Urdu/Hindi Questions at the Syntax-Pragmatics-Prosody Interface

Joint work by
María Biezma, Miriam Butt, Tina Bögel, Farhat Jabeen

University of Konstanz

Tromsø, May 2018
Part of a Research Unit (FOR 2111) *Questions at the Interfaces* at Konstanz

- Looking at non-canonical uses of questions across languages
- We are Project P4, working on Urdu/Hindi
- Biezma is Project P2, looking at Romance

Generally trying to understand the interplay between **prosody, morphosyntax** and **semantics/pragmatics**.

This talk:

- focus on issues at the prosody-syntax interface
- look at polar *kya* ‘what’ in interaction with other question types

Different theoretical frameworks across FOR 2111

But the common assumption: meaning is calculated **compositionally**.
Structure of Talk

1. Some background on questions in Urdu/Hindi
2. Uses of kya ‘what’
3. Polar kya: interaction between prosodic information, syntax and interpretation
4. Ambiguous Strings: Polar kya and wh-constituent (thematic) kya
5. Resolution via a new proposal for the Prosody-Syntax Interface
6. Outlook and Summary
Background: Questions in Urdu/Hindi

Urdu/Hindi has traditionally been characterized as a wh-in-situ language (but also see Bayer and Cheng 2015).

(1) a. sita=ne  d\textsuperscript{h}yan=se  ram=ko  dek\textsuperscript{h}-a  t\textsuperscript{h}-a  
Sita.F=Erg carefully  Ram.M=Acc see-Perf.M.Sg be.Past-M.Sg  
'Sita had looked at Ram carefully'

   b. sita=ne  d\textsuperscript{h}yan=se  kis=ko  dek\textsuperscript{h}-a  t\textsuperscript{h}-a?  
Sita.F=Erg carefully  who.Obl=Acc see-Perf.M.Sg be.Past-M.Sg  
'Who had Sita looked at carefully?'

The default word order in Urdu/Hindi is **SOV**.
But: **default** position for questions is actually the **preverbal** focus position (for information structure analyses of Urdu/Hindi, see Gambhir 1981, Butt and King 1996, 1997, Kidwai 2000).

(2) a. \textit{sita}=\textit{ne} \quad \textit{ram}=\textit{ko} \quad \textit{dek}^{h}-\textit{a} \quad \textit{th}-\textit{a}  \\
Sita.F=Erg \quad Ram.M=Acc \quad \text{see-Perf.M.Sg} \quad \text{be.Past-M.Sg}  \\
‘Sita had seen Ram.’

b. \textit{ram}=\textit{ko} \quad \textit{kis}=\textit{ne} \quad \textit{dek}^{h}-\textit{a} \quad \textit{th}-\textit{a}?  \\
Ram.M=Acc \quad \text{who.Obl=Erg} \quad \text{see-Perf.M.Sg} \quad \text{be.Past-M.Sg}  \\
‘Who saw Ram?’
Féry et al. (2016) conducted a comparative study of Hindi and Indian English.

They asked questions like:

- *In front of the well, who is pushing the car?* (Questioning the Subject)
- *In front of the well, what is the man pushing?* (Questioning the Object)

They found the following word orders in the responses.

<table>
<thead>
<tr>
<th></th>
<th>SOV</th>
<th>OSV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Questioned</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>Object Questioned</td>
<td>26</td>
<td>–</td>
</tr>
</tbody>
</table>

Further corroborated by a web-based acceptability judgement test (Jabeen 2017).

⇒ Default information focus position is immediately preverbal.
Scrambling of Wh-Constituents

- Besides the default position, wh-words can appear anywhere in the clause:
  1. They have exactly the same kind of scrambling possibilities as normal NPs (Manetta 2012).
  2. **But:** there is a difference in interpretation which has to do with information structure.

- More research needs to be done on this (e.g., see Butt et al. 2016).

- Focus of this talk:
  - the multifunctional uses of *kya* ‘what’
    - wh-constituent questions
    - polar questions
    - alternative questions
  - and the prosody-syntax interface
Uses of *kya* ‘what’

Thematic wh-word ‘what’

1. **As a wh-constituent**

   (3)  \( \text{sita=ne} \quad \text{kya} \quad \text{dek}^h-a \quad t^h-a \)  
   \( \text{Sita.F=Erg what see-Perf.M.Sg be.Past-M.Sg} \)  
   ‘What had Sita seen?’

