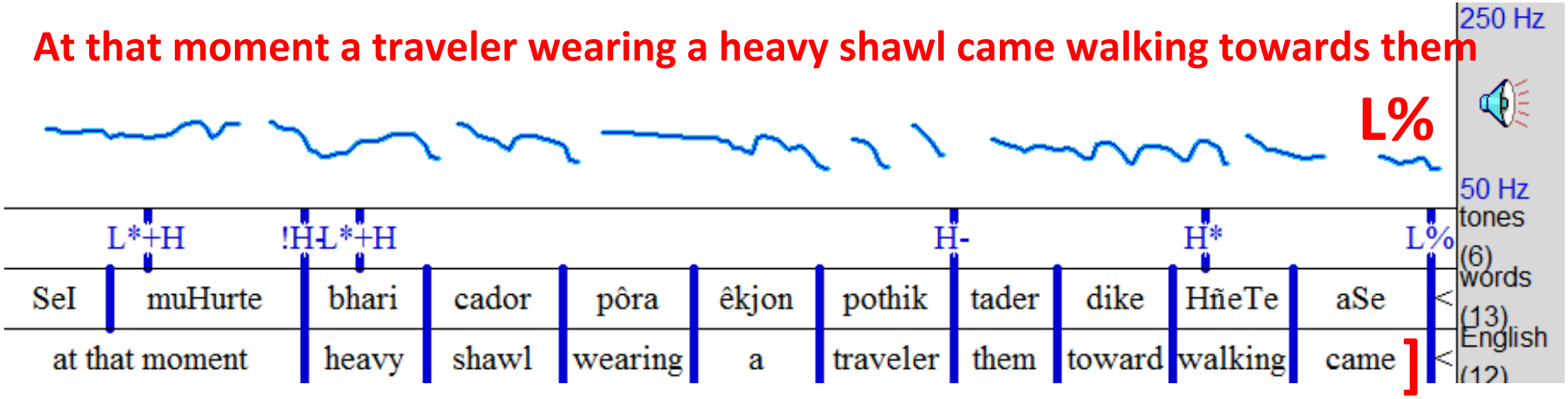


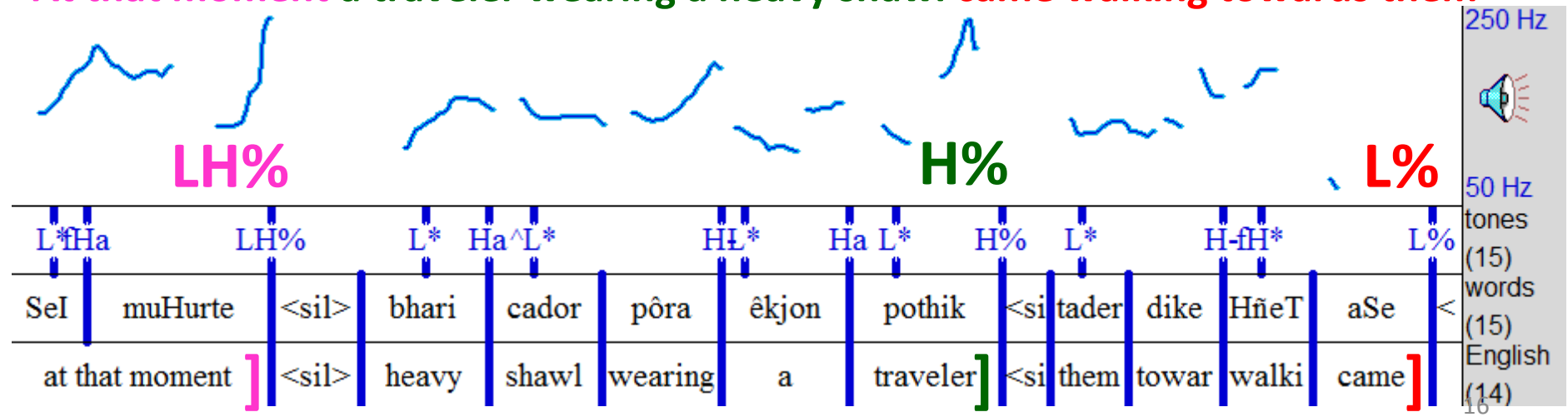
Boundary Tones




Intonational phrase
L%
LH%
HLH%
H%
HL%
L%
LH%
HLH%
H%
HL%
Mean relative frequency
0.0
0.1
0.2
0.3
0.4
0.00
0.05
0.10
0.15
0.000
0.025
0.050
0.075
0.100
0.125
0.00
0.02
0.04
0.06
non–IDS
IDS
Style

