Post-focal compression in Hindi – Pitch register as a phonological category?

Frank Kügler
Goethe-University of Frankfurt
Overview

Background
- A typology of the expression of prosodic focus
- Hindi intonation

Production study
- Characteristic repeated rising contours
- Downstep and post-focal compression (PFC)

Perception study
- Presence and absence of PFC in focus identification

Discussion
- Focus perception, speaker variation, and processing of intonational categories

Conclusion
- Pitch register is a phonological category like pitch accents and boundary tones
A typology of the prosodic expression of focus  
(Kügler & Calhoun to appear)

In many languages prosody plays a key role in encoding focus.

A typology of prosodic focus-marking strategies (Kügler & Calhoun, to appear)

- Stress-based cues
- Phrase-based cues
- Register-based cues

Prosodic cues enhance the prominence of a focused word (e.g. Selkirk 1995; Ladd 2008; Calhoun 2010) and/or align the focus with a prosodic boundary (Féry 2013).
A typology of the prosodic expression of focus
(Kügler & Calhoun to appear)

Stress-based cues:

- Enhancing the prosodic cues of a focused word by means of
  - higher fundamental frequency (f0),
  - greater f0 movement,
  - lengthening,
  - increased intensity and higher spectral tilt on the focused word
  - as well as a drop in f0 after it (cf. Ladd 2008, Fletcher 2010, Turk 2012).

Figure 1. Focus position affects prosodic realisation (from Kügler & Calhoun, to app.).
Phrase-based cues

- Languages predominantly do not have lexical prominence (stress).
- Intonation units are phrase tones / phrase boundaries.
- Focus induces additional phrase boundaries or deletes them post-focally.

(1) a. VP focus: What did he do?
   (Anaményá nyumbá ndí mwáála)φ
   he hit the house with a rock
   ‘He hit the house with a rock.’

   b. Object focus: What did he hit with the rock?
   (Anaményá nyuúmba)φ (ndí mwáála)φ

   c. Verb focus: What did he do to the house with the rock?
   (Anaméenya)φ (nyuúmba)φ (ndí mwáála)φ
   (Chichewa, adapted from Kanerva 1990: 156)
A typology of the prosodic expression of focus
(Kügler & Calhoun to appear)

Register-based cues

- Mandarin (Xu 1999; 2011):
  - Register expansion under focus (high H’s and lower L’s)
  - Post-focus compression
- Lexical tones are not affected, no pitch accents (stress), no phrase boundaries.
- Hindi may belong to this group of languages.
Background – Hindi intonation

Every prosodic phrase is associated with a rising pitch gesture, except the last one (Moore 1965, Harnsberger & Judge 1994, Patil et al. 2008).

Typical areal feature of South Asian languages (Khan 2016)

(Harnsberger 1999)

billi: mami: ke pitSe gayi:
cat Mommy PP behind went
‘The cat went behind Mommy.’
Intonation and Prominence in Hindi

Wide focus / All-new sentence:
(10) A: kyaa huaa?
    ‘What happened?’
B: [graahak ne davaaii ko khariidaa]$_F$
    ‘The customer bought the medicine.’

Narrow focus on the second constituent (object):
(11) A: graahak ne kyaa khariidaa?
    ‘What did the customer buy?’
B: graahak ne [davaaii]$_F$ ko khariidaa
    ‘The customer bought the medicine.’

→ No prominence marking by means of intonation.

(Patil, Kentner, Gollrad, Kügler, Féry, Vasishth 2008)
Intonation and Prominence in Hindi

Wide focus / All-new sentence:

(10) A: kyaa huaa?
    ‘What happened?’
B: [graahak ne davaaii ko khariidaa]_{F}
    ‘The customer bought the medicine.’

Narrow focus on the first constituent (subject):

(12) A: kisa ne davaaii ko khariidaa?
    ‘Who bought the medicine?’
B: [graahak ne]_{F} davaaii ko khariidaa
    ‘The customer bought the medicine.’

→ Pitch register compression on the 2\textsuperscript{nd} constituent (object).

