In Search of the Phrase Accent:
Nuclear Tunes in Mainstream American English

Jennifer Cole & Thomas Sostarics

Northwestern University

Nuclear Tune in English

Pitch trajectory at the end of the intonational phrase:

[Only **Me**lanie ran a mile]

H*   L%

Pitch accent   *Boundary tone*

- Pitch accents lend **prominence** to a stressed syllable
- Boundary tones mark the right edge of an **intonational phrase** (IP)
- *What happens in between?*

Nuclear Tune in English

[Only <u>Melanie ran a mile</u>]

H* ←——— L- ———→ L%

In the Autosegmental-Metrical (AM) model, the *phrase accent* (L-) *s*pans the middle nuclear region

The phrase accent marks the right **edge** of the *intermediate phrase,* a lower level of prosodic phrase structure. It anchors to the final syllable in the (little) ip and spreads leftward to the post-accentual syllable:
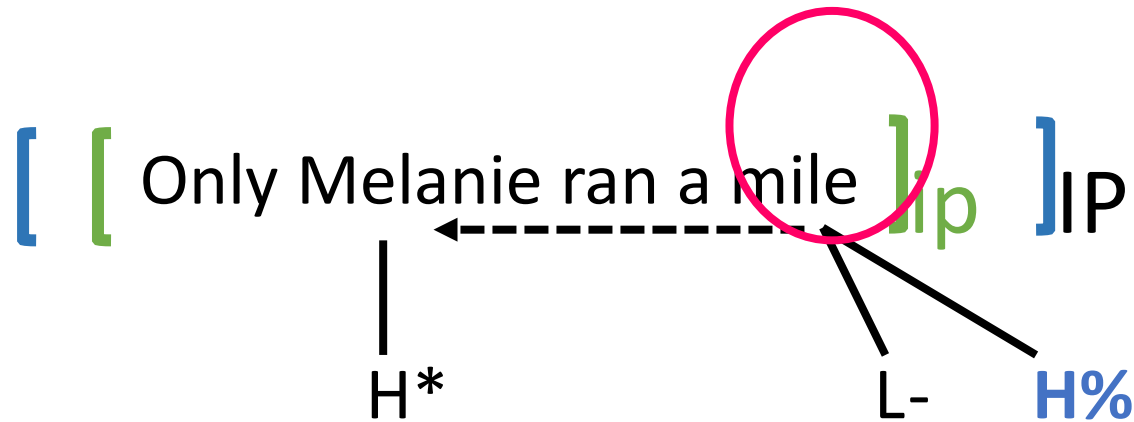
[[Only Melanie ran a mile]ip]IP

H*   L-
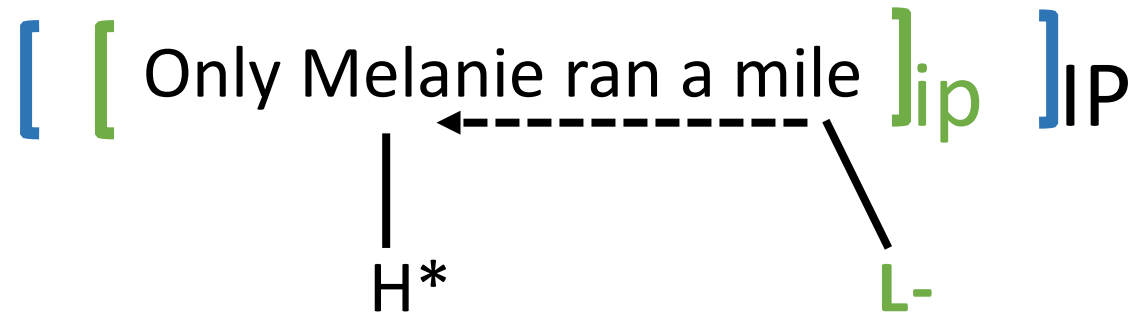
The leftward association of the phrase accent accounts for the fall immediately following the pitch accent, and the sustained low up to the phrase-final syllable.
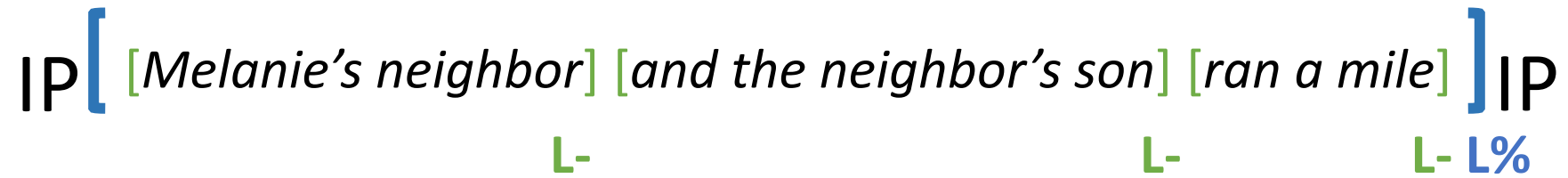
The boundary tone also anchors to the final syllable in the IP.
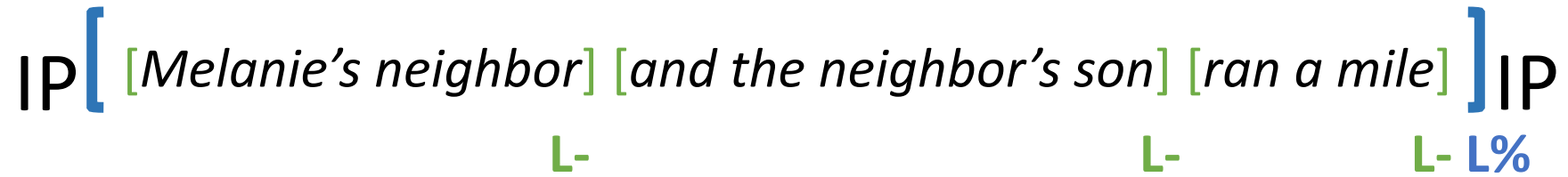Two edge tones anchor to the **same** syllable.

[[Only Melanie ran a mile]ip]IP

H*

L-   L%
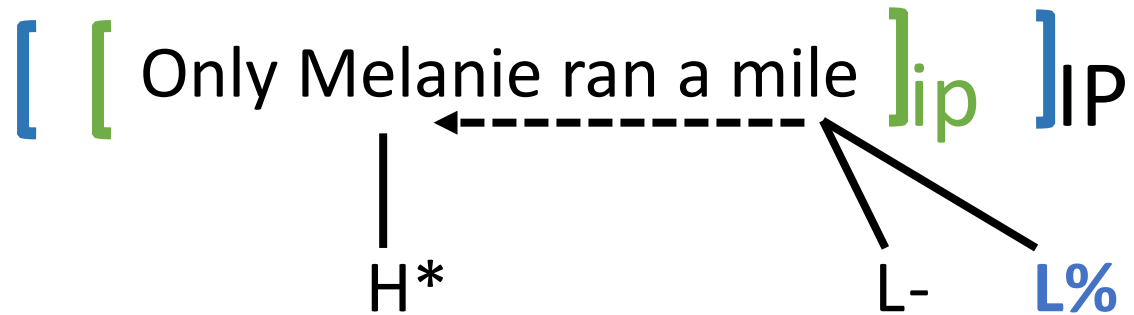
The phrase accent and boundary tone can be specified with different tones. In the case where H* is followed by L-H% this yields a complex "rise-fall-rise" (RFR) pitch trajectory.
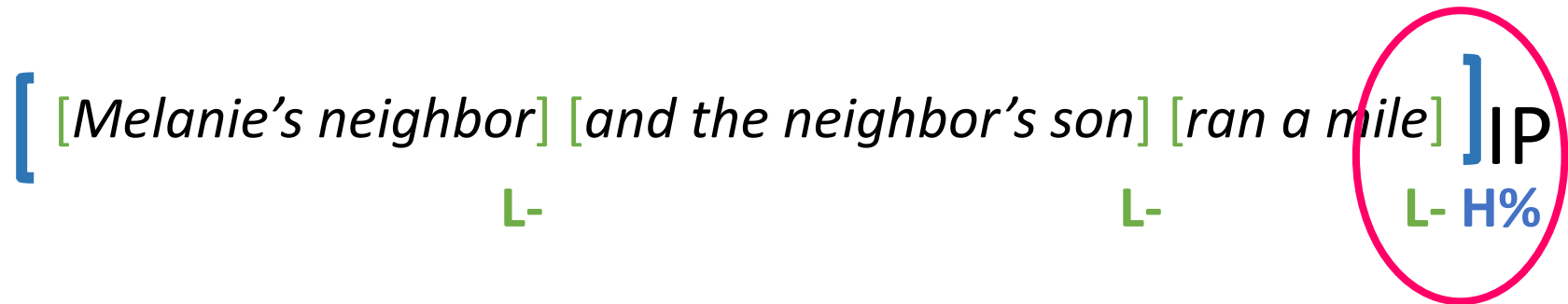
The (big) IP may consist of one or more (little) ip's:

IP[[*Melanie's neighbor*] [*and the neighbor's son*] [*ran a mile*]]IP

                    L-   L-   L- L%

The (big) IP may consist of one or multiple (little) ip's:

IP[[*Melanie's neighbor*] [*and the neighbor's son*] [*ran a mile*]]IP
         L-   L-   L- L%

But only the **final** (little) ip in the (big) IP will be specified for a boundary tone:

[[*Melanie's neighbor*] [*and the neighbor's son*] [*ran a mile*]]IP
         L-   L-   L- H%

The phrase level controversy

- The analysis of the phrase accent rests on the problematic assumption that there are two levels of prosodic phrasing : ip, IP

- Inter-annotator agreement on this level distinction is poor
  - they can have the same dynamic pitch (falling, rising)
  - durational effects of phrase edge are gradient: final lengthening, pause

- An alternative account assumes one level of prosodic phrasing, possibly marked by a tone sequence, e.g. LH% (Gussenhoven 2004)

An alternative perceptual account Barnes et al. 2010

The fall after the accentual peak is due to **postnuclear deaccentuation:**

avoid high targets following a focus-marking nuclear pitch accent

Only [Melanie]_{FOC} ran a mile

(L+)H*   L-H%
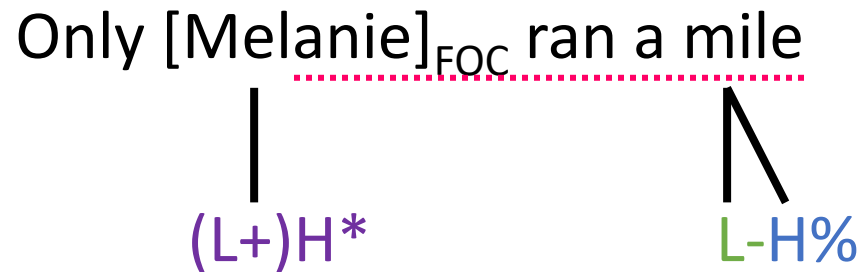
In this analysis, the phrase accent **(L-)** can be locally associated to the final syllable of the ip.
Just like the boundary tone **(H%)**

Other tunes?

The predictions of the perceptual deaccenting analysis are not clear for the middle pitch in other tunes, e.g.

? —

Only Melanie ran a mile

H*   H-H%

Here also, the pitch accents are marking focus due to the focus-sensitive operator "Only".

? —

Only Melanie ran a mile?

L*   H-H%

But the same pitch accents may be used without focus: *Melanie ran a mile*

Filling the empirical middle gap

Goal: examine F0 trajectories in the middle region

Is there evidence of a **tonal target** (the phrase accent) between the pitch accent and boundary tone?

Does the middle region always exhibit **deaccenting**, with low pitch following a focus-marking pitch accent?

Is F0 **interpolated** between the targets of the pitch accent and boundary tone?

Tune imitation experiment

Two model utterances with the same text, M & F model speakers,
with F0 resynthesized (shown schematically here):

Only <u>Damian dined</u>

Only <u>Damian dined</u>
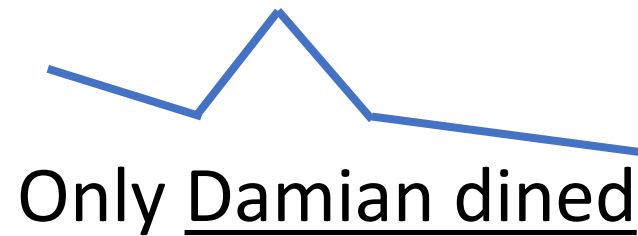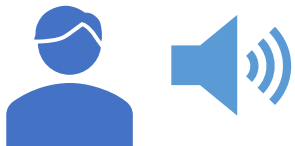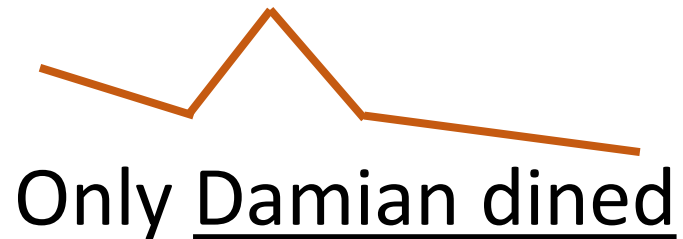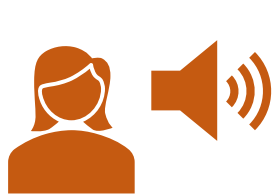
Target sentence

Only Madelyn ran

Steffman, Cole & Shattuck-Hufnagel, 2024
Steffmen & Cole, 2024
Cole, Steffman, Shattuck-Hufnagel & Tilsen, 2023

Methods

37 American English monolinguals from Prolific, 24 analyzed here

144 trials, crossing 3 tunes x 4 trajectories x 2 lengths

F0 tracking errors removed
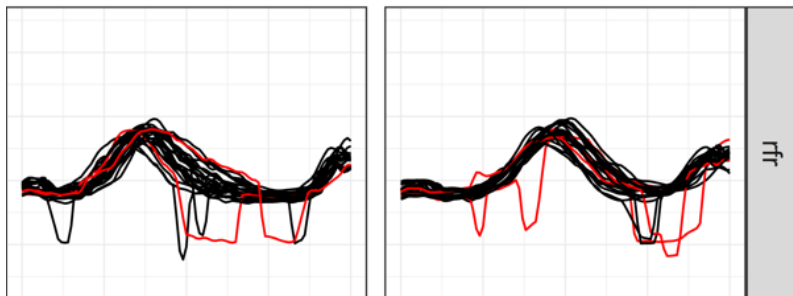
Steffman & Cole 2023

Falls

H*L-L%

Rise-Fall-Rises

L*H L-H%

Rises

L*H-H%

14
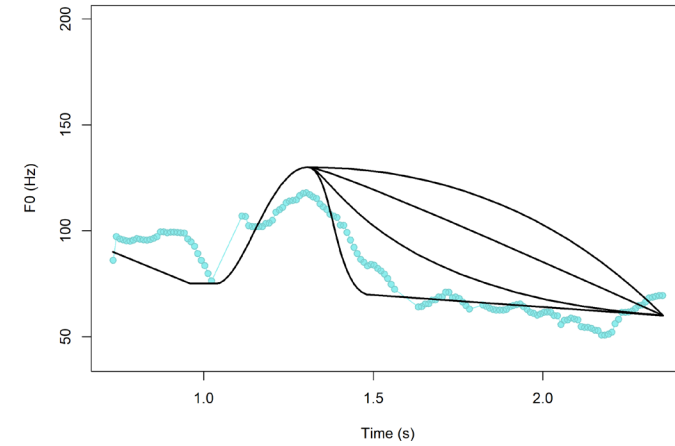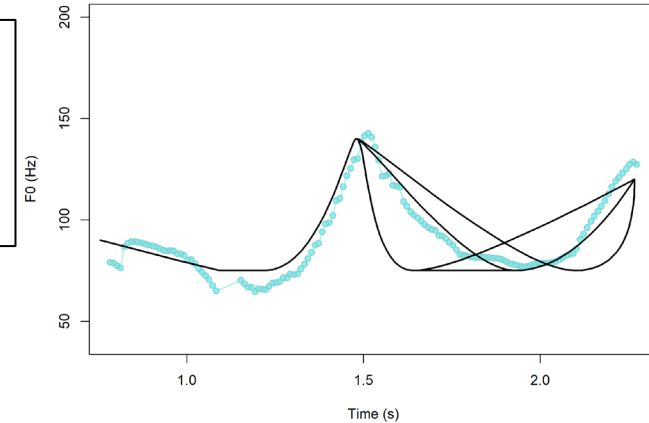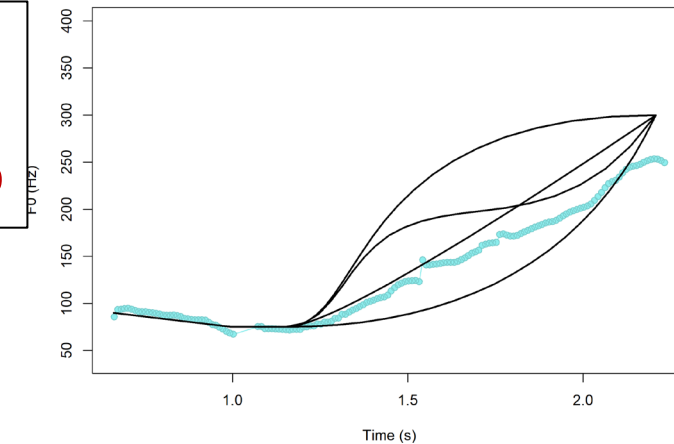
Methods

Short target sentences have 4 syllables in nuclear interval:

*Only **Dam**ian rode*
*Only **Ol**iver dined*

Long target sentences have 6:

*Only **Dam**ian rode away*
*Only **Ol**iver dined alone*

Falls
H*L-L%
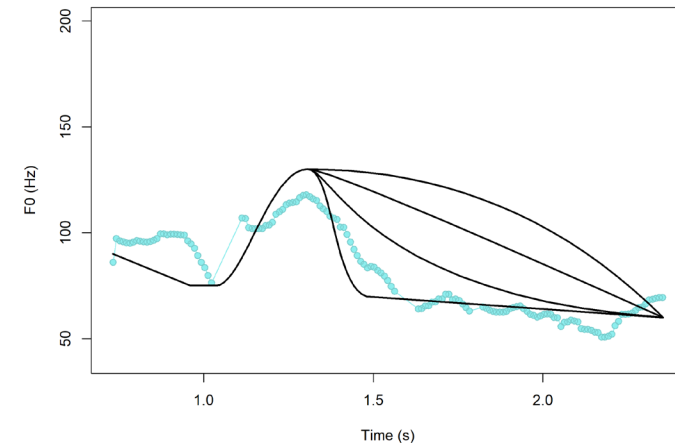
Rise-Fall-Rises
L*H L-H%
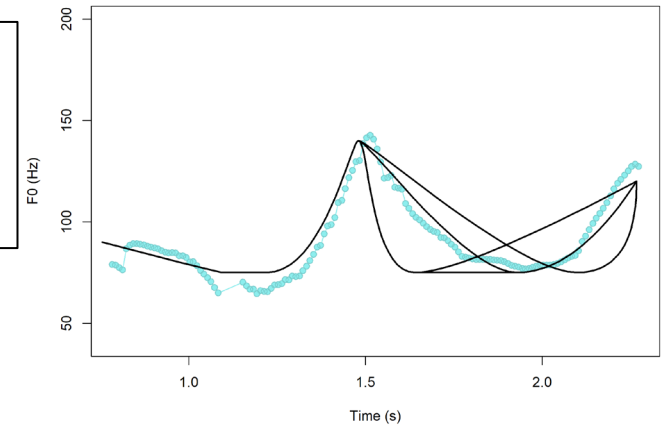
Rises
L*H-H%
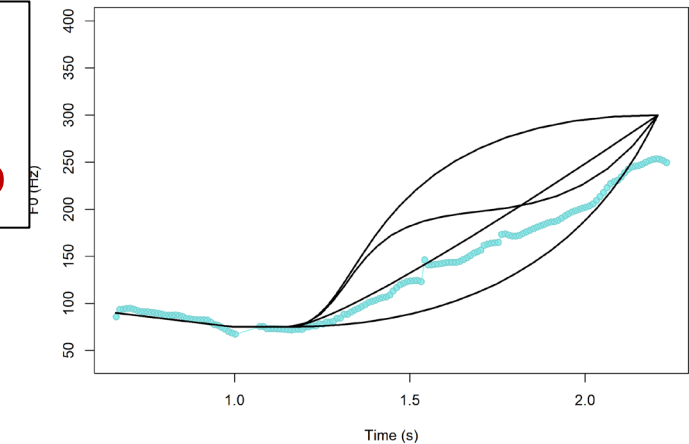
Auditory stimuli

Resynthesis is specified via Bezier curves: discrete (time, F0) targets are specified, where control points affect the curvature between one endpoint and another
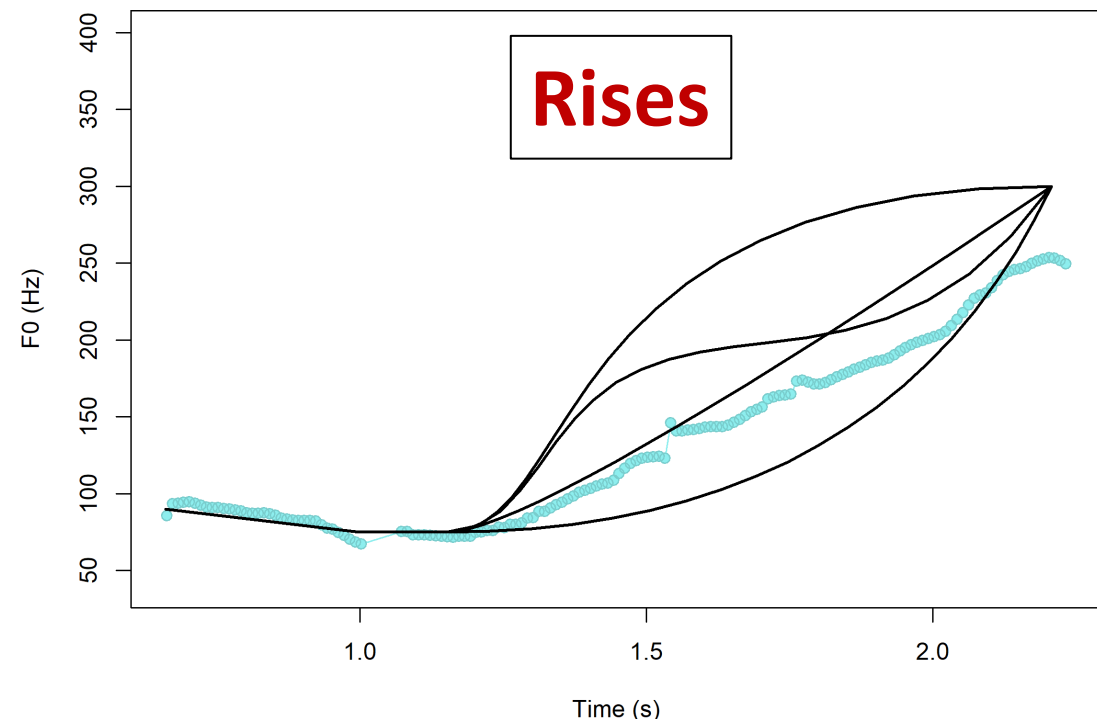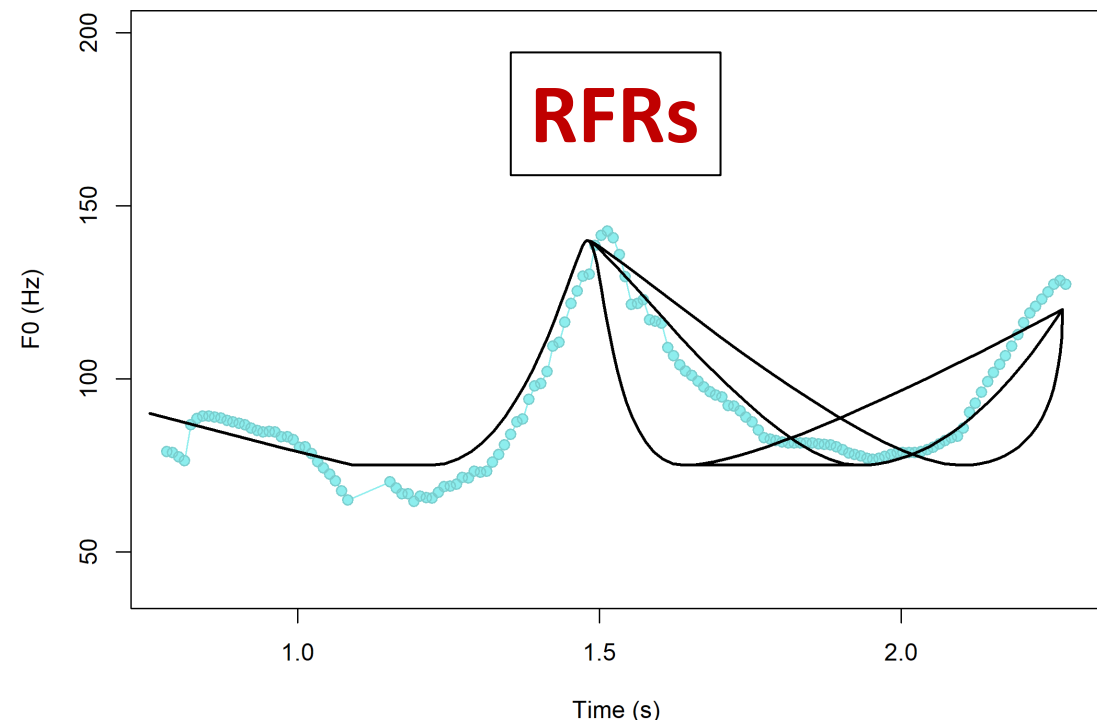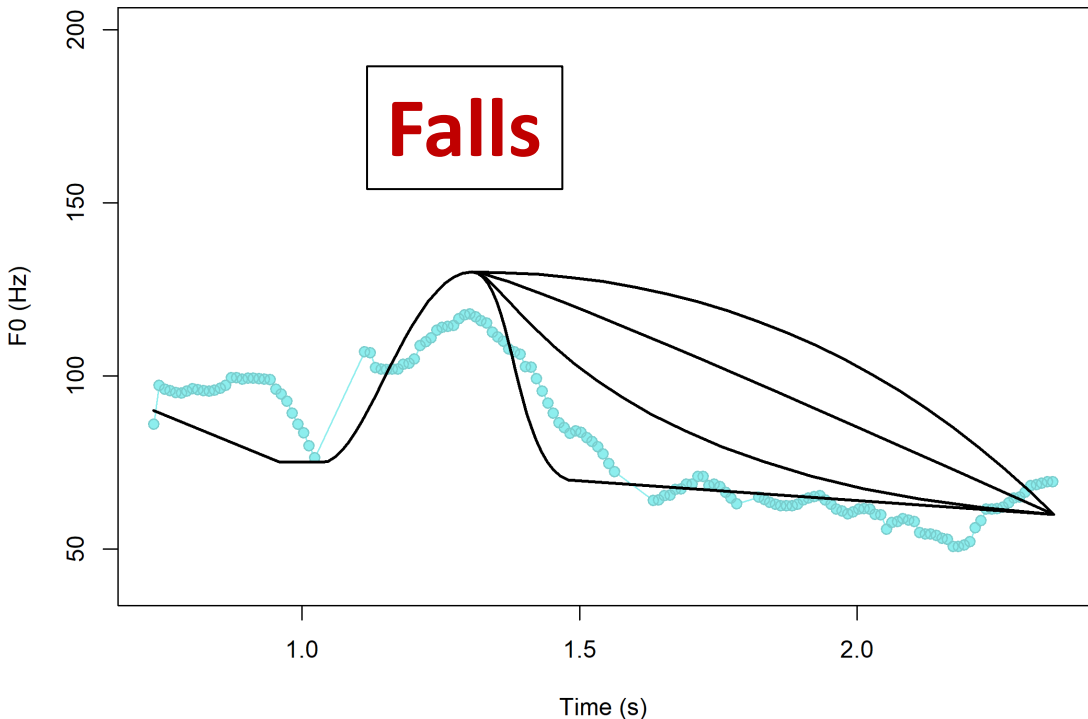
Falls
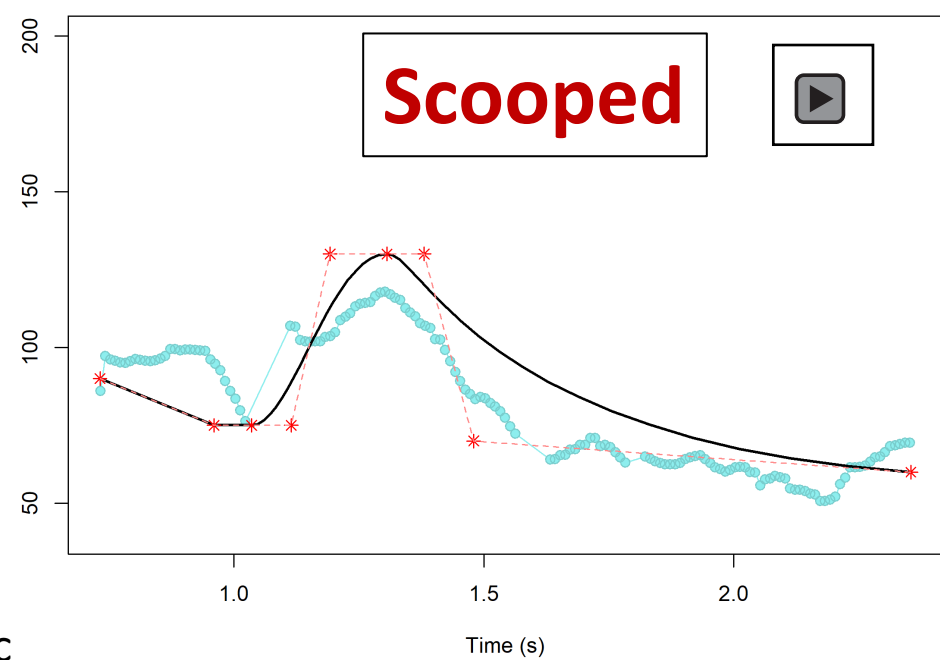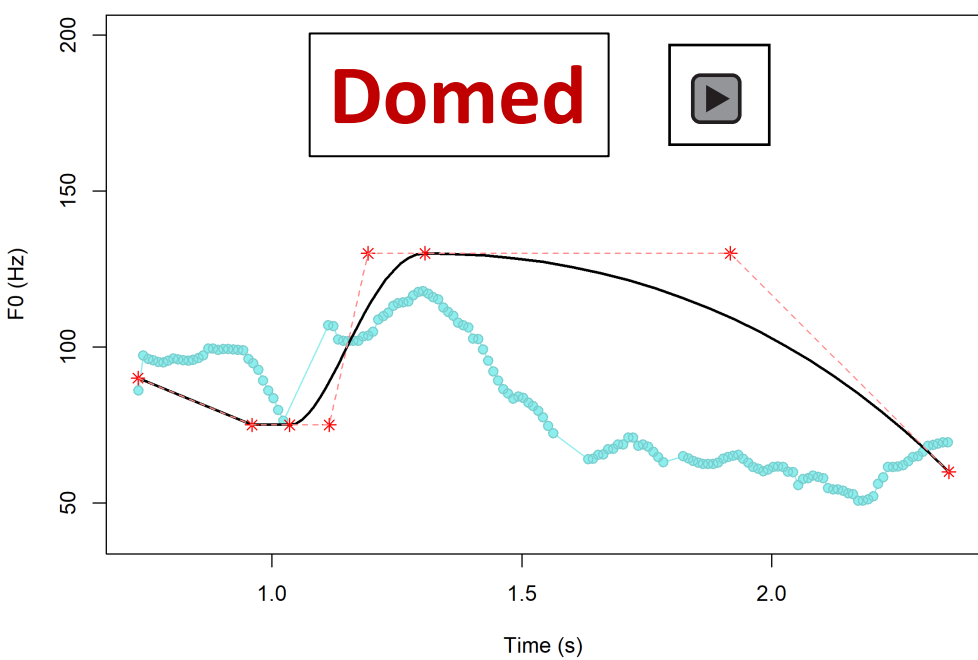
Blue: Original F0 contour