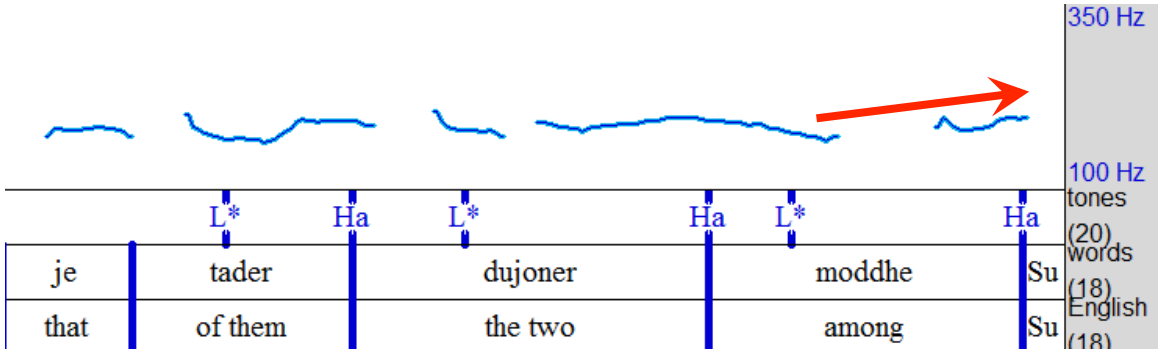
Bengali IPs


At that moment a traveler wearing a heavy shawl came walking towards them
L%
250 Hz
50 Hz
L*+H
!H-L*+H
H-
H*
L%
tones
(6)
SeI
muHurte
bhari
cador
pôra
êkjon
pothik
tader
dike
HñeTe
aSe
words
(13)
at that moment
heavy
shawl
wearing
a
traveler
them
toward
walking
came
English
(12)

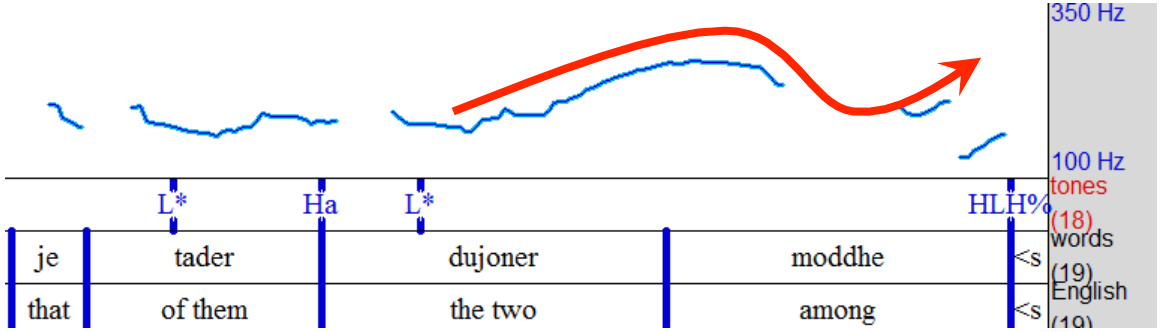

At that moment a traveler wearing a heavy shawl came walking towards them
LH%
H%
L%
250 Hz
50 Hz
L*fHa
LH%
L*
Ha^L*
H-L*
Ha L*
H%
L*
H-fH*
L%
tones
(15)
SeI
muHurte
<sil>
bhari
cador
pôra
êkjon
pothik
<si
tader
dike
HñeT
aSe
words
(15)
at that moment
<sil>
heavy
shawl
wearing
a
traveler
<si
them
towar
walki
came
English
(14)