2. **Within an NP**

   (4)  \( \text{shahina=ne} \quad \text{naz=ko} \quad [\text{kya} \quad \text{tofa}] \quad \text{di-ya?} \)  
   \( \text{Shahina.F=Erg Naz.F=Dat what present.M.Sg.Nom give-Perf.M.Sg} \)  
   ‘What gift did Shahina give to Naz?’
Uses of *kya* ‘what’

- Wh-counterpart of the scope marking construction (Dayal 1996, 2000)
- Licenses matrix scope of wh-in-situ

(5)  

a. sita ye soc-ti hai [ki ram ja-ye-ga]  
Sita.F.Nom this think-Impf.F.Sg be.Pres.3.Sg that Ram go-3.Sg-Fut-M.Sg  
‘Sita thinks that Ram will go.’  
(lit.: Sita thinks this, that Ram will go.)

b. sita kya soc-ti hai [ki kon ja-ye-ga?]  
Sita.F.Nom what think-Impf.F.Sg be.Pres.3.Sg that who go-3.Sg-Fut-M.Sg  
‘Who does Sita think will go?’  
(lit.: What does Sita think, that who will go?)
Polar *kyā* also interacts with alternative questions (Han and Romero 2004, Bhatt and Dayal 2014)

(6) *(kyā) candra=ne kofi p-i ya cai?*
what Chandra.F=Erg coffee.F.Nom drink-Perf.F.Sg or tea.F.Nom
‘Did Chandra drink tea or coffee?’

- The existing discussion on this interaction is already quite complex.
- There is an interesting unresolved puzzle involving the interaction of polar *kyā* with alternative questions (next slide).
- Focus of this talk: polar *kyā*.
Polar kya and Alternative Questions

- **Interesting Puzzle:** Bhatt and Dayal (2014) show that when polar kya is initial, one can get two readings with sentences containing ‘or’.

  (7) \[\text{kya} \quad \text{candra}=\text{ne} \quad \text{kofi} \quad \text{ya cai} \quad \text{p-i?}\]
  what Chandra.F=Erg coffee.F.Nom or tea.F.Nom drink-Perf.F.Sg
  ‘Did Chandra drink tea or coffee?’
  Alternative Question Reading: Did Chandra drink tea or did she drink coffee?
  Polar Question Reading: Is it the case that Chandra drank either tea or coffee?’

- But when polar kya is final, the alternative question reading is out.

  (8) \[\text{candra}=\text{ne} \quad \text{kofi} \quad \text{ya cai} \quad \text{p-i} \quad \text{kya?}\]
  Chandra.F=Erg coffee.F.Nom or tea.F.Nom drink-Perf.F.Sg what
  ‘Did Chandra drink tea or coffee?’
  *Alternative Question Reading: Did Chandra drink tea or did she drink coffee?
  Polar Question Reading: Is it the case that Chandra drank either tea or coffee?’
Polar Questions

- Urdu/Hindi has basic SOV word order.
- Question vs. declarative status is signaled via intonation (with some variation).

**Declarative:** Intonational phrase boundary is L-L%

(9) \(\text{\text{shahina}=ne \text{norina}=ko \text{mara}_{L-L}\%}\)

\[\text{Shahina.F=Erg Norina.F=Acc hit-Perf.M.Sg} \]

‘Shahina hit Norina.’

(Declarative)

**Polar Question:** Intonational phrase boundary is L/H-H%

(10) \(\text{\text{shahina}=ne \text{norina}=ko \text{mara}_{L/H-H}\%}\)

\[\text{Shahina.F=Erg Norina.F=Acc hit-Perf.M.Sg} \]

‘Did Shahina hit Norina?’

(Polar Question)
**Polar Questions**

**Figure:** $F_0$ contour of a string identical declarative and polar question.
Polar kya ‘what’

- Polar questions can optionally be expressed with *kya* ‘what’.

  (11) \( (\textit{kya}) \text{ sahina}=ne \text{ norina}=ko \text{ mara} ? \)


  ‘Did Shahina hit Norina?’