Black: Resynthesis schematic
Red: Bezier control points
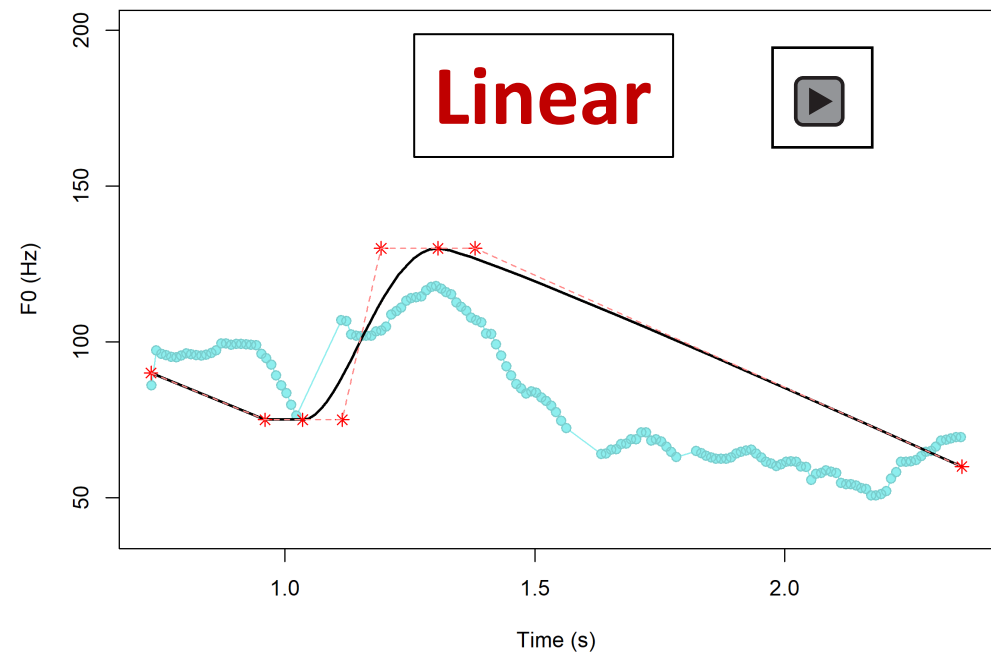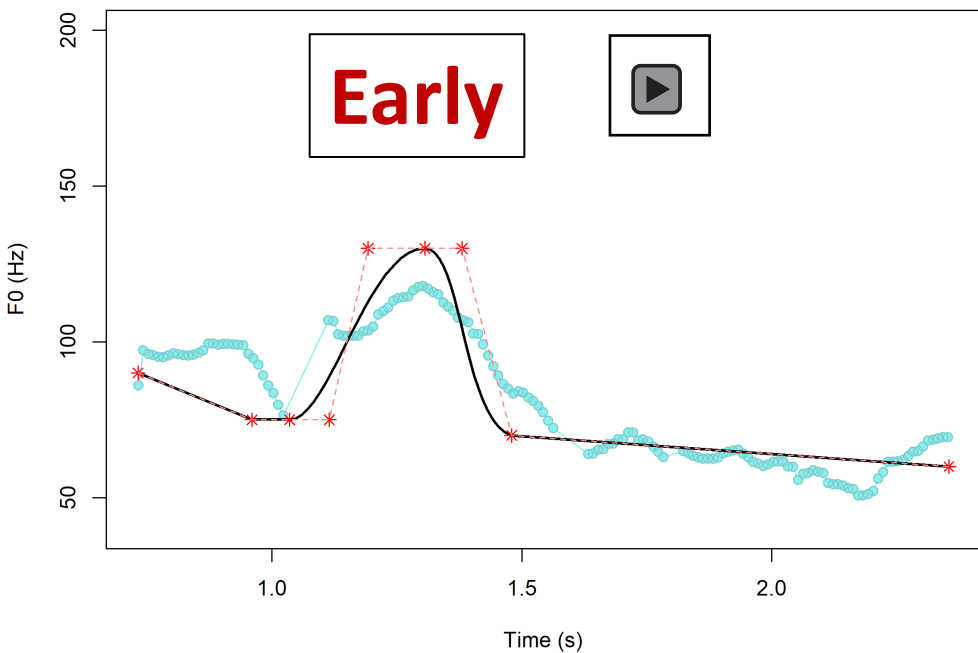
Rises
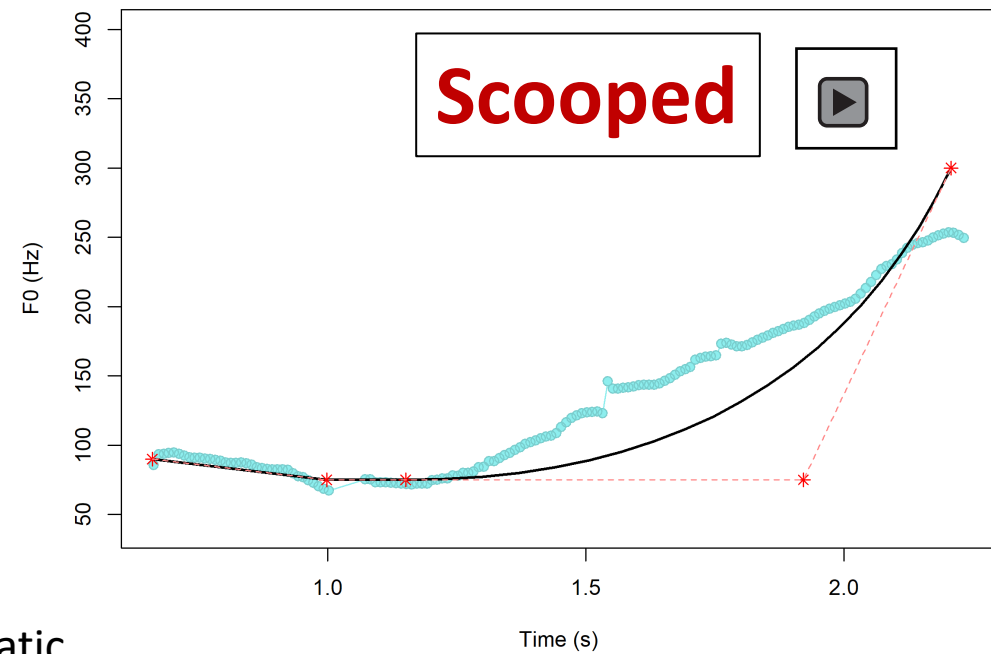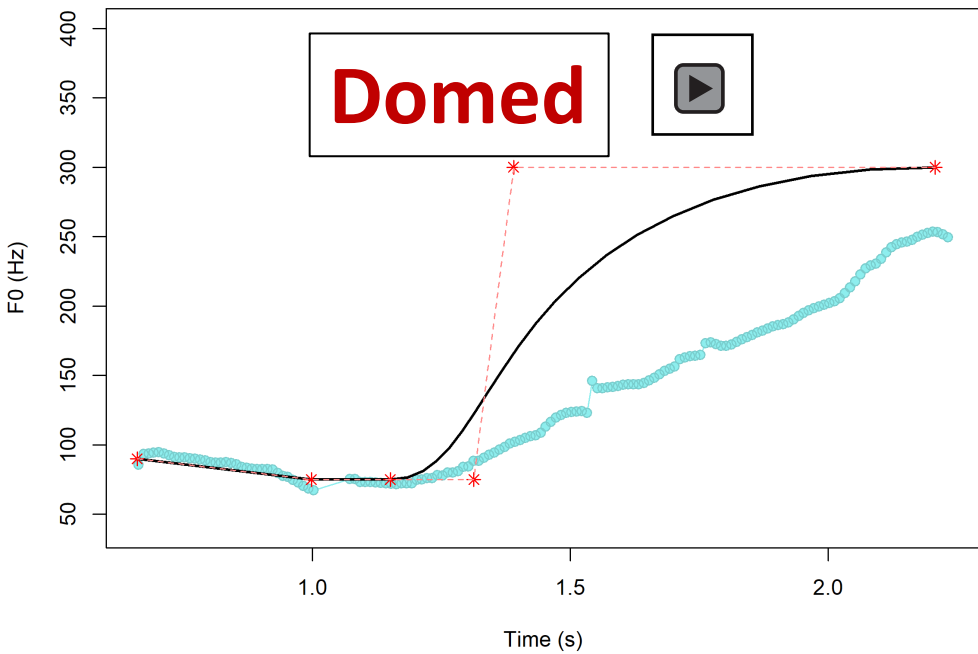
Blue: Original F0 contour
Black: Resynthesis schematic
Red: Bezier control points
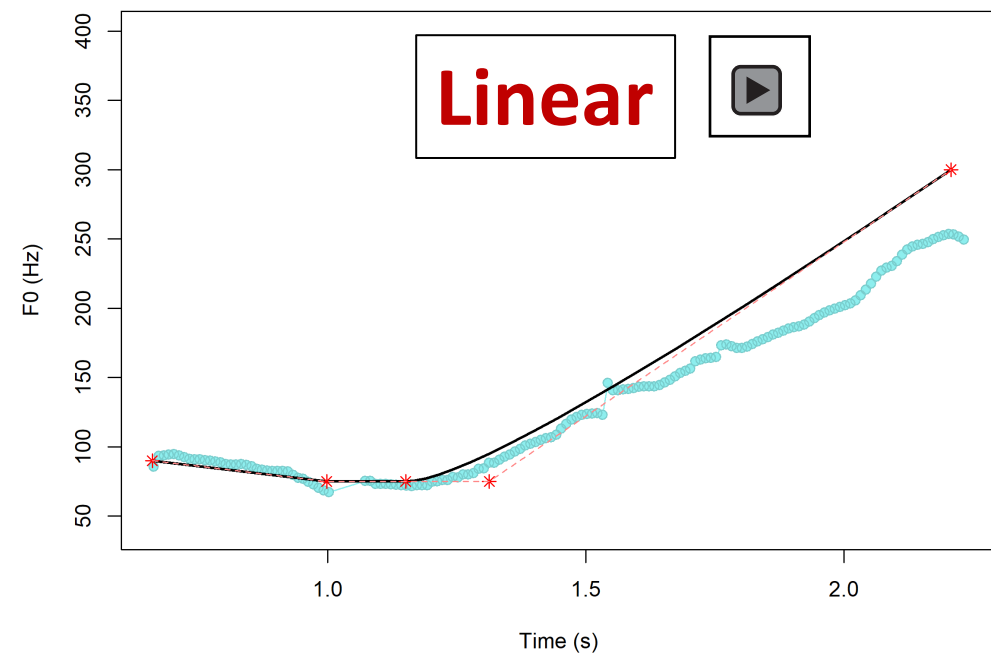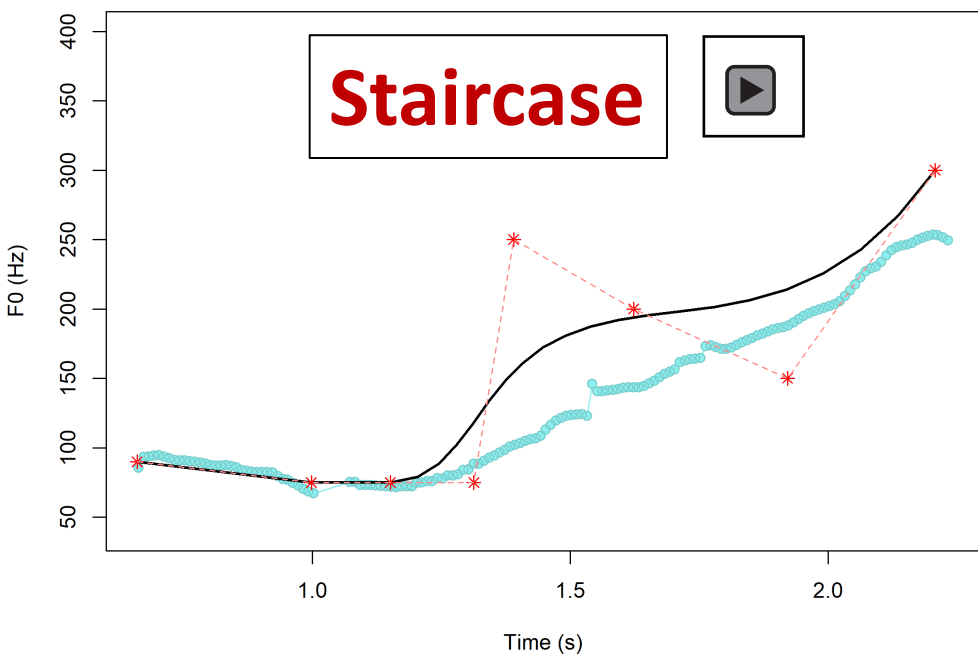
RFRs
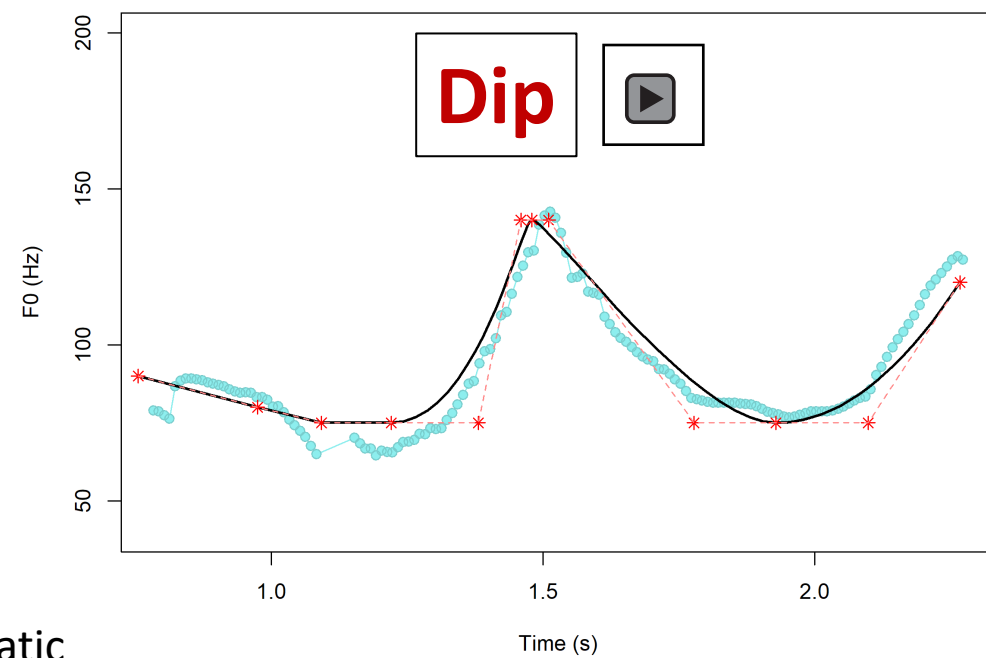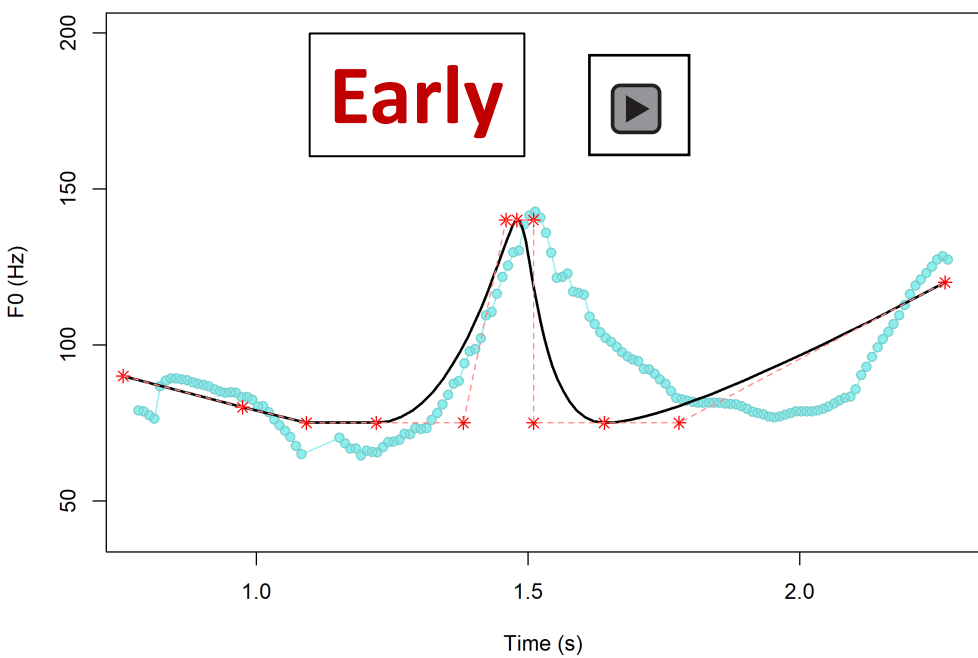
Blue: Original F0 contour
Black: Resynthesis schematic
Red: Bezier control points
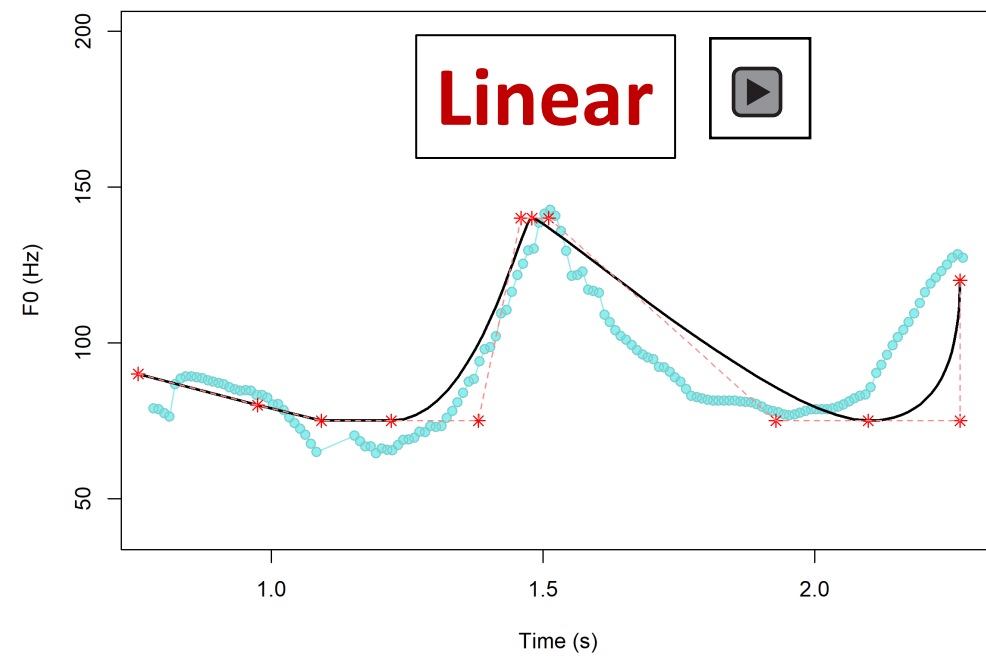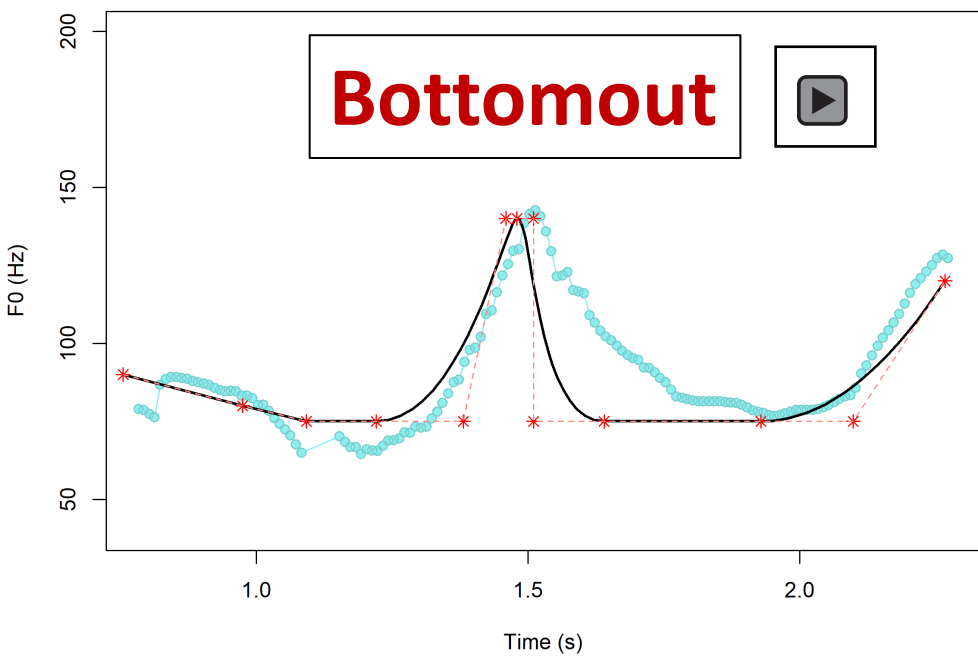
Example trial