(Patil, Kentner, Gollrad, Kügler, Féry, Vasishth 2008)
Results – Hindi intonation (Patil et al. 2008)

graahak ne davaaii ko khariidaa

customer ERG medicine ACC buy.PST

“The customer bought the medicine.”
Contrastive focus in Hindi (Genzel & Kügler 2010)

Contrast on adjective

• Adjective from one to three syllables

1. Q: Kyaa Romina ne purane gaane ko khojaa?  ‘Did Romina search for an old song?’
   A: Nahi! Romina ne ek naye gaane ko khojaa.  ‘No! Romina searched for a new song.’

2. Q: Kyaa Mohinder ne miithe biskit ko khyaa?  ‘Did Mohinder eat a sweet biscuit?’
   A: Nahi! Mohinder ne ek nankim bikit ko khayaa.  ‘Mohinder ate a salty biscuit.’

3. Q: Kyaa Naveena ne khurdare lungii ko bechaa?  ‘Did Naveena sell a raspy lunghi?’
   A: Nahi! Naveena ne ek mulayam lungii ko bechaa.  ‘Naveena sold a silky lunghi.’
Contrastive focus in Hindi (Genzel & Kügler 2010)

- Significant effect on duration:
  - Each syllable of a target word is lengthened compared to wide focus renditions

- Significant effect on F0
  - The L tone gets significantly lower under contrastive focus (speaker variation!)
  - The H tone gets significantly higher
  - Hence, pitch span is significantly enhanced under contrastive focus

Similar findings for Urdu (assumed to be structurally similar to Hindi (Jabeen & Braun 2018))
- Longer duration of first and second syllable of words in contrastive focus
- Larger F0-range on words in corrective focus
- Early alignment of H tone in contrastive focus
Hindi intonation


Typologically, Hindi is classified as a “phrase language” (Féry 2010, 2017) – no pitch accents, but only boundary tones – rising pitch gestures represent L and H φ–phrase tones – presumably, intonational function is to demarcate prosodic word boundaries or φ–phrases

Hindi uses register-cues to encode focus (Kügler & Calhoun, to appear)
Research question

Some thoughts on typological classification:
Intonation languages use pitch to highlight information (pitch accent, e.g. German: Féry & Kügler 2008, English: Breen at al. 2010)
Hindi – as a phrase language – shows no clear prosodic effect on the focused constituent (Patil et al. 2008)
However, Hindi does show a post–focal prosodic effect, i.e. compressed pitch register or post–focal compression (PFC)
Similar to e.g. Mandarin (Xu 1999)
Question arises whether this post–focal cue serves as a perceptual cue for focus interpretation?
If so, could pitch register be interpreted as a phonological category like pitch accents and boundary tones are in intonation languages?

➢ A production study and a perception study
A production study

Do Hindi speakers realize prosodic cues of focus in contrastive ellipsis?

(1) a. raahul ne māā ko [davaaii]F dii naa ki [gaaDii]F
Raahul ERG mother DAT medicine give NEG that car
“Raahul gave the medicine to the mother and not the car.”

b. raahul ne [māā ko]F davaaii dii naa ki [naanii ko]F
Raahul ERG mother DAT medicine give NEG that granny DAT
“Raahul gave the medicine to the mother and not to granny.”

Do speakers differentiate sentences like (1) prosodically in order to resolve the ambiguity?

Predictions:
Focus marking is realized by means of PFC (cf. Patil et al.2008).
(1a/b) contain a contrastive focus (object/indirect object).
PFC is expected to be present after the focused object/indirect object.
A production study – Method

Speaker & recordings
30 female native Hindi speakers, age 20-25
Recordings on DAT at Delhi university

Speech materials
Two conditions: (i) indirect object contrast (ii) direct object contrast
Six sentence pairs as (1)
Interspersed with filler (material from Patil et al. 2008)

Data pre-processing
Annotation and F0 analysis in Praat of 30 x 6 x 2 = 360 sentences
Per constituent: F0-minimum before F0-maximum, duration
A production study – Example

![Waveform and transcription example]

27. November 2019
Marburg - Linguistisches Kolloquium
A production study – Results

- Prosodic differentiation of (1):
  Contrast on indirect object shows PFC on the following object
  Contrast on direct object shows no PFC on following verb
- Speaker variation: lesser & higher degree of PFC
A production study – Results (F0)