- Grammars and previous literature report polar *kya* as appearing only clause initially in Urdu/Hindi.
In contrast, Bhatt and Dayal (2014) point out that it can appear anywhere in the clause.

(12) (kya) anu=ne (kya) uma=ko (kya) kitab (%kya) d-i what A.F=Erg what U.F=Dat what book.F.Sg.Nom what give-Perf.F.Sg (kya)? what ‘Did Anu give a/the book to Uma?’

However, it is strongly dispreferred in immediately preverbal position.

Hypothesis: this is because the immediately preverbal position is the default position for:

- focus
- and therefore wh-constituent questions.
Polar kya ‘what’

- Bhatt&Dayal establish that polar *kya* is NOT a question marker.
  - It is optional in matrix clauses.
  - Generally disallowed in embedded clauses (complements of "rogative" predicates like 'wonder' and 'ask' are an exception).

- **Current State of Our Art** (Biezma et al. 2018)
  - Polar *kya* is a focus sensitive item which serves to constrain the set of possible answers viable in the context of an utterance.
  - It imposes restrictions on what the question is about.
  - Polar *kya* questions convey some assumptions regarding the possible answers that plain information-seeking questions do not convey.
Context: A group of teachers is putting together the costumes for the kids’ end of the year play. There is a list of things that each kid needs and it’s not clear who is going to be able to bring what for each kid or if they are going to be able to find everything.

(13) A: Did Ravi bring a hat for Amra?
B: #Why are you asking about a hat and not about Amra?
B’: #Why are you asking about Amra and not about a hat?
Speaker’s assumptions — Urdu Plain Polar

Context: A group of teachers is putting together the costumes for the kids’ end of the year play. There is a list of things that each kid needs and it’s not clear who is going to be able to bring what for each kid or if they are going to be able to find everything.

(14) A: ravi amra=ke liye ṭopi la-ya?
   ‘Did Ravi bring a hat for Amra?’

B: #tum ṭopi=ke bare=mē kyū puc rah-i ho,
   you.Fam.Nom hat.M.Sg=about=in why ask Prog-F.Sg be.Pres.2.Sg
   amra=ke bare=mē kyū nahi?
   Amra=about=in why not
   ‘Why are you asking about a hat and not about Amra?’

B’: #tum amra=ke bare=mē kyū puc rah-i ho,
   you.Fam.Nom Amra=about=in why ask Prog-F.Sg be.Pres.2.Sg
   ṭopi=ke bare=mē kyū nahi?
   hat.M.Sg=about=in why not
   ‘Why are you asking about Amra and not about a hat?’
Context: A group of teachers is putting together the costumes for the kids’ end of the year play. There is a list of things that each kid needs and it’s not clear who is going to be able to bring what for each kid or if they are going to be able to find everything.

(15) A: ravi amra=ke liye kya ṭopi la-ya?
‘Did Ravi bring a hat for Amra?’

B: ✓ tum ṭopi=ke bare=mē kyū puc rah-i ho,
you.Fam.Nom hat.M.Sg=about=in why ask Prog-F.Sg be.Pres.2.Sg
amra=ke bare=mē kyū nahi?
Amra=about=in why not
‘Why are you asking about a hat and not about Amra?’

B’: # tum amra=ke bare=mē kyū puc rah-i ho,
you.Fam.Nom Amra=about=in why ask Prog-F.Sg be.Pres.2.Sg
ṭopi=ke bare=mē kyū nahi?
hat.M.Sg=about=in why not
‘Why are you asking about Amra and not about a hat?’
Speaker’s assumptions — Urdu Polar kya

Context: A group of teachers is putting together the costumes for the kids’ end of the year play. There is a list of things that each kid needs and it’s not clear who is going to be able to bring what for each kid or if they are going to be able to find everything.

(16) A: ravi kya amra=ke liye ṭopi la-ya?
‘Did Ravi bring a hat for Amra?’

B: #tum ṭopi=ke bare=mē kyū puc rāh-i ho,
you.Fam.Nom hat.M.Sg=about=in why ask Prog-F.Sg be.Pres.2.Sg
amra=ke bare=mē kyū nāhī?
Amra=about=in why not
‘Why are you asking about a hat and not about Amra?’