Bengali boundary tones

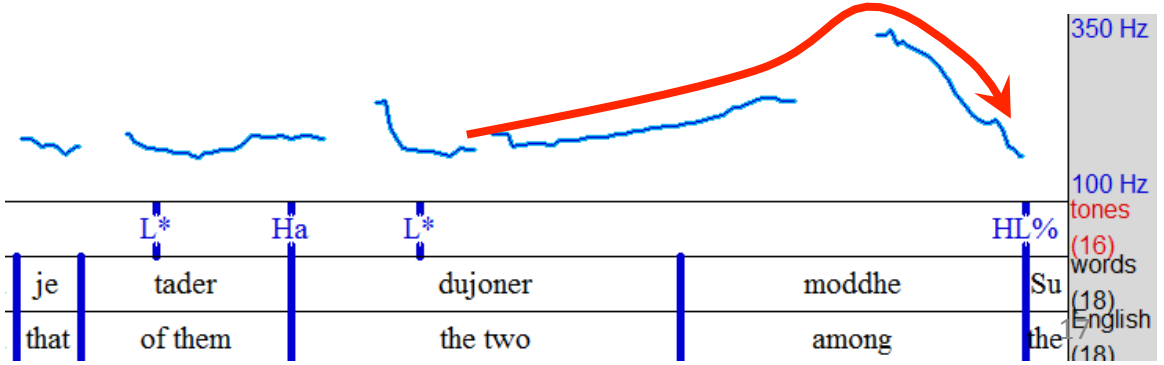
Ha in non-IDS
Default



HLH% in IDS
Continuation

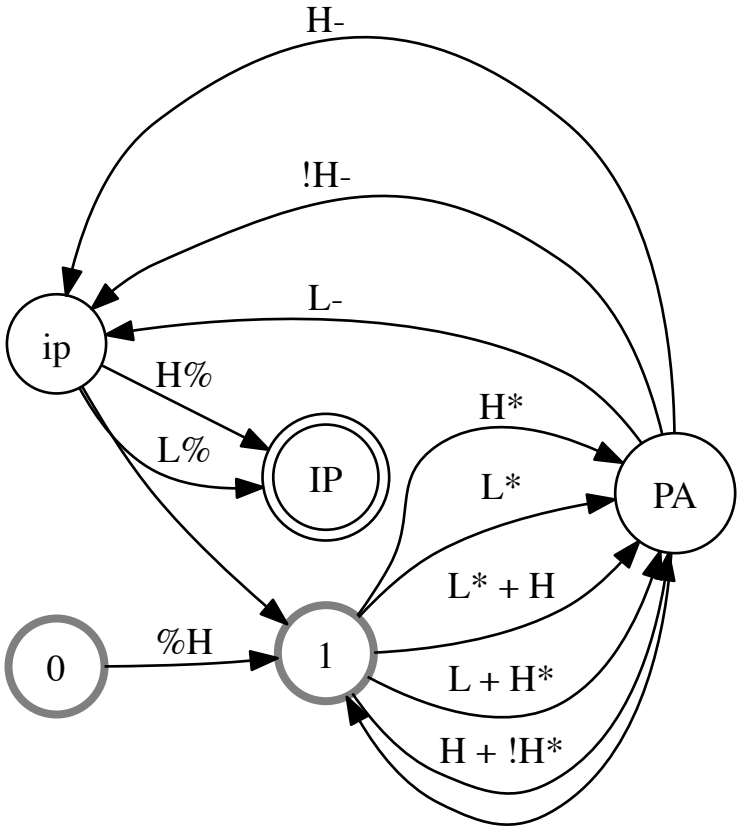


HL% in IDS
Topicalization



ENGLISH

MAE-ToBI: (Ladd, 1996; Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990)