Two context sentences are presented in text format, providing the discourse context for a focused word in the response

Did Oliver and Damian row with Jessie?

Did Harmony and Madelyn move with Jessie?

You'll first see two context sentences

Next

Example trial

Participants are instructed to pay attention to the melody of the next turn in the dialogue, presented for the discourse prompt on top

Did Oliver and Damian row with Jessie?

Did Harmor... ...ith Jessie?

You'll hear two computerized speech samples respond to the first sentence. Pay attention to the melody that's used.

Previous   Next

Example trial

Play then fade out ⟶

Participants hear two versions of the response that differ only in the model speaker (M & F)

Did Oliver and Damian row with Jessie?

Only DAMIAN rowed...

Only OLIVER rowed...

Did Harmony and Madelyn move with Jessie?

Example trial

Participants read aloud the sentence that responds to the second question, using the same melody they just heard

Trial structure
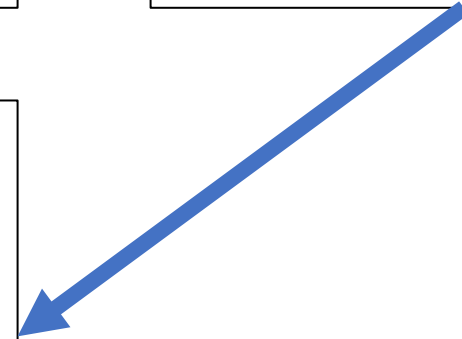
Empirical data

- F0 over the nuclear interval

    Only *Damian dined*

  - Files with >2 flagged F0 errors removed

  - Time normalized to 50 samples

- Speaker means for each trajectory in thin lines

- Grand means for each trajectory overlaid

- Color coded trajectory shape

F0 trajectories with no internal alignment;
Duration of tune-internal segments varies
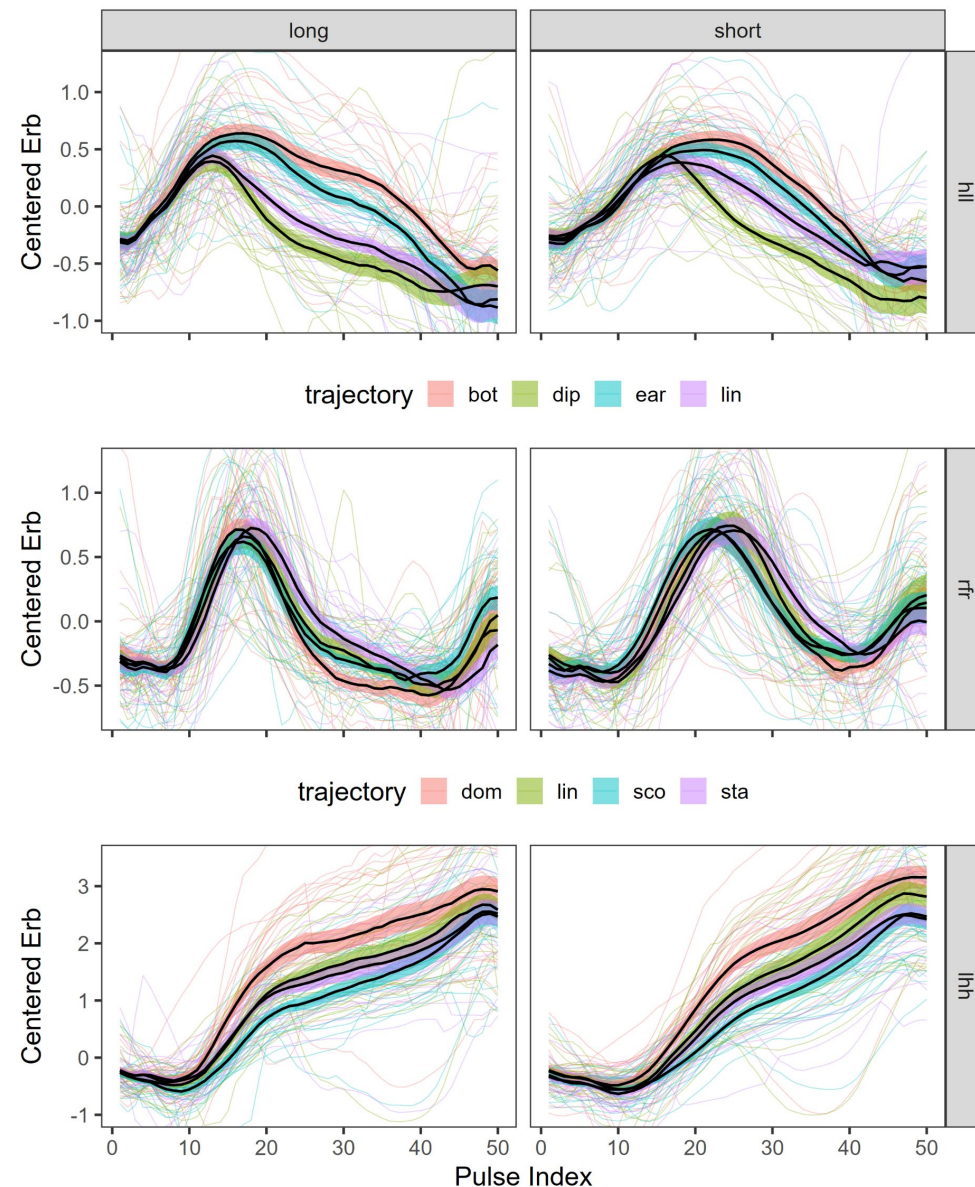
F0 trajectories aligned at the target for the nuclear accent
Duration of tune-internal segments (start to accentual
target; accentual target to end) is held constant

F0 trajectories with no internal alignment
Duration of tune-internal segments varies
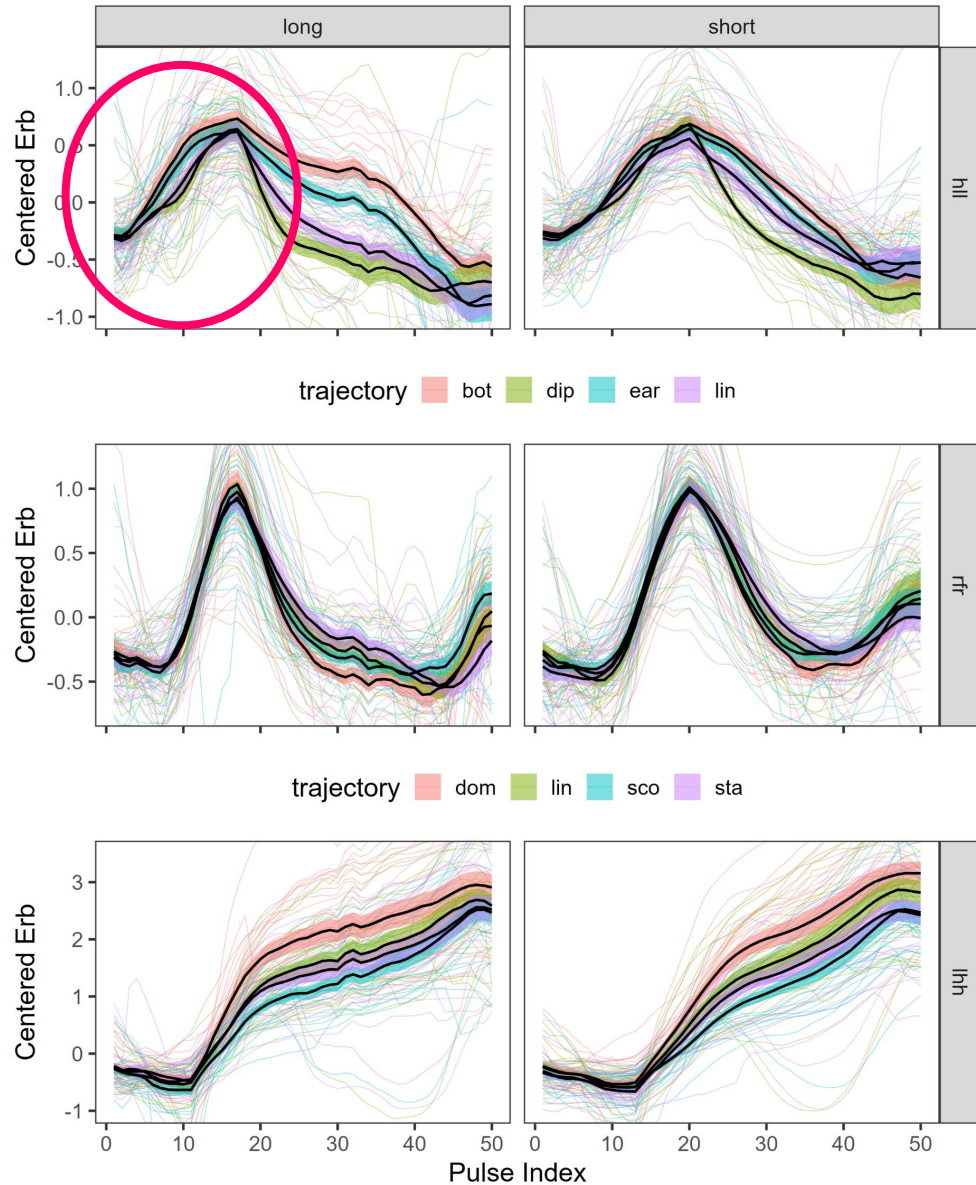
Two ways of
viewing and
analyzing F0
trajectories
that vary in
the location
of F0 targets
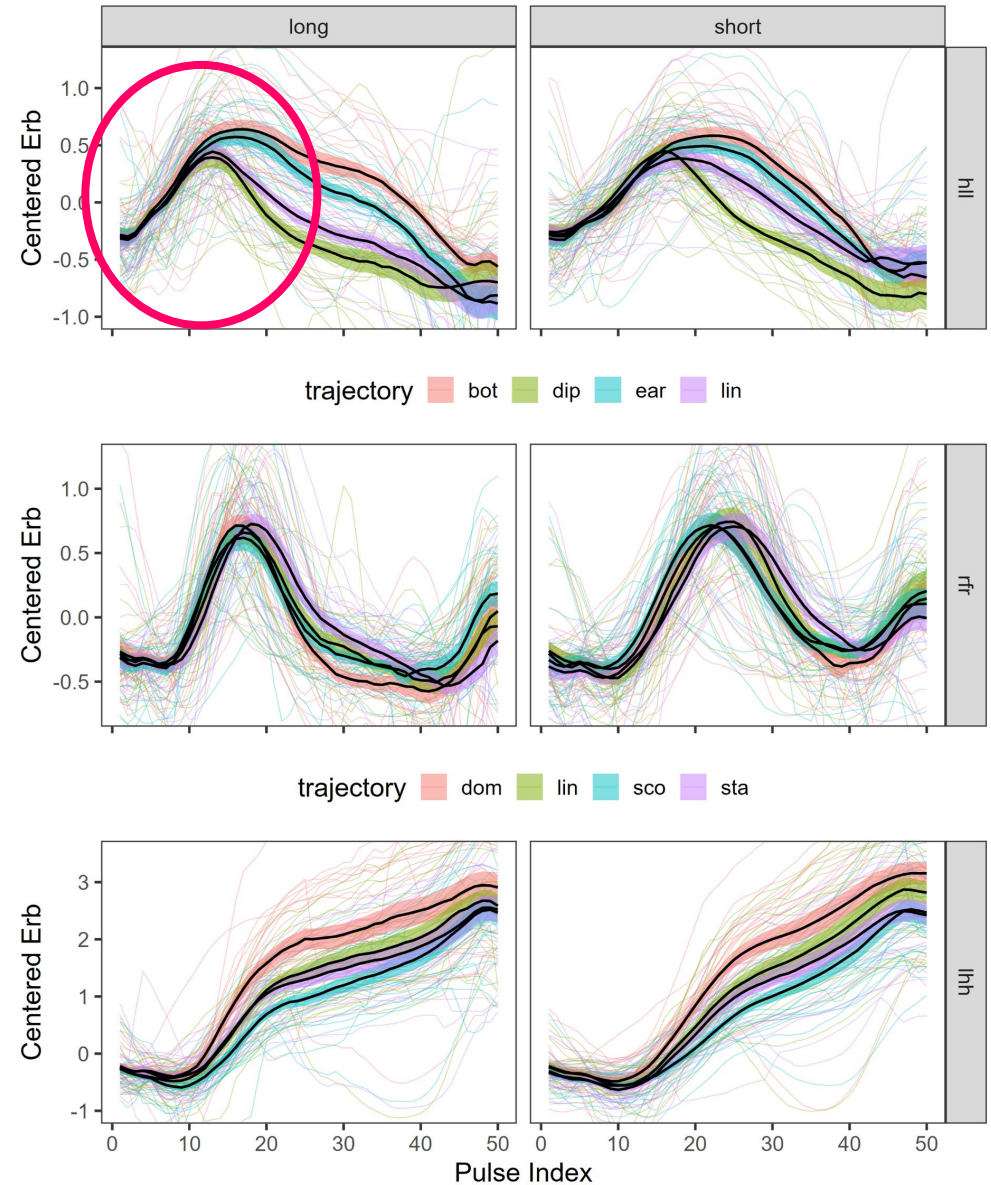
Accented syllable landmark & accentual targets

Option 1: F0 samples spaced at equal distances across the entire nuclear interval
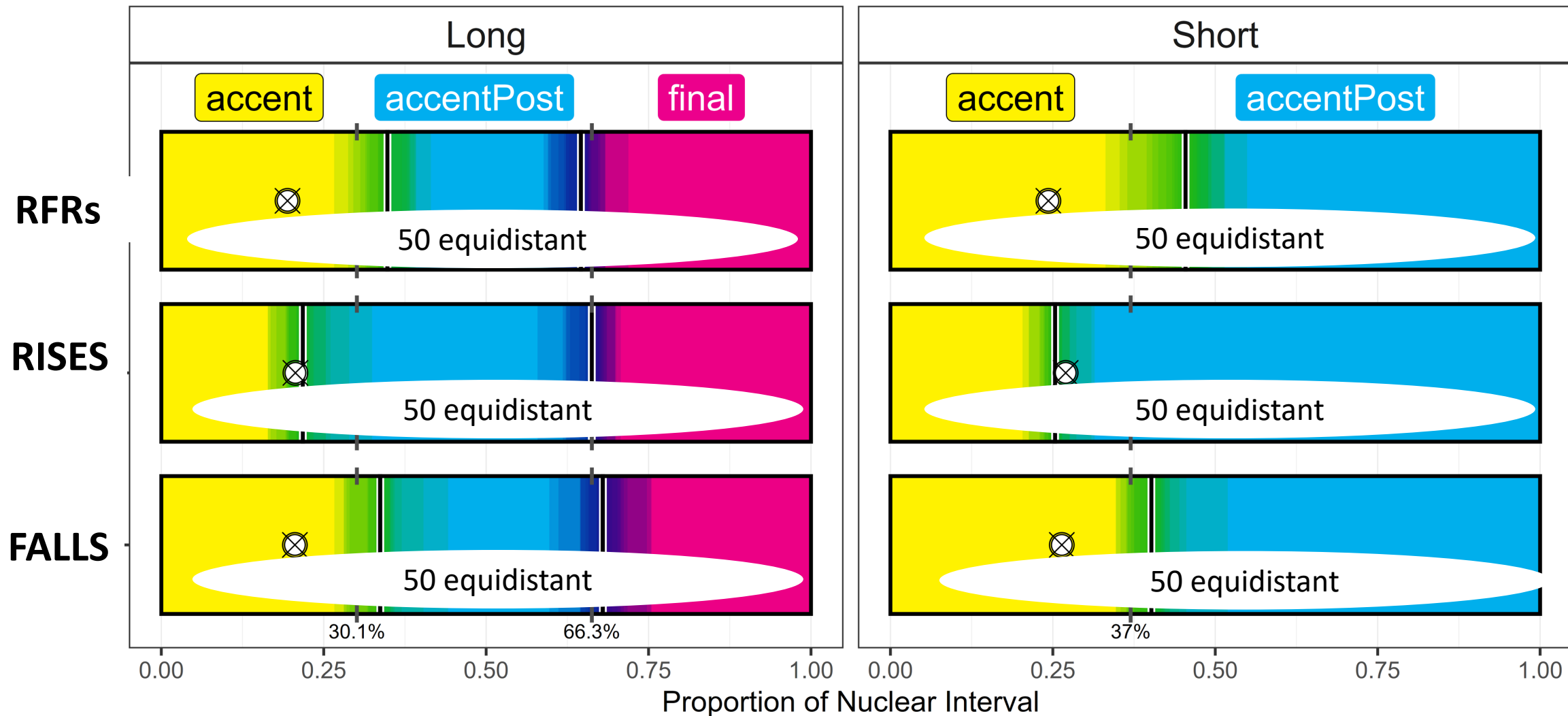
Accented syllable landmark & accentual targets