B’: ✓tum amra=ke bare=mē kyū puc rāh-i ho,
you.Fam.Nom Amra=about=in why ask Prog-F.Sg be.Pres.2.Sg
ṭopi=ke bare=mē kyū nāhī?
hat.M.Sg=about=in why not
‘Why are you asking about Amra and not about a hat?’
Our Current Proposal (Biezma et al. 2018)

- We build on Biezma and Rawlins (2012), which bridges (Hamblin) semantics and discourse using the Q(uestion)U(nder)D(iscussion) discourse model.

- According to Biezma and Rawlins (2012), polar questions
  a. state that the content proposition is a possible answer
  b. inquire whether the content proposition holds (its semantics is merely the singleton set)
  c. require that other alternatives are available in discourse.

- Polar kya further conventionally imposes that the possible answers be a subset of the focus alternatives of the utterance.
Our Current Proposal (Biezma et al. 2018)

Polar questions denote singleton sets (based on Biezma and Rawlins 2012), see also (Roberts 1996, Farkas and Bruce 2010, a.o.)

(17) \[
\llbracket [Q \alpha] \rrbracket^c = \llbracket \alpha \rrbracket^c
\]
defined only if
a. \(\llbracket \alpha \rrbracket^c \subseteq \text{QUD}(M_{\alpha})\)
b. \(|\llbracket \alpha \rrbracket^c \cup \text{QUD}(M_{\alpha})| > 1\)

Polar kya-questions

(18) \[
\llbracket [Q [\_ \_ kya m_F \_ \_]] \rrbracket^c = \llbracket [\_ \_ m \_ \_] \rrbracket^c
\]
defined only if
a. \(\llbracket [\_ \_ m \_ \_] \rrbracket^c \subseteq \text{QUD}(M_{kya})\)
b. \(|\llbracket [\_ \_ m \_ \_] \rrbracket^c \cup \text{QUD}(M_{kya})| > 1\)
c. \(\text{QUD}(M_{kya}) \subseteq \llbracket [\_ \_ m_F \_ \_] \rrbracket^f\)
An Example:

- Polar-*kya* is a focus sensitive question operator.
- It constrains the alternatives that the speaker is entertaining.

(19)  
Ravi gave *kya* [a toy]$_F$ to Amra?  
\[ \text{[Ravi gave *kya* [a toy]$_F$ to Amra?]}^c = \{\text{Ravi gave a toy to Amra}\} \]

defined only if

a. \(\{\text{Ravi gave a toy to Amra}\} \subseteq \text{QUD}(M_{kya})\)

b. \(\{|\{\text{Ravi gave a toy to Amra}\} \cup \text{QUD}(M_{kya})| > 1\)\)

c. \(\text{QUD}(M_{kya}) \subseteq \{\text{Ravi gave a toy to Amra; } \text{Ravi gave a book to Amra; } \text{Ravi gave a game to Amra; } \ldots\} \)

\(\approx\) What did Ravi give to Amra?
Further Predictions — The Answer “Nothing”

- The ‘topic’/QUD has to be regarding what Ravi gave to Amra in (20).

\[(20)\quad ri\bar{v}i=ne\quad amra=ko\quad k\bar{y}a\quad k^{h}i\bar{l}o\bar{n}a\quad di-ya?\]
Ravi=Erg Amra=Dat what toy.M.Sg.Nom give-Perf.M.Sg
‘Did Ravi give a toy to Amra?’

- The QUD conventionally enforced with a polar \textit{kya}-question entails that Ravi gave something to Amra in (20).

Given the constraints (conventionally) imposed by polar \textit{kya}-questions on the QUD, we rule out the possibility of having ‘Ravi didn’t give anything to Amra’ as an answer to (20).
Non-serious Invitations

A doesn’t feel like offering coffee to their visitor and wishes the offer to be declined. If the speaker is not afraid of this coming across, s/he wouldn’t utter the *kya*-question:

(21)  (kya) ap (kya) coffee l-ē-g-e?
what you.Hon what coffee.F.Sg take-2.Pl-Fut-M.Pl
‘Will you have coffee?’

The polar *kya*-question excludes *nothing*. 
Surprise, incredulity...