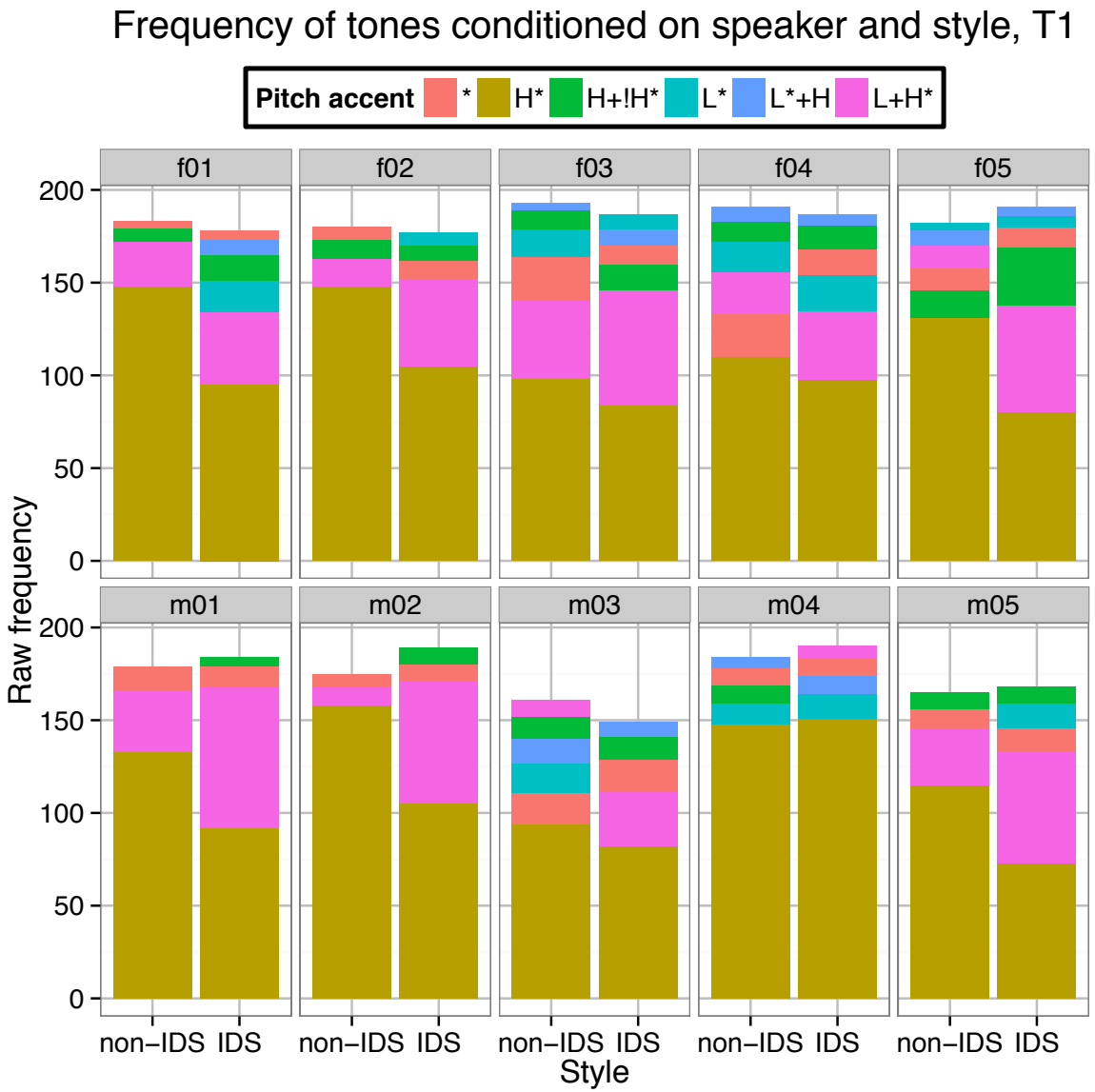
Stacking observed at ip and IP boundaries

Grammar of licit sequences of tonal elements proposed in the MAE_ToBI model of American English intonation. Sequences can only begin at states with a thickly outlined circle ("0", "1") and terminate at states with a double outlined circle ("IP").