Option 2: F0 samples assigned based on overall mean duration of internal intervals
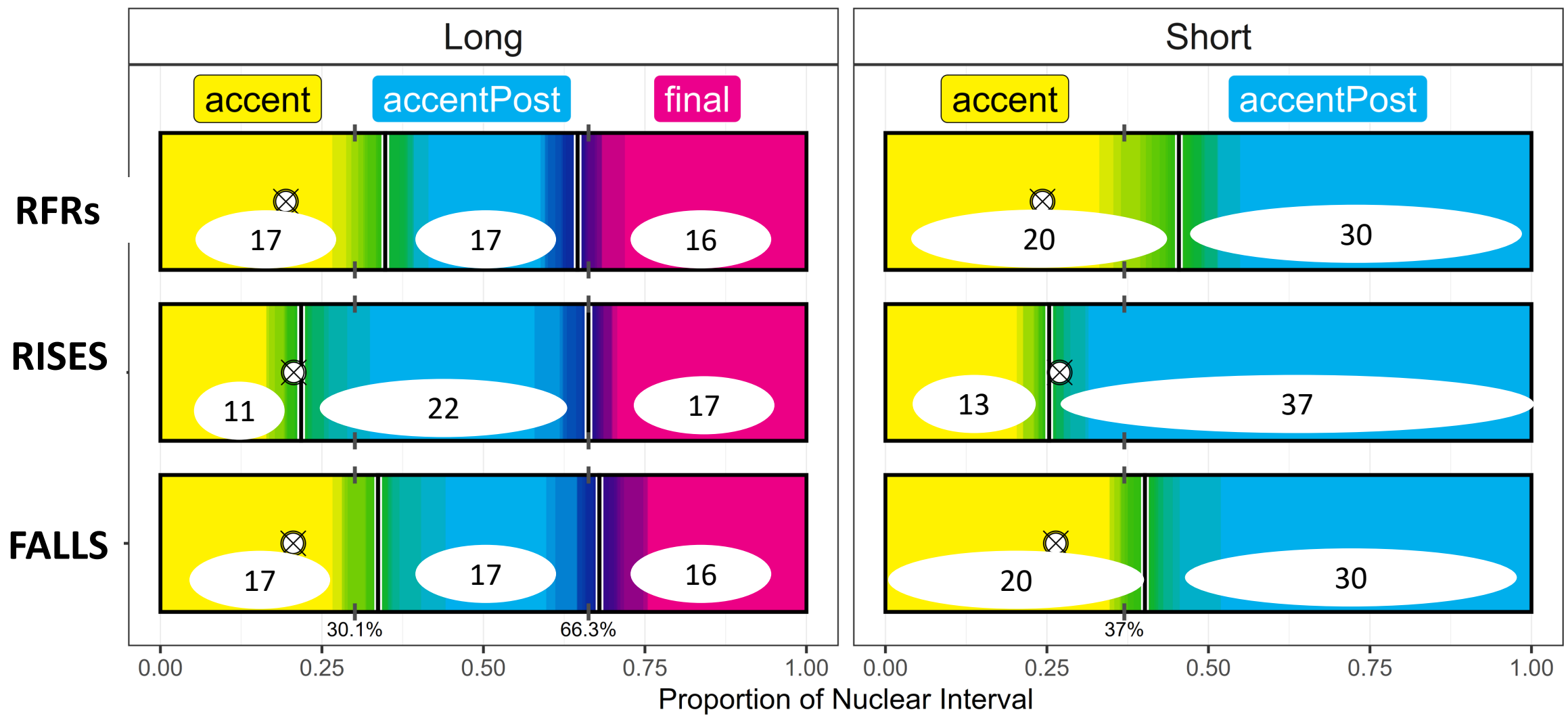
Empirical means:
F0 trajectories, aligned and segmented
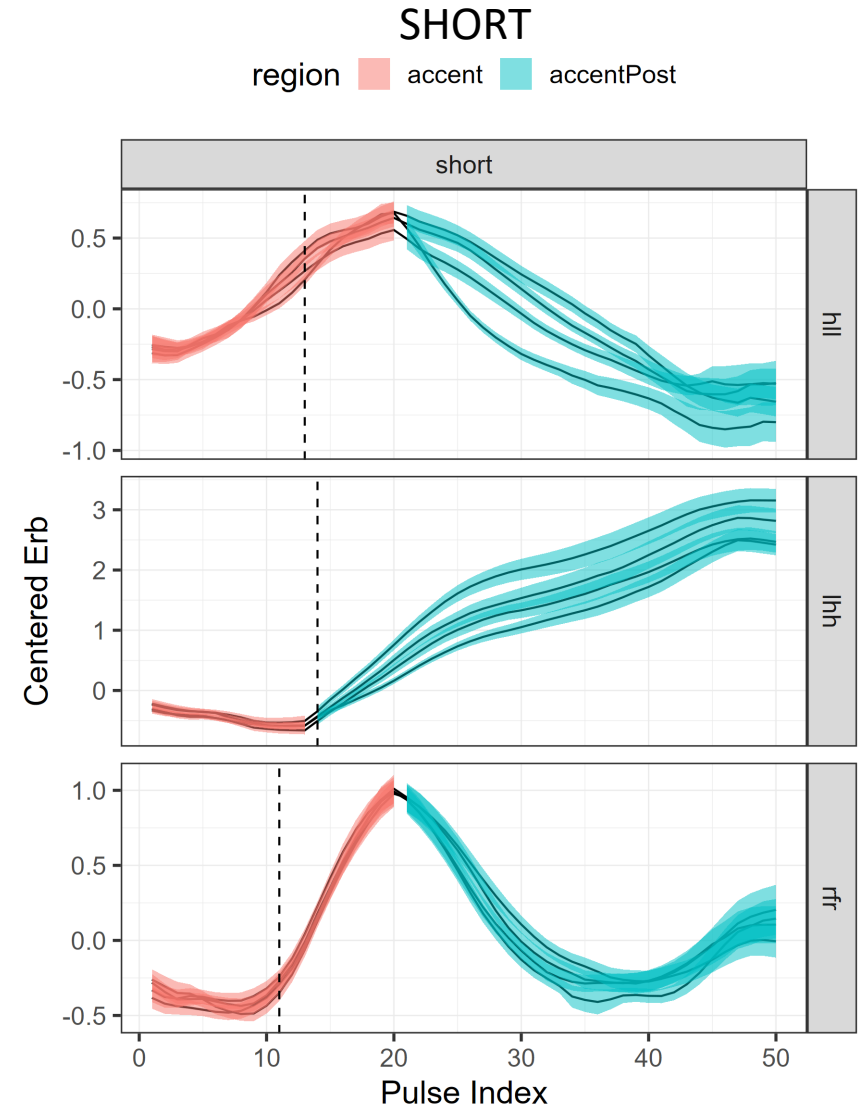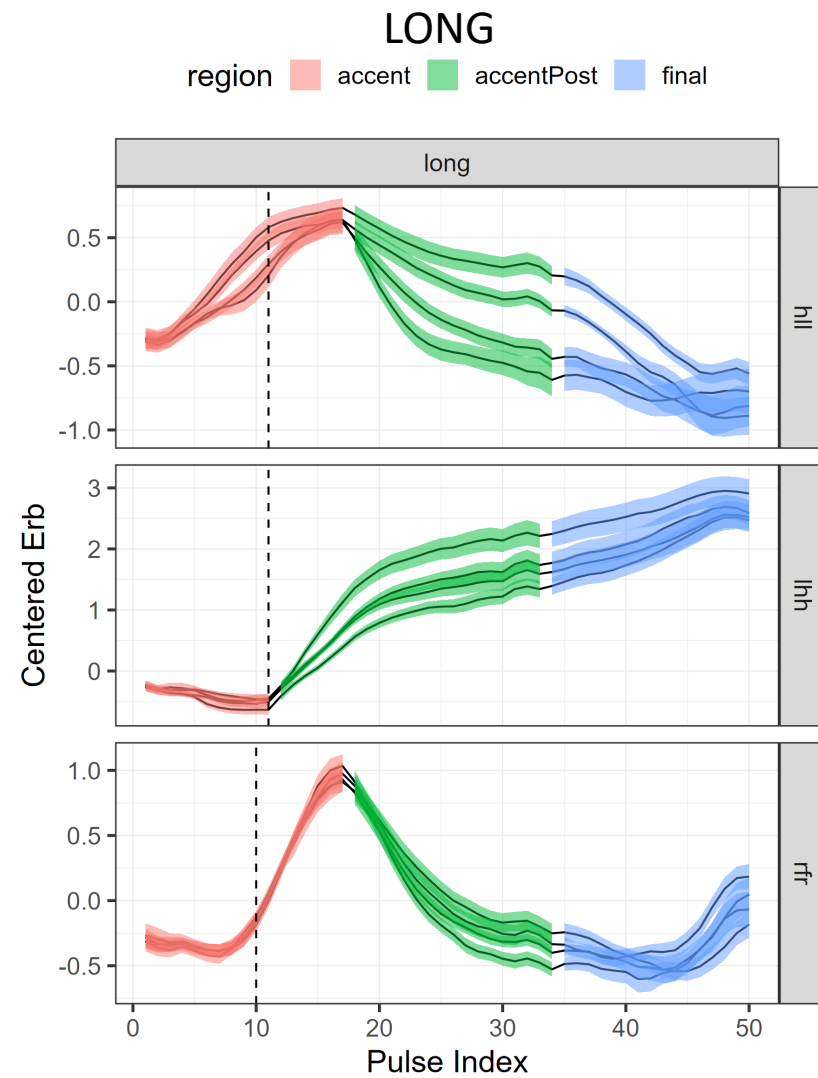
For each tune, what are the distinct F0 shapes speakers produce when imitating these stimuli?

Hypothesis from the AM model:
The middle interval sustains the pitch target of the phrase accent:
- Early fall
- Bottom-out
- Staircase (with upstepped H%)

Falls
H*L-L%
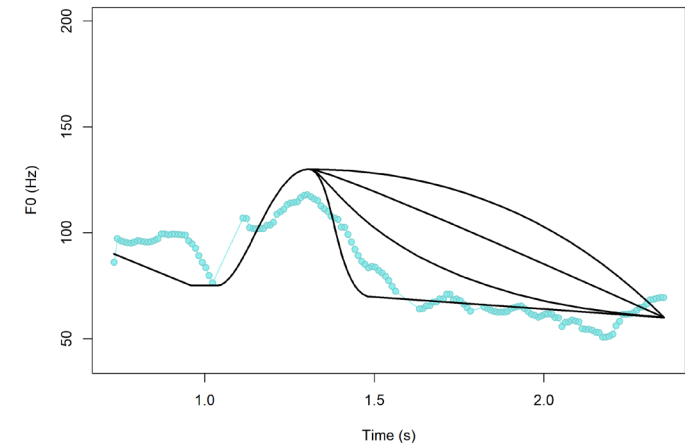
Rise-Fall-Rises
L*H L-H%
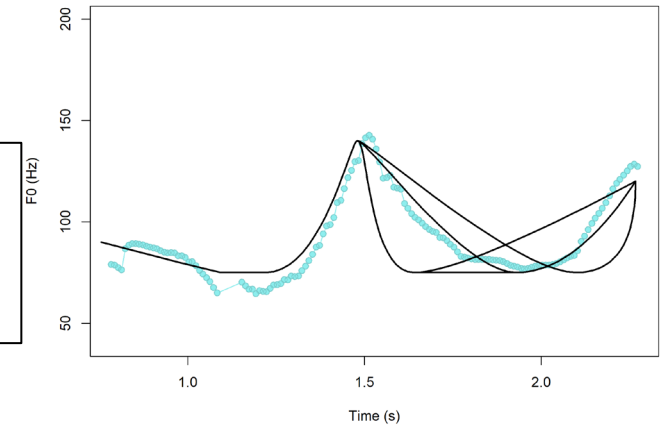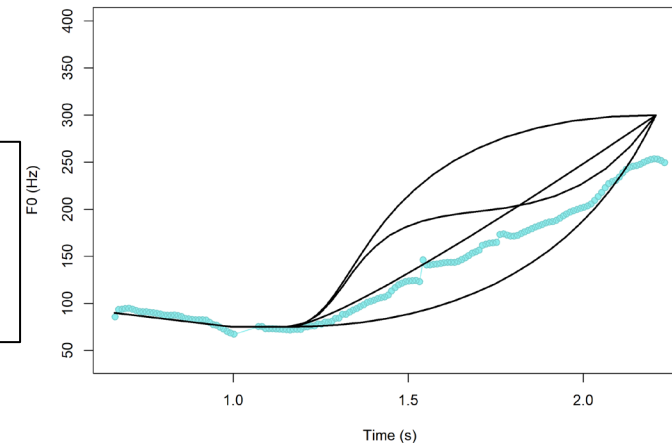
Rises
L*H-H%

For each tune, what are the distinct F0 shapes speakers produce when imitating these stimuli?

Hypothesis from the AM model:
The middle interval sustains the pitch target of the phrase accent:
- Early fall
- Bottom-out
- Staircase (with upstepped H%)

Interpolation Hypothesis:
The middle interval is a straight-line interpolation between the accentual target and the boundary tone:
- Linear

Exploratory: Are 'domed' or 'scooped' F0 curves after the accentual target reproduced?
Differences in the Tonal Center of Gravity
(Barnes et al. 2010, 2021)

Falls
H*L-L%
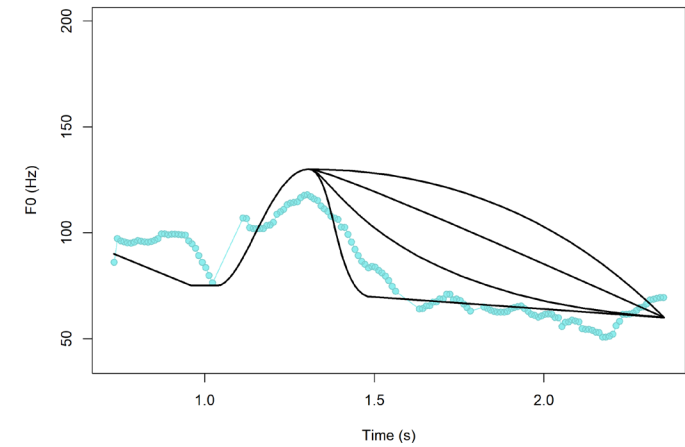
Rise-Fall-Rises
L*H L-H%
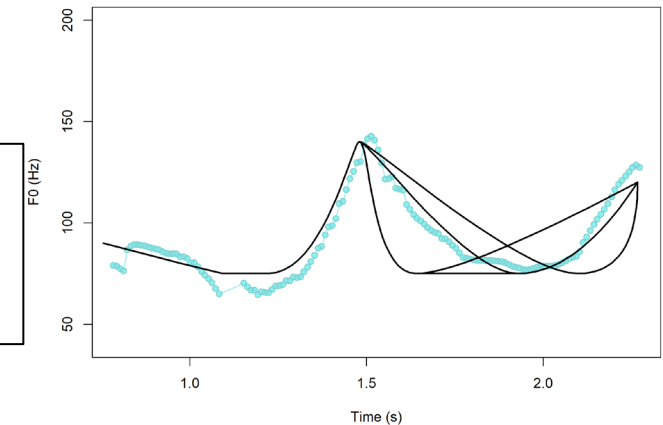
Rises
L*H-H%
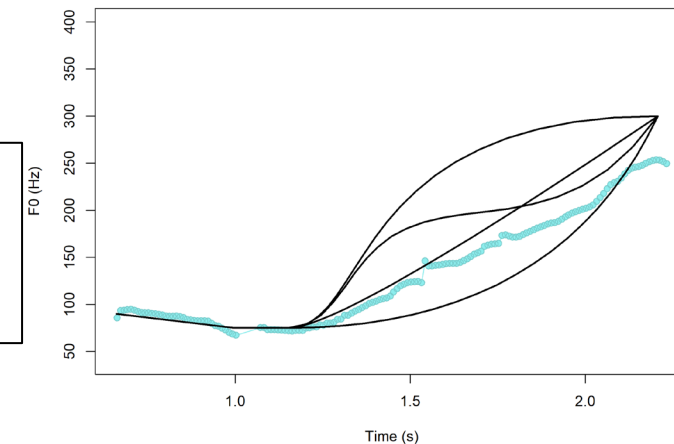
Clustering analysis: What are the robust distinctions in F0 trajectories?

F0 trajectories, segmented and aligned at accentual target
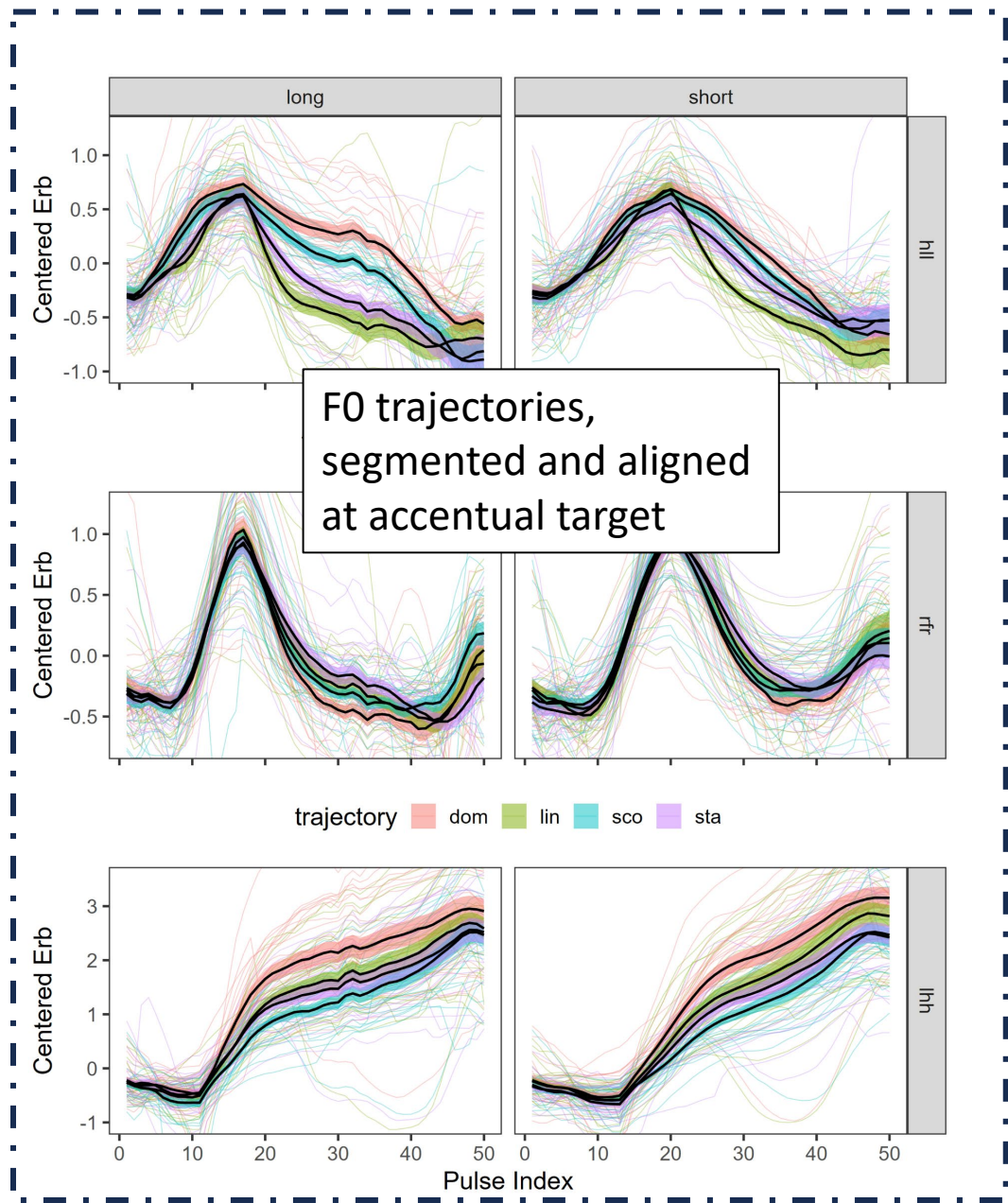
Unlabeled trajectories, time-normalized (50 samples) are submitted to k-means clustering for time-series data.