Q1a – Difference on Indirect Object
→ \( \text{lmer}(\log(IOF0\text{max}) \sim \text{cond} + (1 + \text{cond}|\text{speaker}) + (1 + \text{cond}|\text{item}), \text{data}) \)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.603747</td>
<td>0.02257</td>
<td>248.27</td>
<td></td>
</tr>
<tr>
<td>Condition=IO</td>
<td>0.015450</td>
<td>0.008016</td>
<td>1.93</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Q1b – Difference on Direct object, Indirect object is in contrast
→ \( \text{lmer}(\log(OF0\text{max}) \sim \text{cond} + (1 + \text{cond}|\text{speaker}) + (1 + \text{cond}|\text{item}), \text{data}) \)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.51883</td>
<td>0.02491</td>
<td>221.55</td>
<td></td>
</tr>
<tr>
<td>Condition=IO</td>
<td>-0.03287</td>
<td>0.00694</td>
<td>-4.74</td>
<td>*</td>
</tr>
</tbody>
</table>

Q1c - Difference on Verb, Direct object is in contrast
→ \( \text{lmer}(\log(VF0\text{max}) \sim \text{cond} + (1 + \text{cond}|\text{speaker}) + (1 + \text{cond}|\text{item}), \text{data}) \)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.46828</td>
<td>0.02856</td>
<td>191.47</td>
<td></td>
</tr>
<tr>
<td>Condition=IO</td>
<td>-0.01611</td>
<td>0.01907</td>
<td>-0.84</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
A production study – Results (duration)

Q1a – Difference on Indirect Object
→ lmer(IOdur~cond+(1+cond|speaker)+(1+cond|item),data)

<table>
<thead>
<tr>
<th>Q1a</th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.53529</td>
<td>0.04529</td>
<td>11.819</td>
<td></td>
</tr>
<tr>
<td>Condition=IO</td>
<td>0.01310</td>
<td>0.01079</td>
<td>1.213</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Q1b – Difference on Object, Indirect object is in contrast
→ lmer(Odur~cond+(1+cond|speaker)+(1+cond|item),data)

<table>
<thead>
<tr>
<th>Q1b</th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.418799</td>
<td>0.043654</td>
<td>9.594</td>
<td></td>
</tr>
<tr>
<td>Condition=IO</td>
<td>-0.048867</td>
<td>0.005335</td>
<td>-9.160</td>
<td>*</td>
</tr>
</tbody>
</table>

Q1c - Difference on Verb, Direct object is in contrast
→ lmer(Vdur~cond+(1+cond|speaker)+(1+cond|item),data)

<table>
<thead>
<tr>
<th>Q1c</th>
<th>Estimate</th>
<th>SE</th>
<th>t value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.326627</td>
<td>0.033614</td>
<td>9.717</td>
<td></td>
</tr>
<tr>
<td>Condition=IO</td>
<td>-0.010851</td>
<td>0.007492</td>
<td>-1.448</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
A production study – Speaker variation

![Waveform graphs with labeled segments and pitch values.]
Summary

- F0-maximum of each constituent is the acoustic cue of the constituent-final H tone.
- Downstep of each H tone was observed, similar to previous findings (Harnsberger & Judge 1996, Patil et al. 2008, Féry 2017)
- No difference of scaling of the H tone on the indirect object:
  → no prosodic marking of focus on the focused constituent
  → in line with Patil et al. (2008)
- Difference in scaling of the H tone on the direct object:
  → post-focal compression after indirect object contrast
  → PFC in line with Patil et al. (2008)
- If there is no prosodic marking of focus on the focused word, would listeners be able to identify a focus perceptually?
- What about speaker variation (speaker with PFC and without PFC)?
A perception study

Do listeners use the prosodic cue PFC in sentence processing? In particular, do listeners identify the focus on the basis of PFC?

Sentence completion task:

(2) raahul ne māā ko davaaīi dii naa ki …
Raahul ERG mother DAT medicine give NEG that
“Raahul gave the medicine to the mother and not …”

Predictions

If PFC is present, listeners will identify the indirect object contrast correctly.
If PFC is absent, listeners will show a chance level identification of the contrast (both direct object contrast and speakers without PFC).
A perception study – Method

Speech materials and experimental design

Data from 3 groups of speakers:

(i) 0 Hz PFC  (ii) 10 Hz PFC  (iii) 30 Hz PFC

Sentence fragments from five sentences (cf. (2))

– Original data from production study were cut after the conjunction “ki”
– Random presentation of fragments with PRAAT–MFC
– 18 listeners x 5 sentences x 2 conditions x 6 speakers = 1080 trials