- A corpus study (Bollywood movies, WhatsApp) has yielded the observation that polar *kya* questions tend to be used in situations when an extra pragmatic import is to be conveyed (e.g., rhetoricity, sarcasm, surprise, ...).

- We see these as derivative, following from the analysis of polar *kya* as a focus sensitive operator.

(22)  \begin{align*}
\text{kya} & \text{ ye sac hai?} \\
\text{what} & \text{ this true be.Pres.3Sg} \\
\text{‘Could this be true?’} & \text{ } \text{Script, Socha Na Tha}
\end{align*}

(23)  \begin{align*}
\text{acc}^{h-i} & \text{ nahi lag-i kya?} \\
\text{good-F.Sg. not attach-Perf.F.Sg what} & \text{ } \text{WhatsApp}
\end{align*}

*kya* can associate with the entire proposition.
Sarcasm

(24) Context: A is telling B how to behave in a situation. B says (with sarcasm):

B: (kya) tum (?)kya) meri ammā ho (kya)?
what you what my mother.F.Sg be.Pres what
‘Are you my mother?’

kyā can associate with the entire proposition.
Bhatt and Dayal (2014) invoke given vs. new in their analysis and see the polar kya as a question operator that interacts with topicalization.

Syed and Dash (2017) compare polar ‘what’ across Hindi, Bangla and Oriya and also see polar ‘what’ as a focus sensitive operator.

- Both approaches treat polar ‘what’ on a par with plain polar questions — not aware of the extra constraints signaled about the speaker’s assumptions.
- Neither proposal seriously factors in the prosodic dimension.
Focus Sensitivity and Prosody vs. Syntax

Bhatt and Dayal argued for the following descriptive generalization.

- Polar *kya* appears to partition a clause into given vs. new (cf. the “watershed” idea of Krivonosov 1977, Grosz 2016).
- Per default, material to the left of polar *kya* is taken as given and not available for being questioned (Bhatt and Dayal 2014).

\[(25) \quad \begin{align*}
\text{A.} \quad & \text{anu=ne} \quad \text{kya} \quad \text{uma=ko} \quad \text{tofa} \\
& \text{Anu.F=Erg what Uma.F=Dat present.M.Sg.Nom} \\
& \text{di-ya?} \\
& \text{give-Perf.M.Sg} \\
& \text{’Did Anu give a/the present to Uma?’} \\
\text{B.} \quad & \#nah\̄i, \quad \text{asim=ne} \quad \text{di-ya} \\
& \text{no} \quad \text{Asim.M=Erg give-Perf.M.Sg} \\
& \text{’No, Asim did.’}
\end{align*}\]

- However, this generalization only seems to be true with “default” prosody.
Prosody vs. Syntax in Focus Sensitivity

It turns out that when a constituent to the left of polar *kya* is stressed, it can be questioned.

\[(26) \quad \text{A. } \text{anu}=\text{ne}_{\text{stressed}} \text{ kya } \text{uma}=\text{ko} \quad \text{tofa} \quad \text{di-ya?} \\
\text{Anu.F=Erg} \quad \text{what Uma.F=Dat} \quad \text{present.M.Sg.Nom} \quad \text{give-Perf.M.Sg} \\
\text{’Did ANU give a/the present to Uma?’} \\
\text{B. } \text{nahi, asim}=\text{ne} \quad \text{di-ya} \\
\text{no} \quad \text{Asim.M=Erg} \quad \text{give-Perf.M.Sg} \\
\text{’No, Asim did.’} \]

**Distribution of polar *kya* and interpretations:**

- **Syntactic position:** the default is to interpret these utterances as asking about the constituent in its immediate scope.
- **Prosody:** prosodic focus marking overrides syntactic encoding – polar *kya* associates with the prosodically focused element.
Taking Stock

- We have presented a pragmatic analysis of polar *kya*.
- This work is on-going (Biezma et al. 2018, Bhatt and Dayal 2017) and is being extended to account for cross-linguistic patterns (Turkish, Sinhala, Albanian) together with project P5 (Romero, Meertens).

**On To Do List:**

- Investigate the idea of one underlying *kya* ‘what’ that is overtly realized either as a focus sensitive operator in polar questions vs. *wh*-questions → aiming at unified account of both
- Delve more deeply into the prosody of Urdu/Hindi.
- Develop a computational, psycholinguistically motivated model of the complex interface.