Two analyses: with and without internal segmentation

F0 trajectories with no internal segmentation
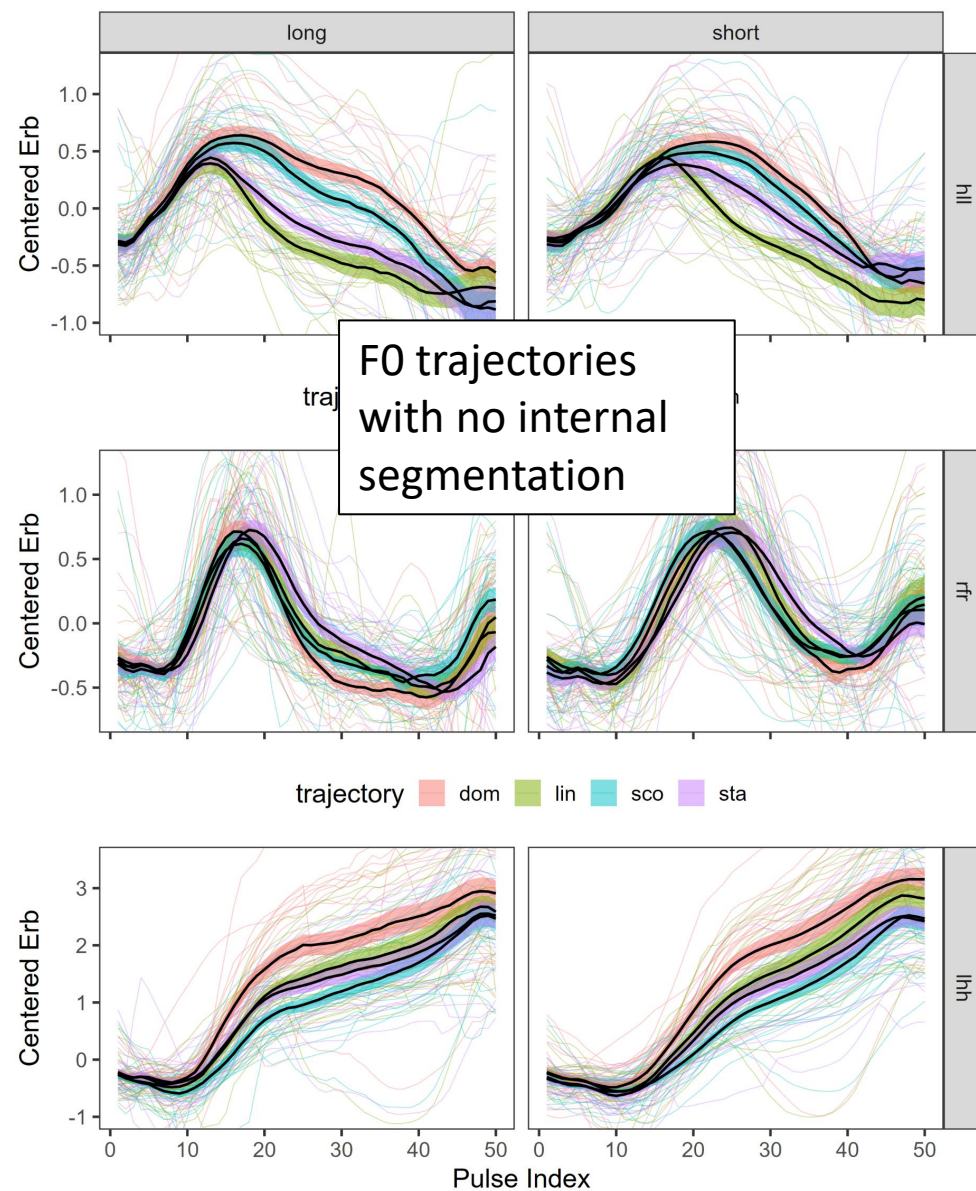
All data – aligned, segmented trajectories

Falls
3 clusters
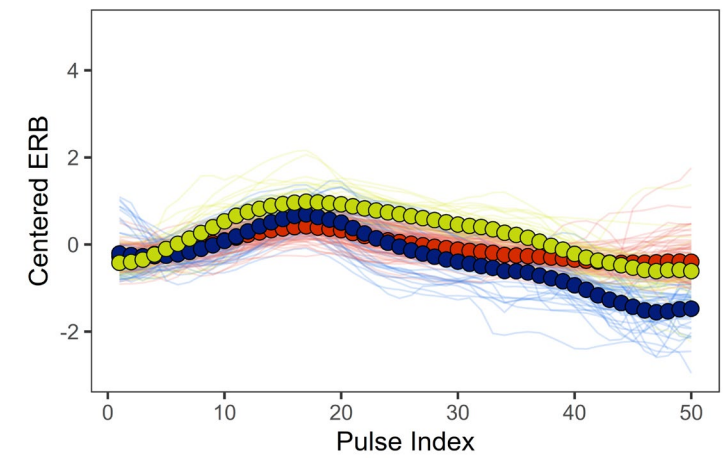
Rise-Fall-Rises
2 clusters
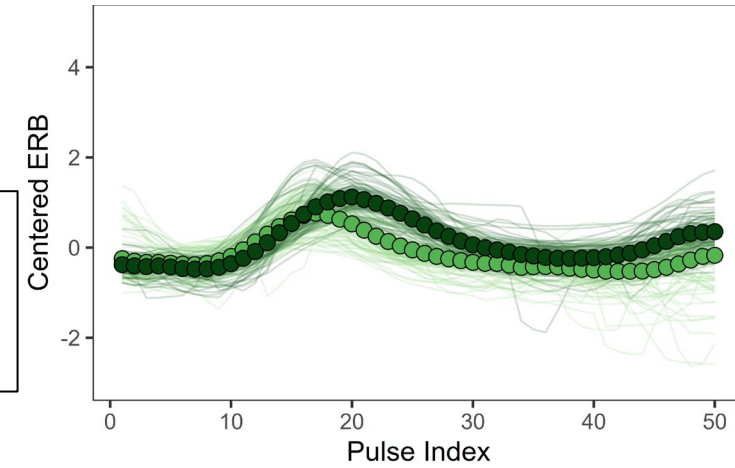
Rises
2 clusters
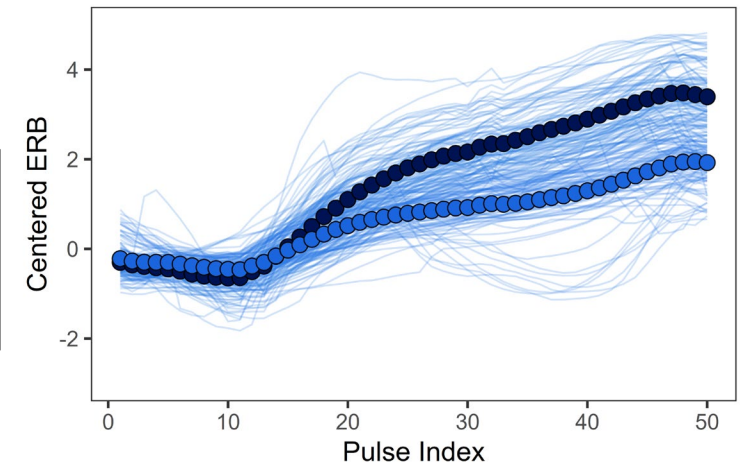
Aligned, segmented trajectories

LONG

SHORT

Only 2 Fall clusters in each length subset
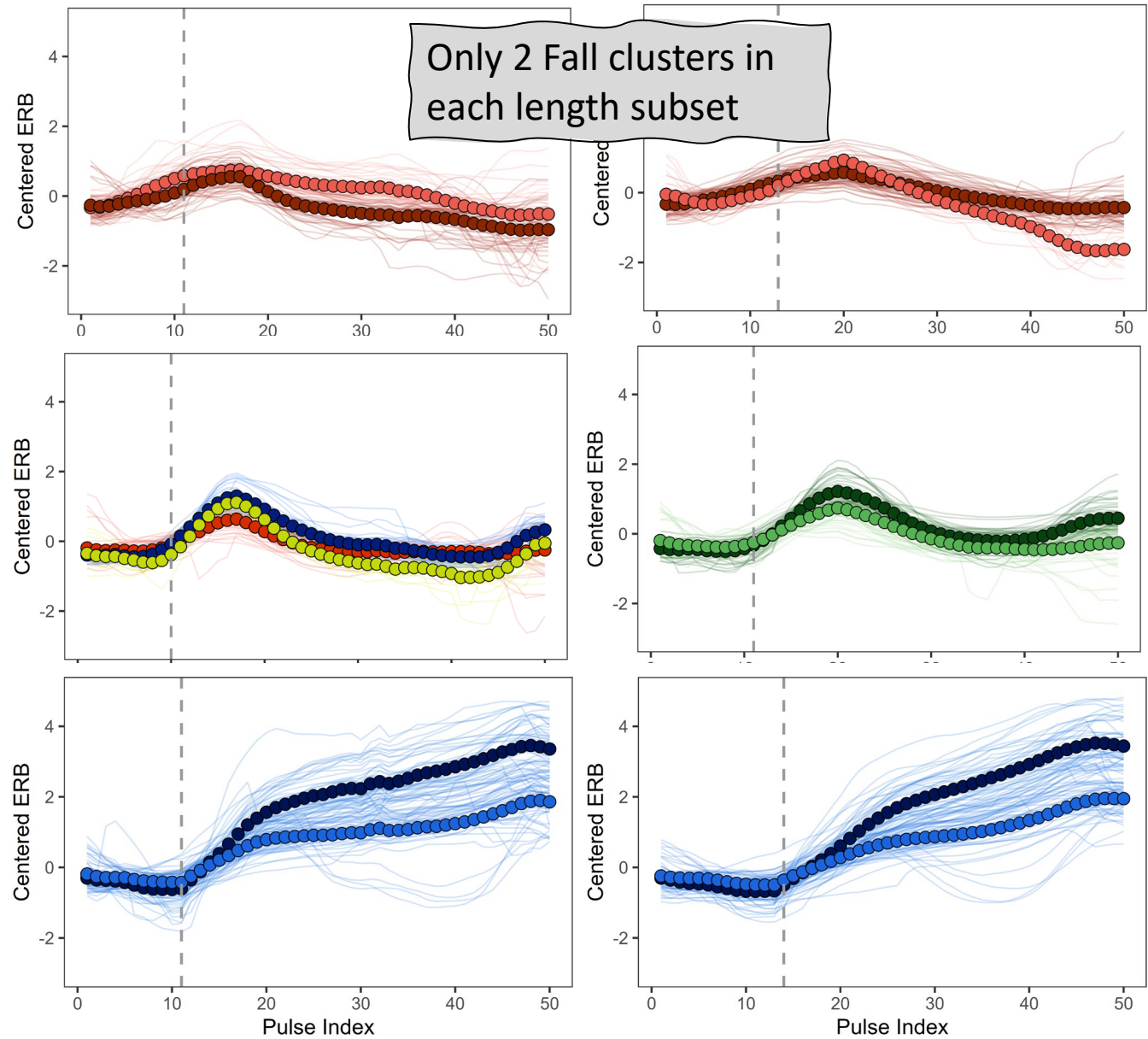
Falls
3 clusters
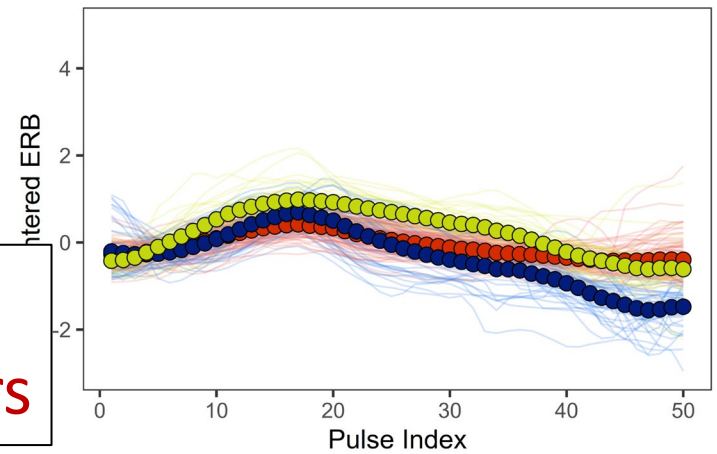
RFR
2 clusters
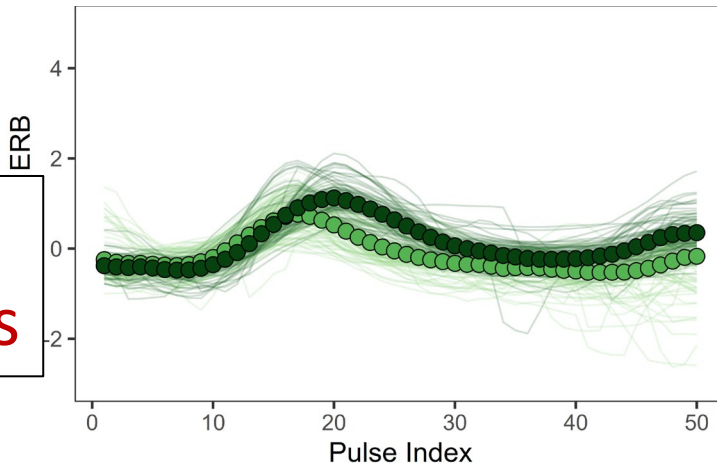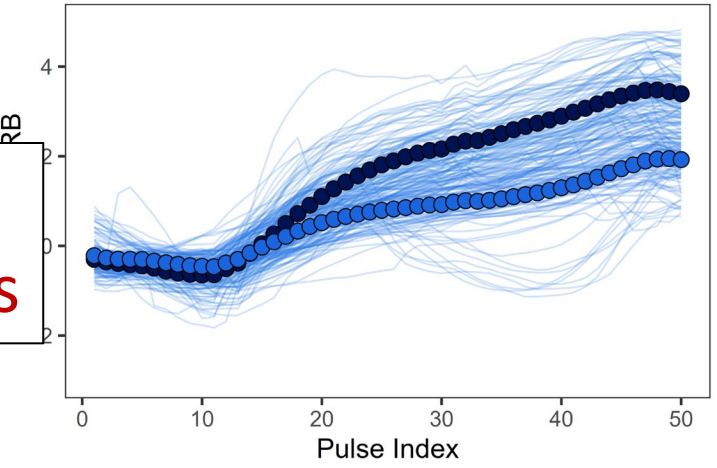
Rises
2 clusters

Aligned, segmented trajectories

LONG

SHORT

A post-focal !H*

Scaling boundary tone

Falls
3 clusters

RFR
2 clusters
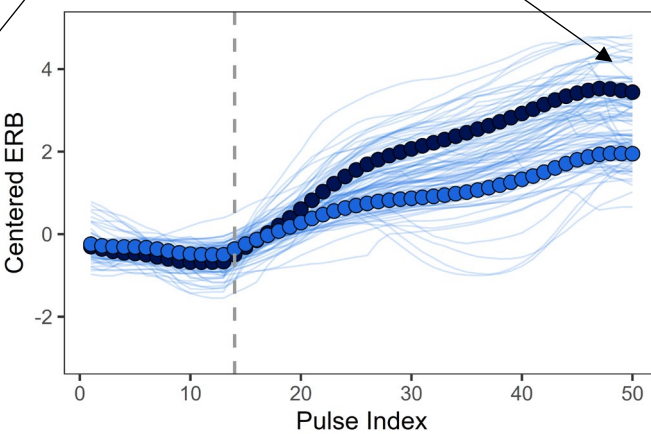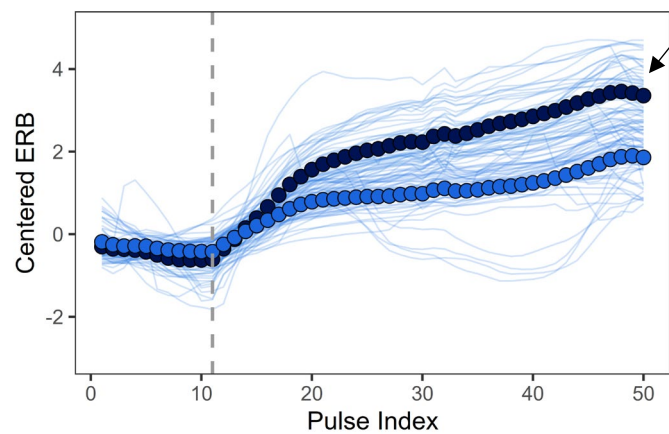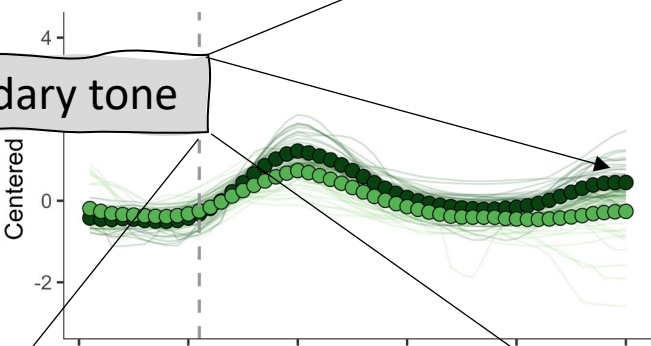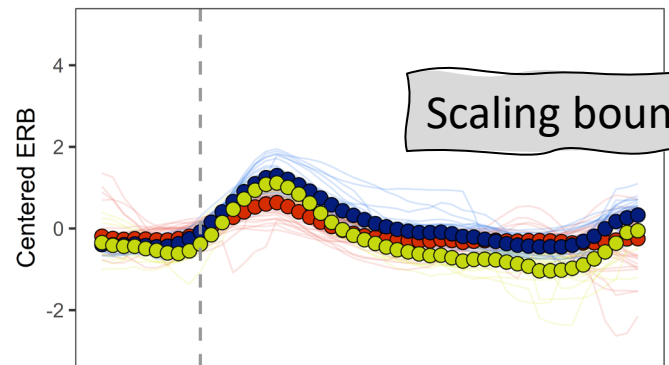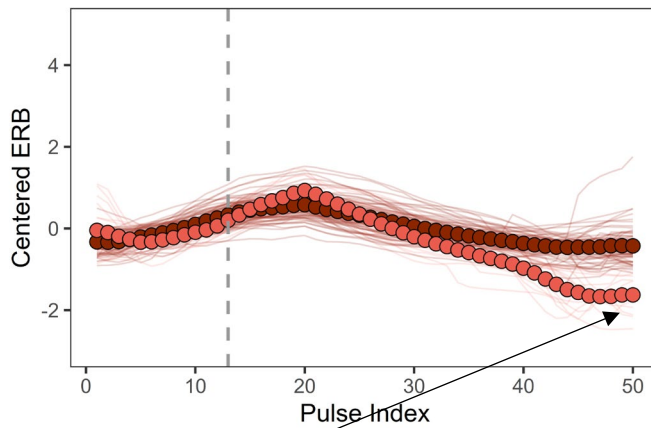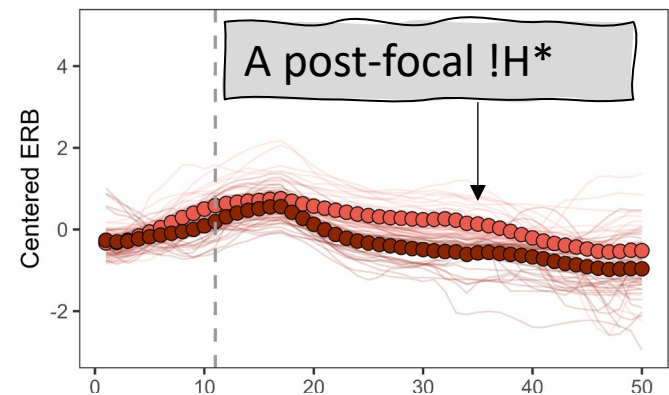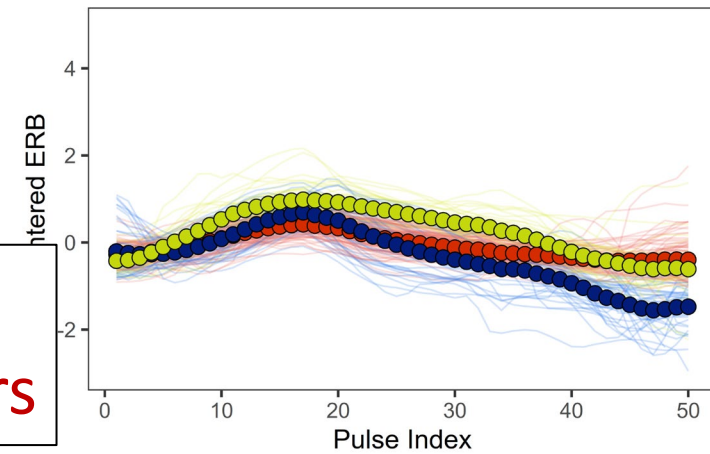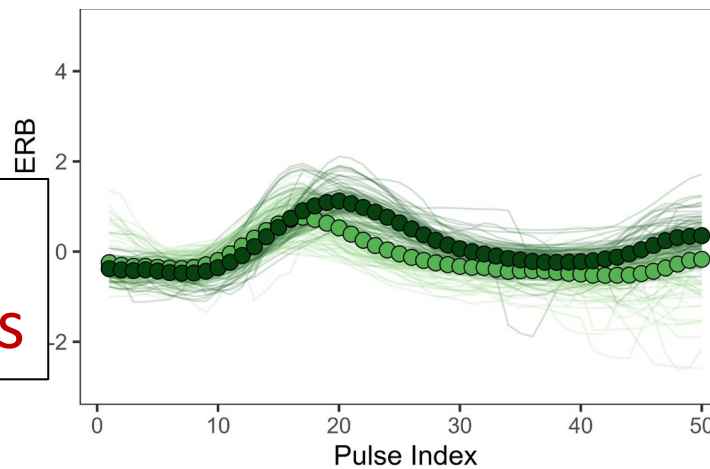
Rises
2 clusters
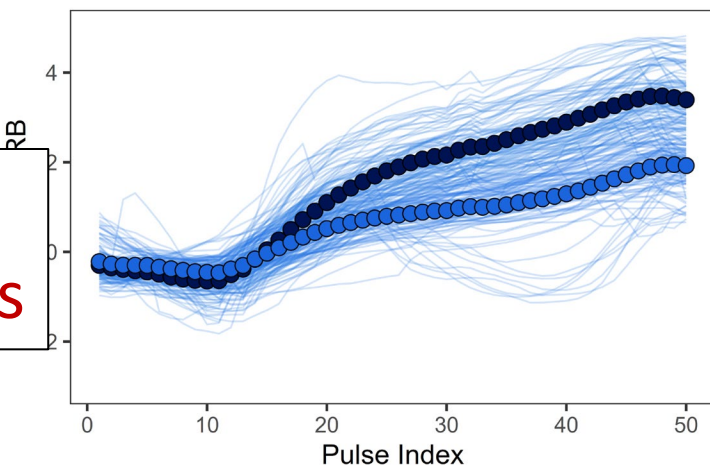
Aligned, segmented trajectories