Listeners 18 native Hindi listeners (mean age 27 years)

Task

Forced choice sentence completion
A perception study – Results

**Tab.1:** Mean identification of correct sentence completion split by sentence

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Condition Direct Object</th>
<th>Condition Ind. Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.287</td>
<td>0.824</td>
</tr>
<tr>
<td>2</td>
<td>0.519</td>
<td>0.639</td>
</tr>
<tr>
<td>3</td>
<td>0.370</td>
<td>0.769</td>
</tr>
<tr>
<td>4</td>
<td>0.472</td>
<td>0.741</td>
</tr>
<tr>
<td>5</td>
<td>0.556</td>
<td>0.574</td>
</tr>
</tbody>
</table>

**Tab.2:** Mean identification of correct sentence completion split by item (speaker)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Condition Direct Object</th>
<th>Condition Ind. Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.489</td>
<td>0.611</td>
</tr>
<tr>
<td>21</td>
<td>0.422</td>
<td>0.622</td>
</tr>
<tr>
<td>8</td>
<td>0.478</td>
<td>0.800</td>
</tr>
<tr>
<td>18</td>
<td>0.422</td>
<td>0.678</td>
</tr>
<tr>
<td>26</td>
<td>0.456</td>
<td>0.767</td>
</tr>
<tr>
<td>30</td>
<td>0.378</td>
<td>0.778</td>
</tr>
</tbody>
</table>

0 Hz  10 Hz  30 Hz
A perception study – Results

Linear mixed effects model with “condition” as fixed factor and “listener” and “item” as random factors

\[
\text{glmer(}\text{rating} \sim \text{cond} + (1 + \text{cond}|\text{subject}) + (1|\text{item}) \ldots
\]

- Conditions differ significantly in terms of correct sentence completion.
- Condition OBJECT CONTRAST (no PFC): On average, correct sentence completion in 44.1 % of the cases (chance level).
- Condition INDIRECT OBJECT CONTRAST (PFC): On average, correct sentence completion in 70.9 % of the cases.
- What about speaker variation? Speakers with and without PFC.
A perception study – Results

A model that includes speaker variation
(2 speakers with no PFC compared to 4 speakers with PFC)

\[
\text{glmer}(\text{rating} \sim \text{cond} \times \text{speakervar} + (1 + \text{cond}|\text{subject}) + (1|\text{item})...)
\]

|                      | Estimate | SE     | z value | Pr(>|z|) | Sign. |
|----------------------|----------|--------|---------|----------|-------|
| (Intercept)          | -0.5075  | 0.4835 | -1.050  | 0.29393  |       |
| Condition =          | 1.3397   | 0.8648 | 1.549   | 0.12135  |       |
| Ind. Obj. Contrast   |          |        |         |          |       |
| Speaker=PFC          | -0.1276  | 0.2164 | -0.590  | 0.55552  |       |
| Interaction          | 0.9584   | 0.3096 | 3.096   | 0.00196  | **    |

- No significant effect of CONDITION and SPEAKER GROUP
- Significant interaction in condition INDIRECT OBJECT CONTRAST (PFC)
  → speaker variation matters
  - Items (speakers) with PFC appear to lead to more correct sentence completion
  - Items (Speakers) without PFC appear to lead to chance performance
A perception study – Results

**Interaction plot:** Speakers with no PFC (dotted line) compared to speakers with PFC (solid line) split by condition (Obj. vs. Ind. Obj. contrast)

- Speakers with PFC have a higher (75.6 %) correct sentence completion identification than speakers with no PFC (61.6 %).
- Post-hoc paired samples t-test confirms this pattern: 
  \[ t = -8.39, \text{df} = 359, p < 0.001 \]
- Model comparison: Model with speaker variation shows significant improvement compared to the simple model measured as the difference in the deviance (Bates et al. 2015): \( \Delta D = 14.2, p < 0.001 \).
Summary

The presence of PFC matters for focus perception:
If PFC is present, listeners identify the correct sentence completion.
If PFC is not present, listeners perform on chance level.

Speaker variation found in the production of PFC has an impact on focus perception:
Sentences of speakers with no PFC in the ‘indirect object contrast’ condition were completed by chance.
Sentences of speakers with PFC in the ‘indirect object contrast’ condition were completed correctly (75.6 % on average).