Next part of talk — focus on disambiguating polar *kya* vs. *wh*-question *kya* via such a model.
Some utterances are ambiguous between polar kya and wh-constituent questions.

(27) mē kya bol-ū?
I.Nom what speak-1.Sg
Constituent Question: ‘What should I say?’
Polar Question: ‘Should I say (something)?’

Script, Ankhon Dekhi

(28) kya taklif ho rah-i hai [...]?
what bother.Nom be Prog-F.Sg be.Pres.3.Sg
Constituent Question: ‘What’s bothering (you)?’
Polar Question: ‘Is something bothering (you)?’

Script, Ankhon Dekhi
However, the strings are prosodically distinct.

(29) a b k y a m a fi m ā g - ē t u m = s e ?
now what forgiveness. M. Sg. Nom ask- Pl you. Fam = Inst
‘It’s no use apologizing now.’
‘Am I supposed to ask for your forgiveness now?’

Script of Ankhon Dekhi

So — prosodic information crucial for the overall analysis.

Questions

- What is the relevant prosodic information?
- How should it be integrated into the analysis?
Polar *kya* ‘what’

- Prosodic investigations show that polar *kya* always has a flat (or falling) intonation.
- Contrast between plain polar question (left) and *kya* polar (right).
Prosodic Information

- In Urdu/Hindi, focused/contrastively stressed items can be marked by:
  - A larger pitch excursion on the L H pattern found generally on all prosodic phrases (Genzel and Kügler (2010)).
  - Longer duration of the focused constituent (Genzel and Kügler 2010, Jabeen and Braun 2017)
  - Pitch compression after the focused/stressed element (Patil et al. 2008, Jabeen 2017)

- kya ‘what’
  - The thematic wh-word kya has a high tone: H*.
  - The polar kya is always flat or falling.

- Boundary Intonation
  - Polar questions: L/H-H%
  - Declaratives and constituent-questions: L-L% (with some variation).
We follow the general syntactic analyses as established as part of the Urdu grammar (Butt and King 2007).

For the syntax, strings like the following are ambiguous.

(30) alina=ne zain=ko kya tofa di-ya th-a?
    Alina=Erg Zain=Acc what present.M.Sg give-Perf.M.Sg be.Past-M.Sg
Constituent Question: ‘What gift did Alina give to Zain?’
Polar Question: ‘Did Alina (actually) give a gift to Zain?’

Following Slade (2011), we analyze the kya as a Q.
We assume one underspecified kya ‘what’ for the polar and the wh-readings.
One String — Two Possible Analyses

**Wh-Question**

```
S
  /\   /
 /   / \   
 KP  KP  NP  VC
  |     |    |
 alina=ne zain=ko Q N V Aux
    |    |   |
    kya tofa diya tʰa
```

**Polar kya**

```
S
  /\   /
 /   / \   
 KP  KP  Q  NP  VC
  |     |    |
 alina=ne zain=ko kya N V Aux
    |    |   |
    tofa diya tʰa
```
We show how polar vs. constituent *kya* can be disambiguated via an integration of the prosodic information.

The analysis is based on the syntax-prosody interface for LFG developed in Bögel (2015).

Initial LFG proposals for the p-structure were “syntactocentric” (cf. Jackendoff 2002).

Newer proposals have moved to seeing prosody as a separate level of representation that interacts with morphosyntax, but is not derived from it (e.g., Bögel 2015, Dalrymple and Mycock 2011, Dalrymple and Nikolaeva 2011).
LFG Architecture

- There are two syntactic representations in LFG.
  - c(onstituent)-structure: represents linear order, hierarchical relationships and constituency
  - f(unctional)-structure: represents basic predicate-argument relations and functional information

- Below is a simplified analysis (Butt and King 2015).