LONG
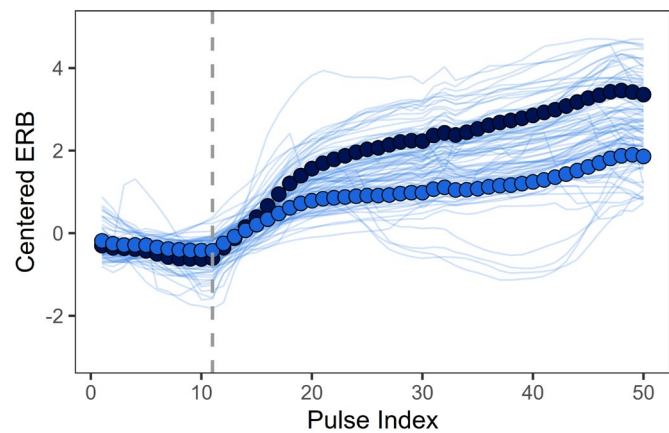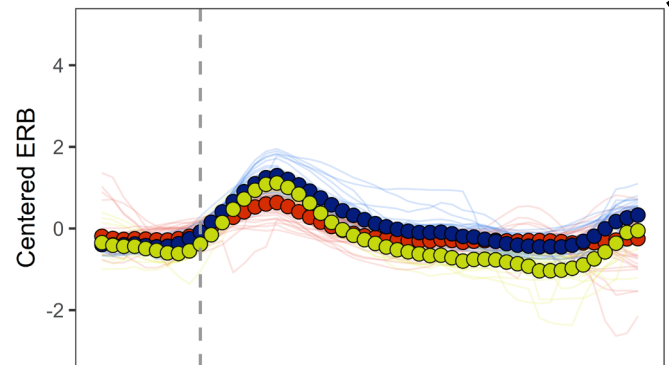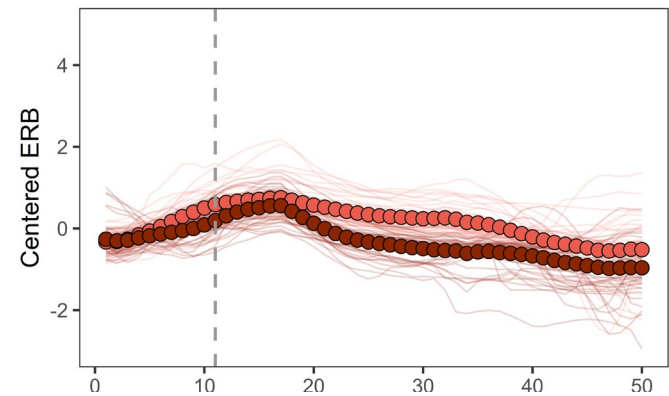
One cluster (red) appears as a Rise-Fall-Plateau – *this is an artifact of averaging. No actual plateau shapes in speakers' data*
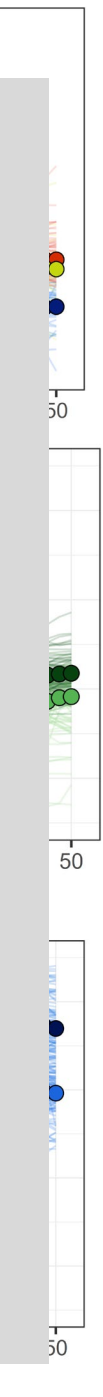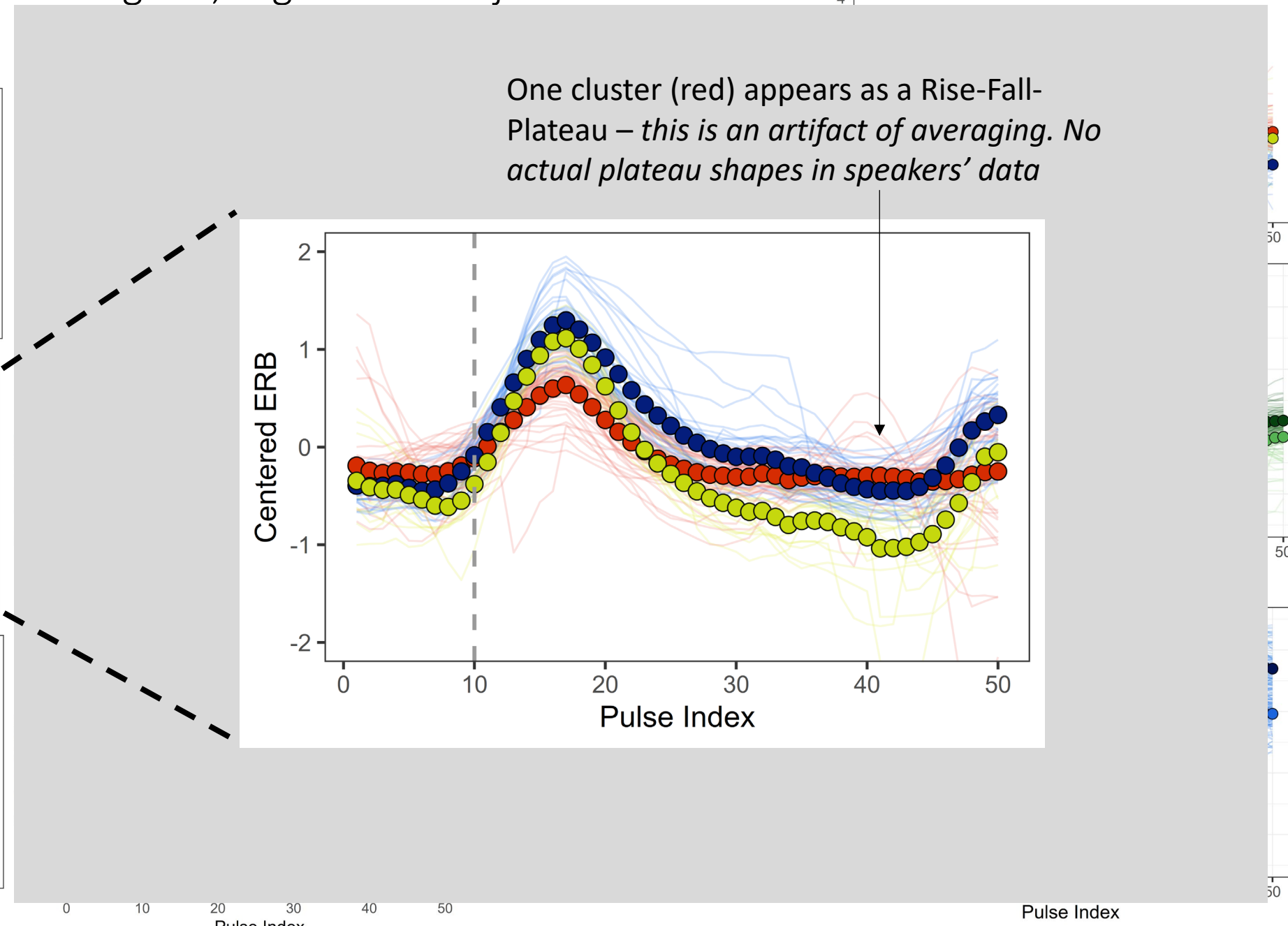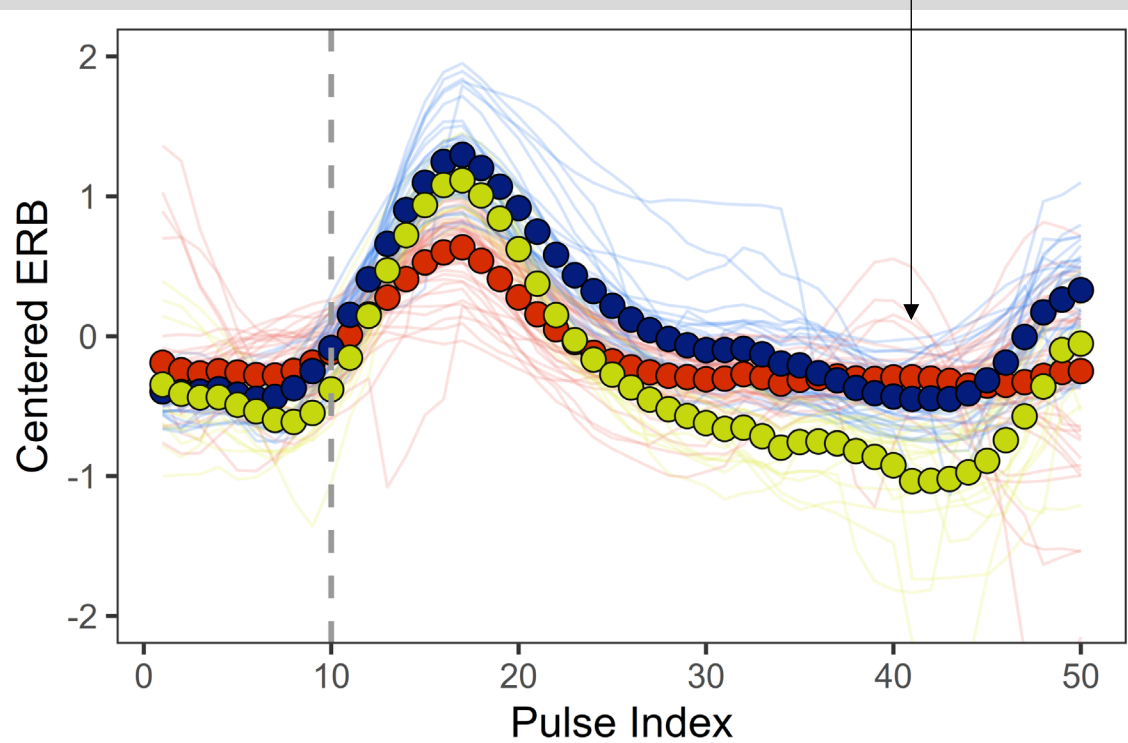
Returning to our hypotheses

Hypothesis from the AM model:
The middle interval sustains the pitch target of the **phrase accent**

Falls
H*L-L%
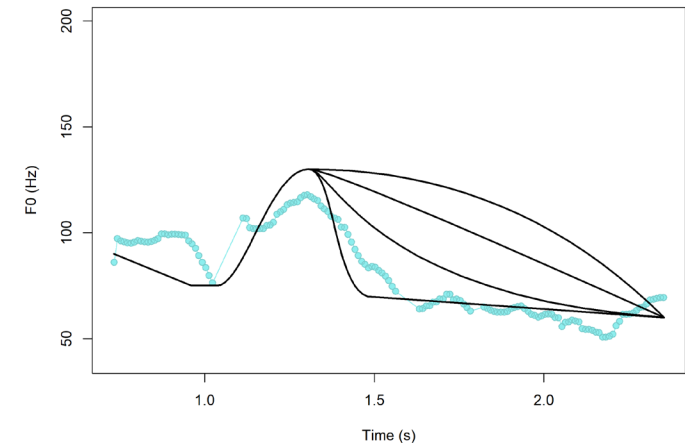
RFR
L*H L-H%
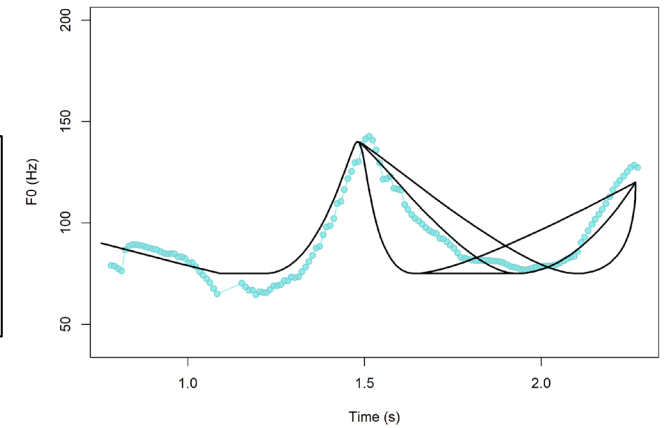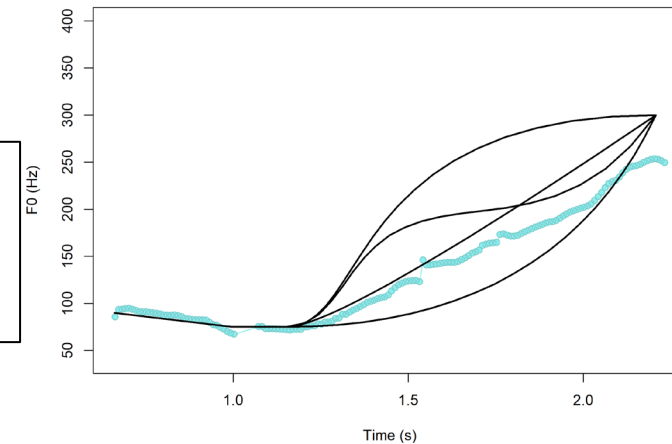
Rises
L*H-H%

Comparison with our results – the long condition:
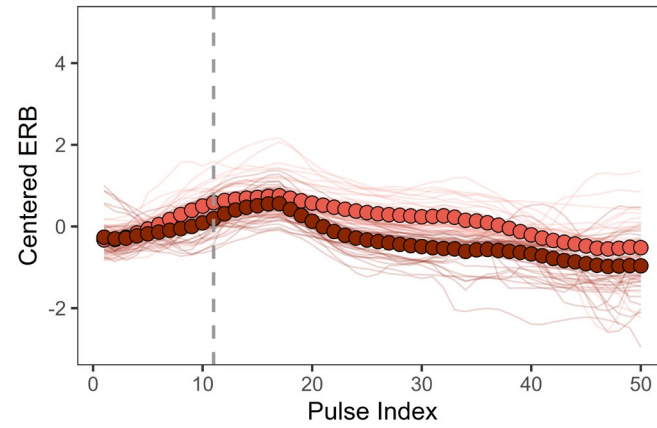
Falls
H*L-L%
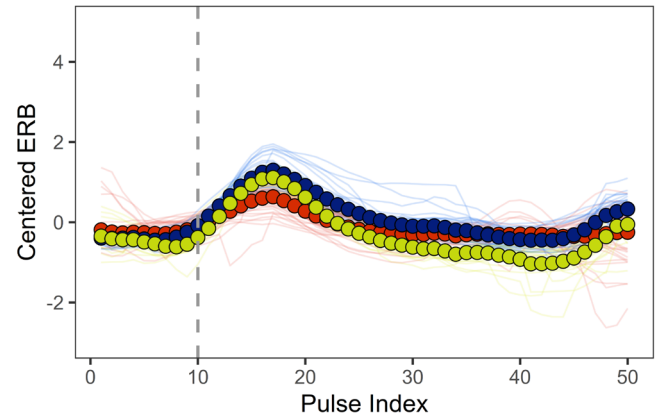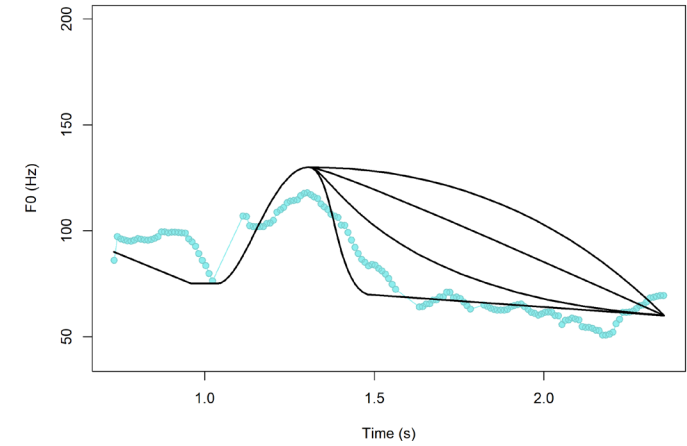
RFR
L*H L-H%
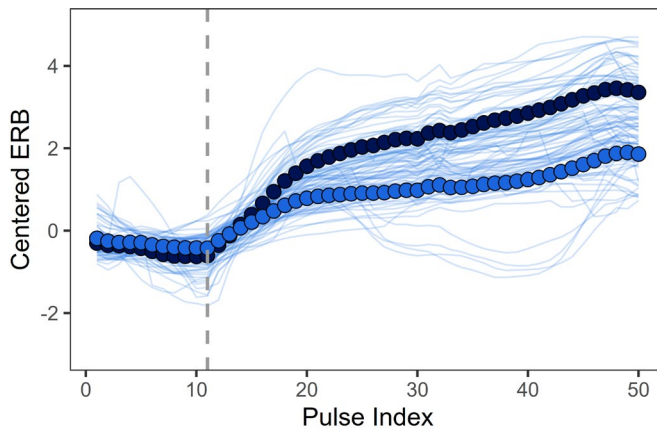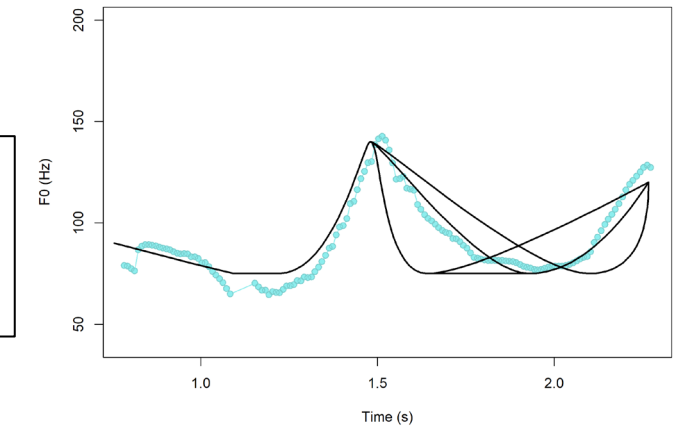
Rises
L*H-H%
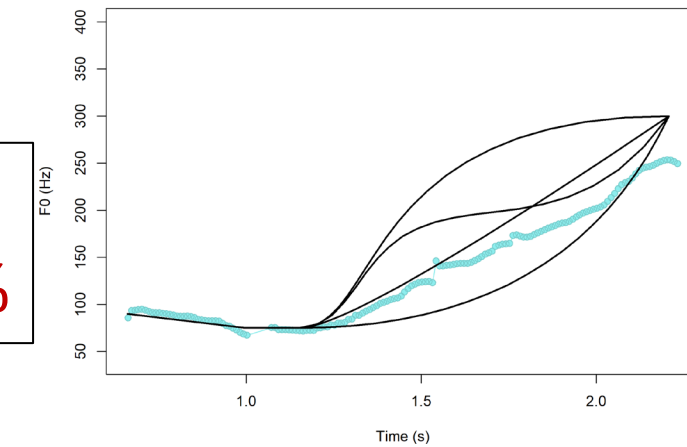
Long condition,
Zooming in y-axis
for Falls and RFR