Comparison of Bengali and English

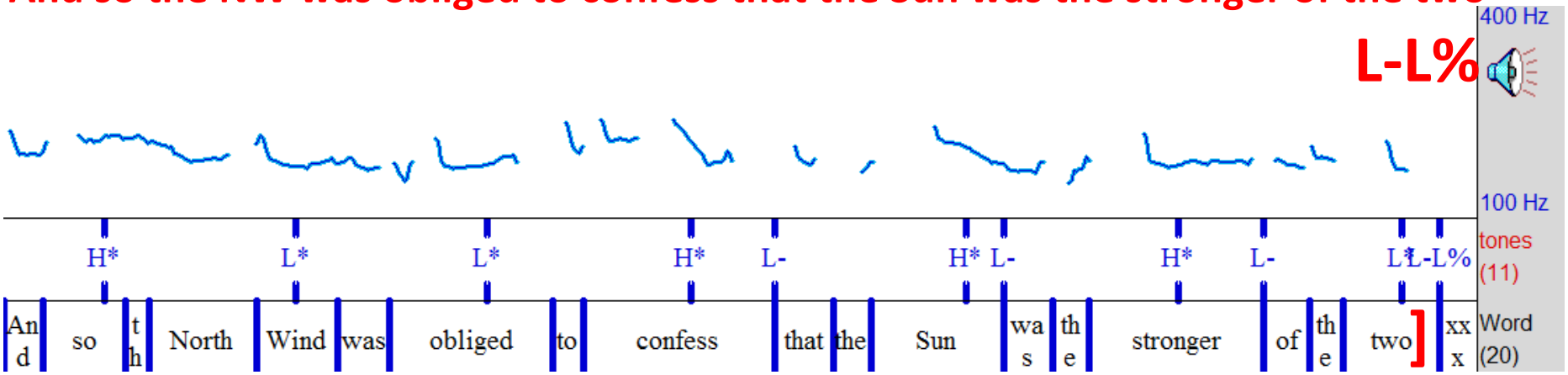
- Bengali not English has a regular repeating pattern
  - Due to restricted PA and AP boundary inventory
- Both language mark focus with PA choice and post-focal de-accenting
- Bengali has 3 – AP, ip, IP (compared to 2 - ip and IP in English) tonally-marked prosodic structures
  - More complex boundary tones involving pitch changes for ip and even more for IP