PFC is thus a functional cue to focus identification in Hindi.

Pitch register information, i.e. PFC, is a cue that resolves local ambiguities.
Discussion

In some tone languages, PFC is a robust phonetic cue to focus perception (cf. Xu et al. 2012 for Mandarin).

However, Mandarin also shows highlighting of focus on the focused constituent (cf. Xu 1999).

➢ Thus, there are more cues to focus perception in Mandarin.

Hindi is a language that does not mark focus by means of F0 on the focused constituent (Patil et al. 2008).

However, PFC is a cue that signals focus and that perceptually was shown to function a cue for focus perception.

➢ Thus, PFC as the only cue in production serves as a cue in perception.
Discussion – Perception of focus

- Only few studies that investigate the perception of focus.
- If a prosodic cue systematically found in speech production is present, listeners identify the focus of a sentence (e.g. Botinis et al. 1999, Liu & Xu 2005, Vainio & Järvikivi 2006, Krahmer & Swerts 2007, Xu et al. 2012, Kügler & Gollrad 2015).
- Reliable focus perception for initial or medial focus, not final focus (Botinis et al. 1999, Xu et al. 2012, Wang et al. 2012).
  - For languages with PFC, cue must be present (sentence-finally it is not!)
- Listener’s language background matters (Wang et al. 2012):
  - Listeners from a non-PFC language do not identify focus based on PFC
  - L2 phonological entities are harder identified (cf. Hume & Johnson 2003)
- PFC might be a phonological category similar to pitch accents or boundary tones.
Discussion – Speaker variation

- Many studies found speaker variation with respect to the prosodic expression of focus
  - German: speakers use predominantly F0 or duration (Baumann et al. 2007)
- Compensating strategy for F0 cues (Gussenhoven 2002):
  - Later aligned F0-peaks ~ higher F0-peaks
  - Proven for perception (Ladd & Morton 1997)
- Different strategies reported have a common aim:
  marking a focused constituent as prosodically prominent
- Speaker variation in Hindi:
  - No cue accumulation as in stress-based languages (F0, duration, intensity)
  - Presence or absence of the cue (PFC)
  - Only presence of cue leads to functionally decisive focus identification

- PFC as a phonological category similar to pitch accents or boundary tones.
Discussion – Intonation categories in sentence processing

- Intonation – pitch accents and boundary tones (intonation contour)
- Pitch accents and processing:
  - E.g. Focus Attraction Hypothesis (Schafer et al. 1996):
  - Default interpretation (in silent reading) (1c)
  - Reading (1b) or (1c) depends on the presence of a pitch accent on either ‘plane’ or ‘propeller’

  (1)
  a. The sun sparkled on the propeller of the plane that the mechanic was so carefully examining.
  b. The mechanic carefully examined the propeller (high attachment)
  c. The mechanic carefully examined the plane (low attachment)

- Phrase boundaries and processing:
  - Phrase break marked by longer phrase-final duration (Lehiste 1973)

  (2)
  a. The hostess greeted the girl with a smile.
  b. ( ) ( ) ‘The girl is smiling’
  c. ( ) ( ) ‘The hostess is smiling’
Discussion – Intonation categories in sentence processing

- Intonation – pitch accents and boundary tones (intonation contour)
- Pitch accents and boundary tones as categories of intonation have functional load.
- Pitch accents and boundary tones are typically found in languages using stress-based cues to express focus.
- If pitch register (PFC) has functional load – as was shown in the identification of focus structure in Hindi – pitch register might also be interpreted as an intonational category (at least for languages that use register-based cues to express focus).
Conclusion

Pitch register – PFC –
- is a cue to focus perception.
- is a cue to disambiguate local ambiguities
- thus carries functional load in Hindi.

Pitch register is thus also an effective prosodic cue for disambiguation – besides pitch accents and phrasing.

Pitch register is a phonological category.
Thank you!

Acknowledgements
This work was supported by the DFG, projects “Prosody in Parsing” (DFG-KU 2323/2-1) and “Recursivity in Prosodic Phonology” (KU 2323/4-1). Special thanks to the Co-PI’s of “Prosody in Parsing” Caroline Féry and Shravan Vasishth, and to the project members Gerrit Kentner, Anja Gollrad and Stella Gryllia as well as our student assistant, at that time, Dinah Baer-Henney. Thanks also to Susanne Genzel.
References


References