Falls: Initially steep, then gradual fall over the middle region to ending F0. Option to insert a post-focal !H*

RFR: Initially steep fall, then gradually falling over the middle region, slight rise on final syllable
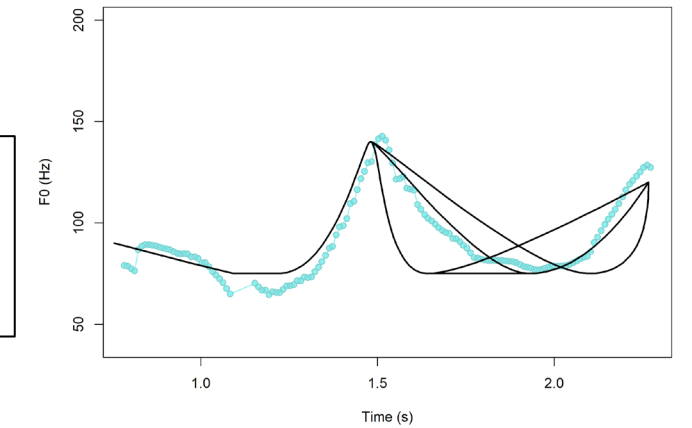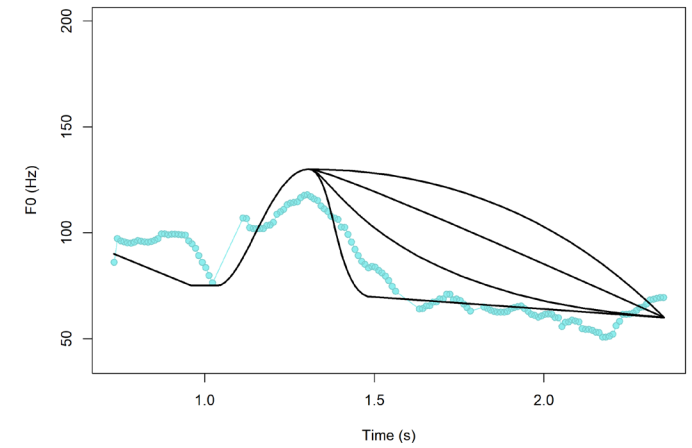
Rise: Initially steep rise, then gradual rise to ending F0
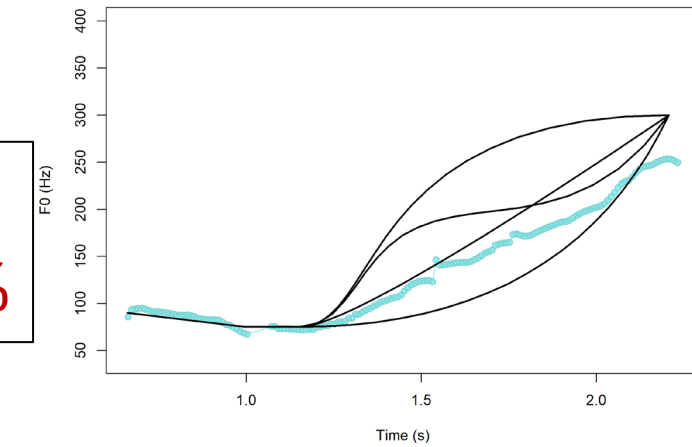
Summing up

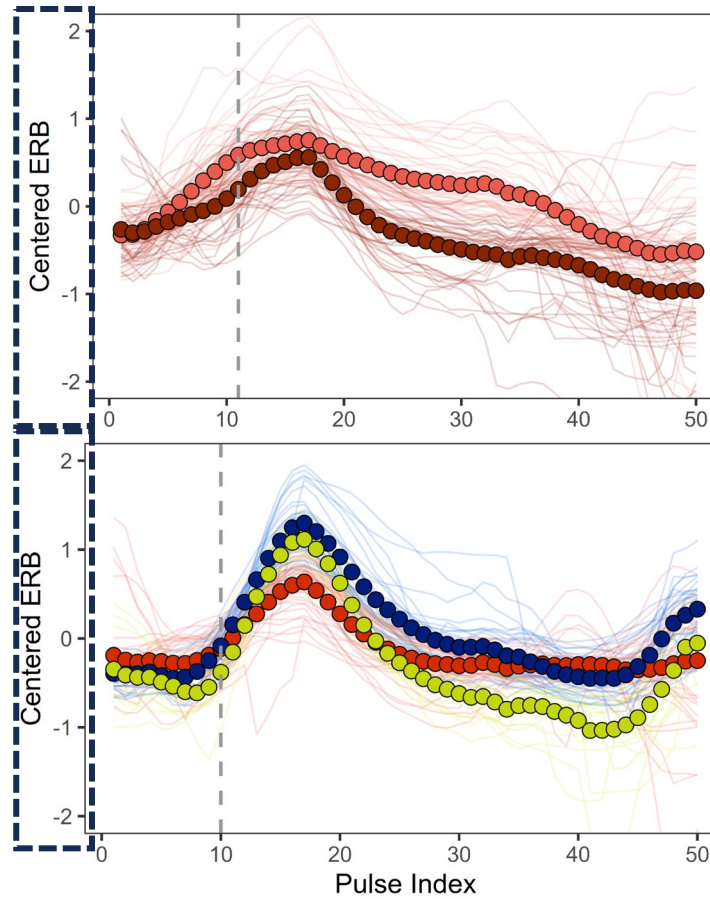Following the accentual target, F0 movement towards a phrase-edge target with the opposite F0 pole is initially

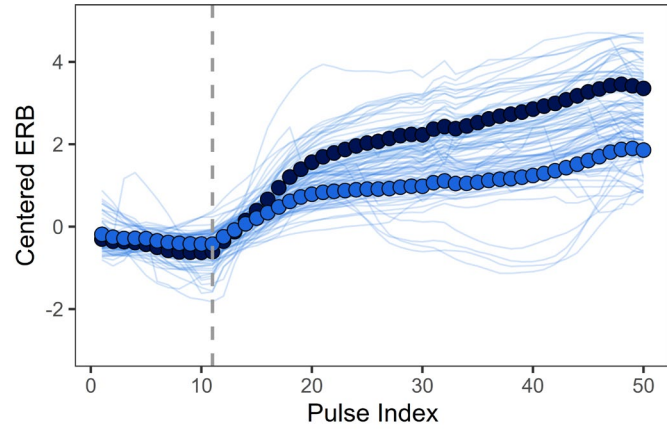🐰 **steep** and may continue in a

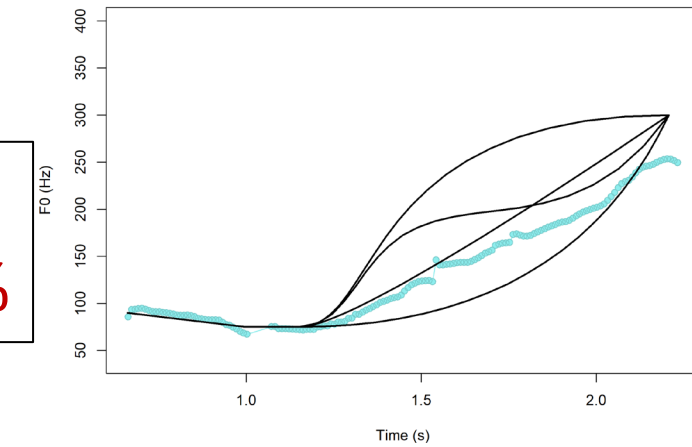🐢 **gradual** pattern of linear change to the final syllable.

→ A dynamic phrase accent?

Falls
H*L-L%

RFR
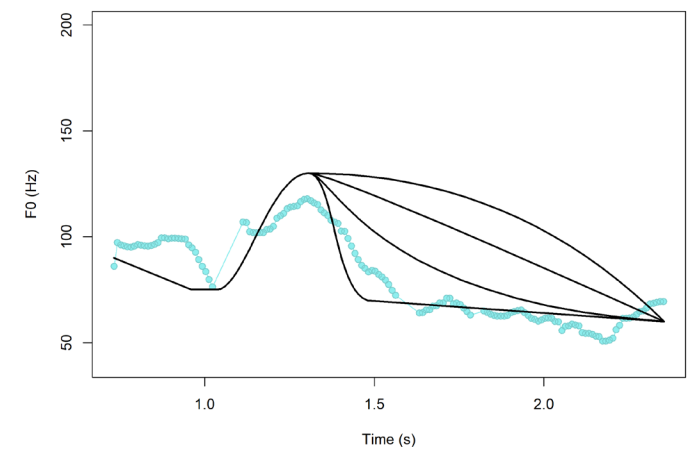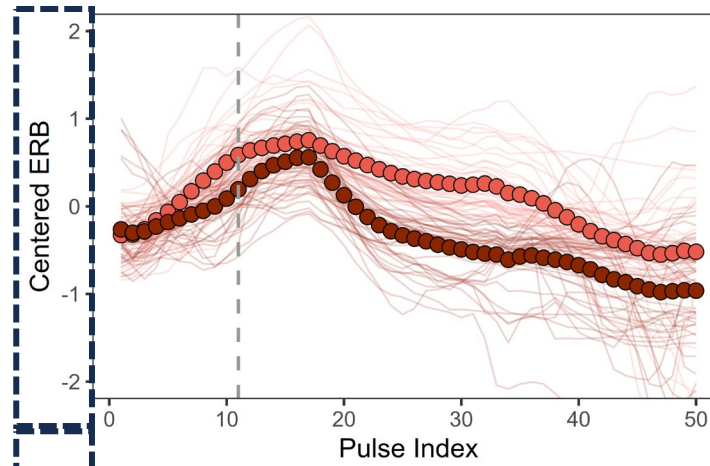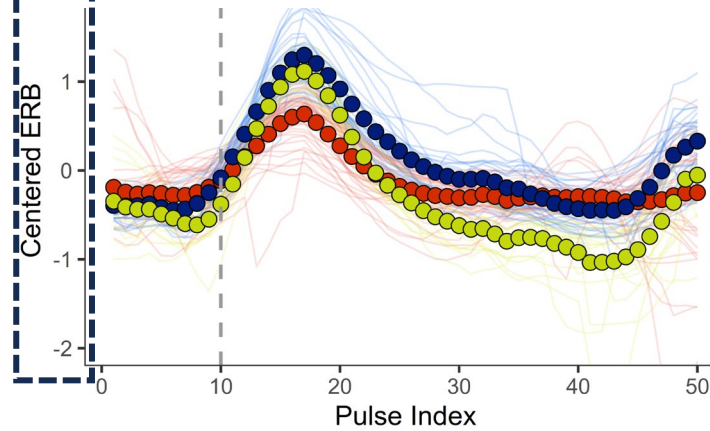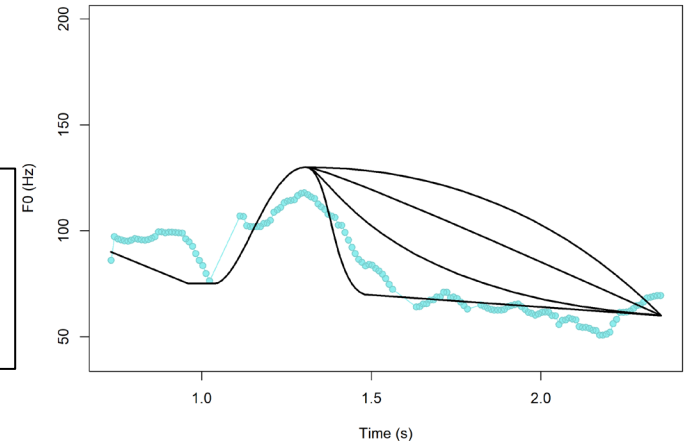L*H L-H%

Rises
L*H-H%
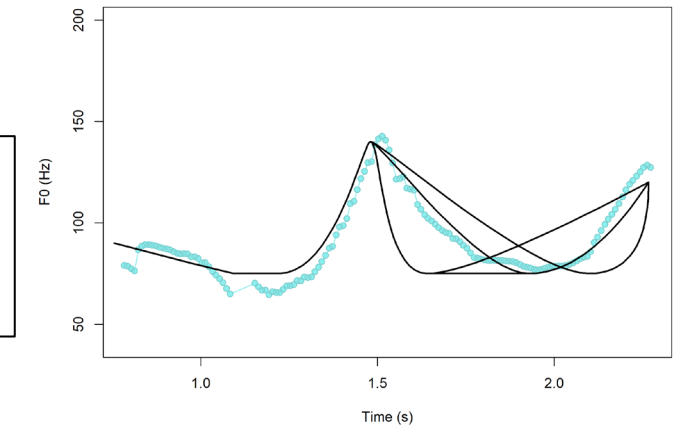
Summing up

Or…

→ Reanalysis of pitch accents: **H*L, L*HL** and **L*H**, with interpolation to boundary tone.

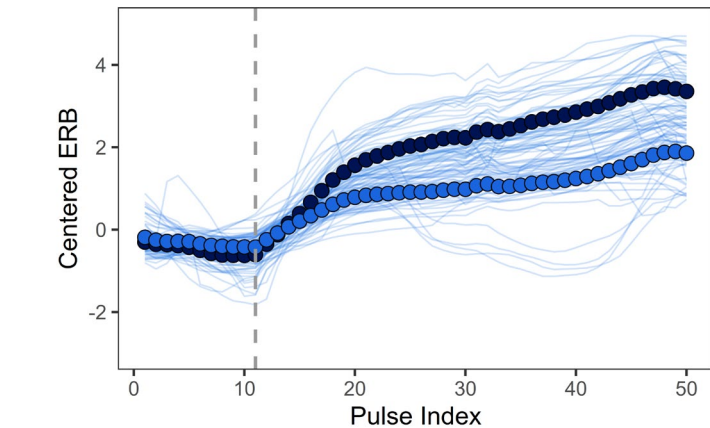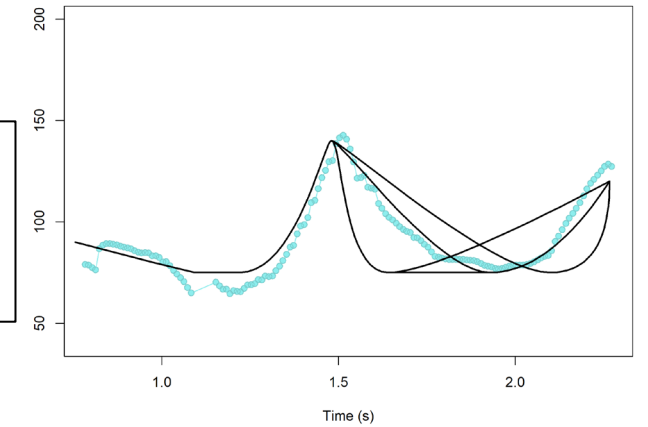→ An alternative account in terms of F0 dynamics over the entire tune (*working on this!*)

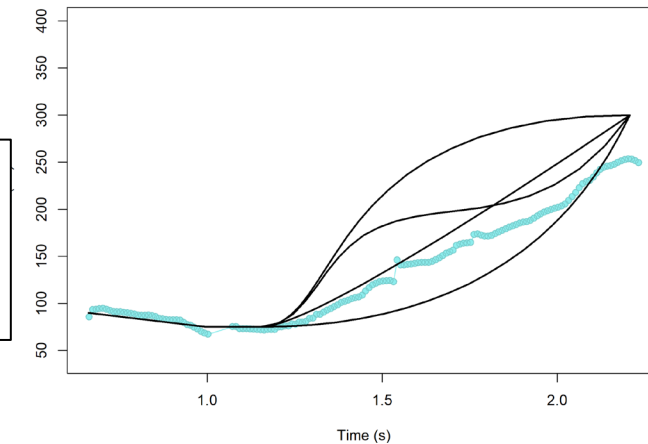The final analysis must take into account other evidence for a phrase accent.

Falls
H*L   L%

RFR
L*HL   H%

Rises
L*H   H%

Thank you!

Falls

H*L L%

RFR

L*HL H%

Rises

L*H H%