Pitch Accents

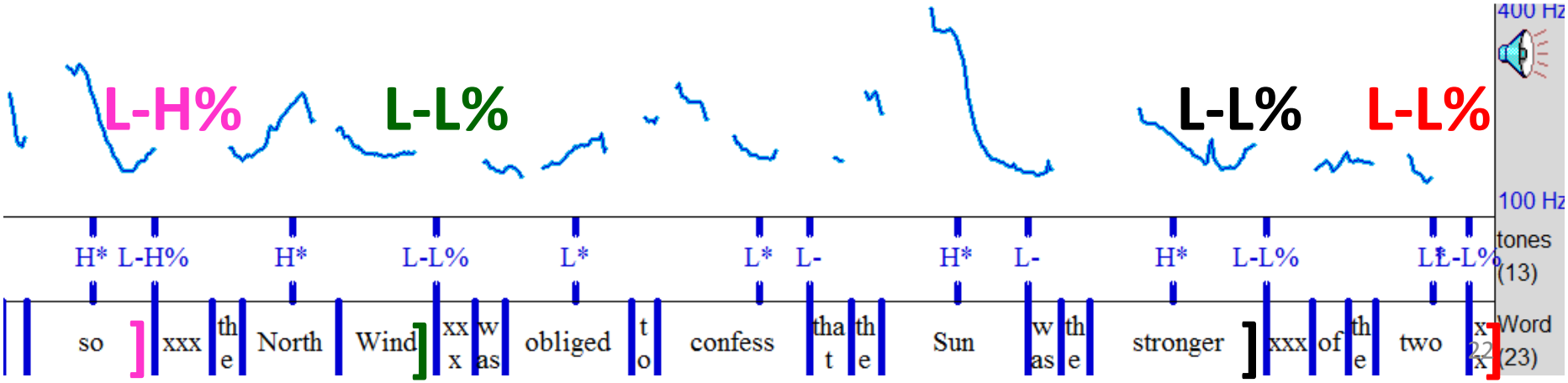


English: IPs

And so the NW was obliged to confess that the Sun was the stronger of the two



And so the NW was obliged to confess that the Sun was the stronger of the two



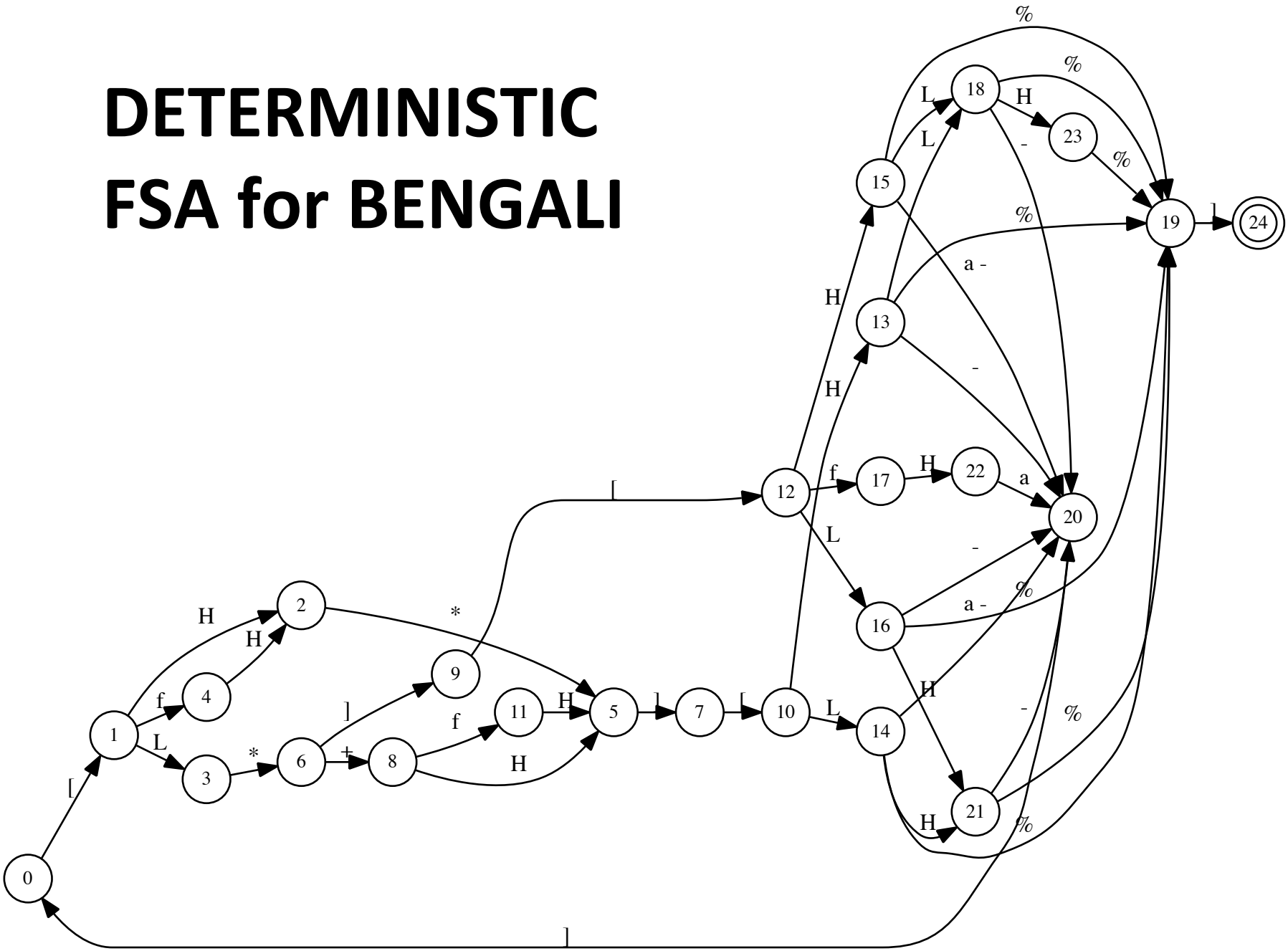
Comparing intonational grammars in ADS and IDS

- Construct a finite state grammar for intonational tonotactics, common across ADS/ IDS
- Estimate weights over transitions in finite state acceptor using production corpus frequencies
- Compare weights between ADS and IDS

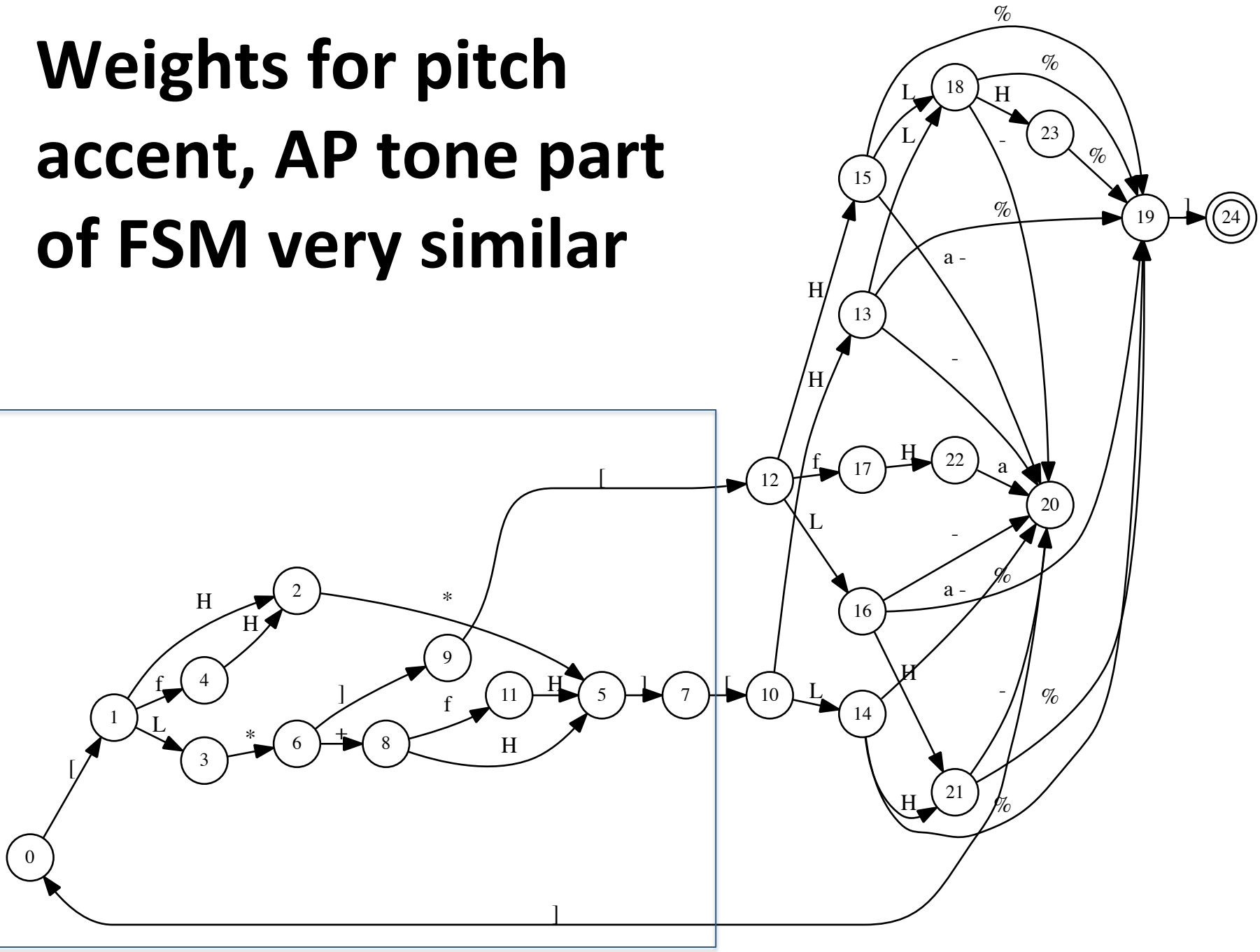
Defining a finite state acceptor for Bengali intonation

- Use `xfst` (Beesley and Karttunen 2003): define "extended" regular grammars with high-level language (regular expressions)
- Example definition:
  - `define ipSequence [ [ APSequence ]* ipEdge ];`
  - " An ip tone sequence consists of zero or more AP tone sequences followed by an "ipEdge" (PA + ip tone)."
- Construct deterministic finite state acceptor
  - *Deterministic:* given input and current state, there is only one possible outgoing path from the current state

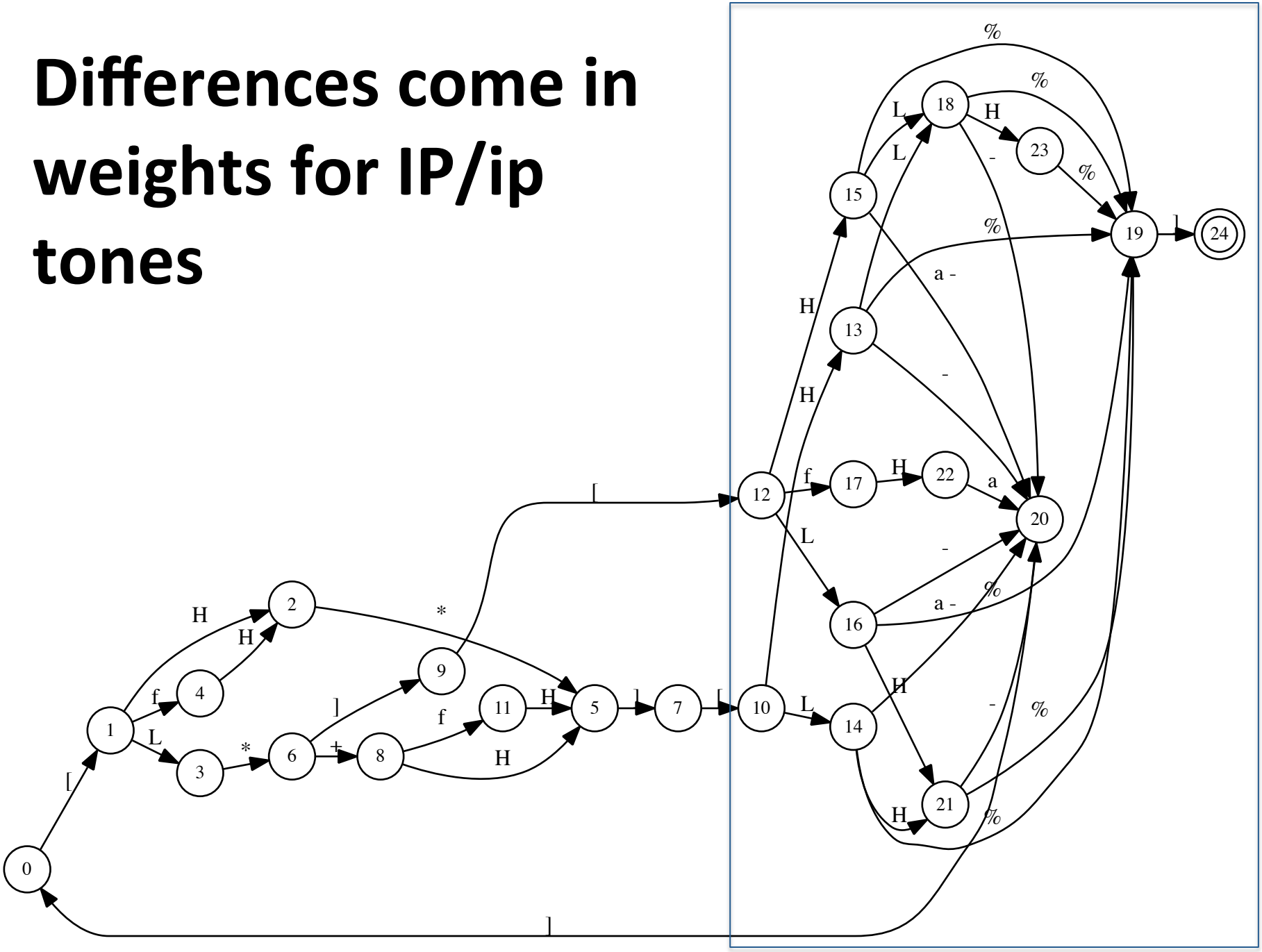
DETERMINISTIC FSA for BENGALI



Weights for pitch accent, AP tone part of FSM very similar



Differences come in weights for IP/ip tones



Patterns in arc weight differences

- At choice point for ip vs. IP tone, IDS goes with an IP tone more frequently
- At a choice point for an IP tone to include more tonal targets or not, e.g. L**%** vs. L**H**%, HL **%** vs. HL**H**%, IDS goes with more tonal targets more frequently