(1) a. Yassin will watch the movie.

   b. **c-structure**

   
   ```
   S
   / | \\
   NP VP
   /   |
   Yassin AUX
   /     |
   will VP
   /       |
   V watch
   /       
   DET N
   the movie
   ```

   c. **f-structure**

   ```
   [PRED 'watch<SUBJ,OBJ>']
   [SUBJ [PRED 'Yassin']
   [OBJ [PRED 'movie']
   [TENSE future]]
   ```
LFG Architecture

- LFG has a *projection architecture*.
- The different levels of representation are related to each other via mathematically defined projections.
- c-structure and f-structure are related to one another by the $\phi$-projection, realized below via f-structural annotations on c-structure.
  
  a. $S \rightarrow \text{NP} \quad \text{VP}$
    $\quad (\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow$
  
  b. $\text{VP} \rightarrow \text{AUX} \quad \text{VP}$
    $\quad (\uparrow \text{TENSE}) = \downarrow \quad \uparrow = \downarrow$
  
  c. $\text{VP} \rightarrow \text{V} \quad \text{NP}$
    $\quad \uparrow = \downarrow \quad (\uparrow \text{OBJ}) = \downarrow$
Each piece of the c-structure thus contributes information to the f-structure.

(2) Yassin will watch the movie.

The f-structure provides the main basis for further semantic analysis.

Within LFG, glue semantics is currently popular.
LFG’s Projections

Over the years, more projections than the original core c-structure, f-structure and s(emantic)-structure have been argued for:

- a rgument)-structure: place for thematic roles and information about predicate composition (complex predicates)
- i nformation)-structure: place for information structural components (inspired mainly by Vallduví 1992).

The architecture of LFG allows for complex interactions across projections.

We analyse kya at the prosody–syntax interface following the proposal made by Bögel (2015):
The Prosody-Syntax interface

Two perspectives:
(Roughly following models as proposed by, a.o., Levelt (1999) and Jackendoff (2002))

- **Production**: from meaning to form (syntax → prosody)
- **Comprehension**: from form to meaning (prosody → syntax)

\[\text{production} \downarrow \]
\[\text{comprehension} \uparrow\]

\[\begin{array}{c}
\text{c-structure} \\
\downarrow \pi \\
\text{string} \\
\downarrow \rho \\
\text{p-structure} \\
\end{array} \]

\[\begin{array}{c}
\text{Lexicon} \\
\end{array} \]

\[\begin{array}{c}
\downarrow \\
\text{production} \\
\end{array} \]

- \(\pi\): The Transfer of structure → Information on (larger) syntactic and prosodic phrasing, and on intonation is exchanged

- \(\rho\): The Transfer of vocabulary → Associates morphosyntactic and phonological information on lexical elements and projects them to their respective structures
P-structure – coming from comprehension

- Input: the ‘raw’ speech signal information
- Linearly represented in the p-diagram:

<table>
<thead>
<tr>
<th>INDEX</th>
<th>S_1</th>
<th>S_2</th>
<th>S_3</th>
<th>S_4</th>
<th>S_5</th>
<th>S_6</th>
<th>S_7</th>
<th>S_8</th>
<th>S_9</th>
<th>S_{10}</th>
<th>S_{11}</th>
<th>S_{12}</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>VALUE</td>
<td>[æ]</td>
<td>[li]</td>
<td>[na]</td>
<td>[ne]</td>
<td>[zæn]</td>
<td>[ko]</td>
<td>[kja]</td>
<td>[t̚o]</td>
<td>[fa]</td>
<td>[di]</td>
<td>[ja]</td>
<td>[tʰa]</td>
<td></td>
</tr>
<tr>
<td>DUR.</td>
<td>0.07</td>
<td>0.17</td>
<td>0.16</td>
<td>0.15</td>
<td>0.28</td>
<td>0.13</td>
<td>0.25</td>
<td>0.23</td>
<td>0.13</td>
<td>0.13</td>
<td>0.11</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>F₀</td>
<td>177</td>
<td>183</td>
<td>204</td>
<td>216</td>
<td>181</td>
<td>177</td>
<td>205</td>
<td>188</td>
<td>166</td>
<td>164</td>
<td>140</td>
<td>136</td>
<td>↓</td>
</tr>
</tbody>
</table>

→ Structured syllablewise
→ Each syllable is part of a vector, which associates the syllable with its values for duration or F₀
⇒ Values can be interpreted in a more categorical way
P-structure – coming from comprehension

- Categorical interpretation on the basis of ‘raw’ information:

<table>
<thead>
<tr>
<th>PHRAS.</th>
<th>(t ... ... ... ... ... ... ... ... )ₜ</th>
<th>INTERPRETATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOBI</td>
<td>... ... ... ... ... H* ... ... ... ... ...</td>
<td>↓</td>
</tr>
<tr>
<td>DUR.</td>
<td>0,07 0,17 0,16 0,15 0,28 0,13 0,25 0,23 0,13 0,13 0,11 0,14</td>
<td>SIGNAL</td>
</tr>
<tr>
<td>F₀</td>
<td>177 183 204 216 181 177 205 188 166 164 140 136</td>
<td>↓</td>
</tr>
<tr>
<td>VALUE</td>
<td>[ə] [li] [na] [ne] [zæn] [ko] [kja] [t̚o] [fa] [di] [ja] [tʰa]</td>
<td></td>
</tr>
</tbody>
</table>
| INDEX  | S₁  S₂  S₃  S₄  S₅  S₆  S₇  S₈  S₉  S₁₀ S₁₁ S₁₂ | ...

- Pauses, patterns in F₀ and other acoustic indicators can be further interpreted

→ Includes language-specific prosodic/phonological readjustments

⇒ Content of p-diagram is then matched against the multidimensional lexicon (→ *Transfer of vocabulary*)
The Transfer of Vocabulary

The acoustic information from the speech signal is matched against the p-form of the multidimensional lexicon.

<table>
<thead>
<tr>
<th>concept</th>
<th>s(yntactic)-form</th>
<th>p(honological)-form</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘gift’</td>
<td>N (↑ PRED) = ‘tola’  &lt;br&gt; (↑ NUM) = sg  &lt;br&gt; (↑ GEND) = masc</td>
<td>SEGMENTS /t o f a/  &lt;br&gt; METR. STRUC. σ σ</td>
</tr>
<tr>
<td>Q (↑ INT-FORM) = kya</td>
<td>SEGMENTS /k j a/  &lt;br&gt; METR. STRUC. σ</td>
<td></td>
</tr>
</tbody>
</table>

- Each lexical dimension can only be accessed by the related module
  → Modular: strict separation of module-related information
  → Translation function: Once a dimension is triggered, the related dimensions can be accessed as well.
  ⇒ Associated s-form is selected and made available to c-structure.
The Transfer of Structure ...

- Relates information on intonation and prosodic domains given in p-structure to c-structure (*comprehension*)
- Relates (a.o.) information on syntactic constituents from c-structure to p-structure (*production*)
- **Comprehension:**
  - (constituent *kya*)

\[
(\tau(T(\ast)) \ S \ \text{ToBI}) =_c H^* 
\]

\[
(\tau(\rho^{-1})) \ \text{ToBI} \quad \text{VALUES} \\
\text{ToBI} \quad \cdots \quad H^* \quad \cdots \quad \cdots \\
\text{...} \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\text{VALUES} \quad [\ldots] \quad [\ldots] \quad [\ldots] \quad [\ldots] \\
\text{VECTORINDEX} \quad S_3 \quad S_4 \quad S_5 \quad S_6
\]

\((\tau(\rho^{-1})), \ T(\ast) = \text{all terminal nodes under this node}, =_c = \text{constraining equation, } S=\text{Syllable})\)
kya as a constituent question (ToBI $\equiv_c H^*$)

**Lexicon**

<table>
<thead>
<tr>
<th>CONCEPT</th>
<th>S-FORM</th>
<th>P-FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUESTION</td>
<td>kya</td>
<td>Q</td>
</tr>
<tr>
<td>‘GIFT’</td>
<td>tofa</td>
<td>N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segments</th>
<th>/k j a/</th>
<th>/t o f a/</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>ToBI</th>
<th>...</th>
<th>H*</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>L-L%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUR.</td>
<td>...</td>
<td>0,25</td>
<td>0,23</td>
<td>0,13</td>
<td>0,13</td>
<td>0,11</td>
<td>0,14</td>
</tr>
<tr>
<td>F$^0$</td>
<td>...</td>
<td>205</td>
<td>188</td>
<td>166</td>
<td>164</td>
<td>140</td>
<td>136</td>
</tr>
<tr>
<td>VALUE</td>
<td>...</td>
<td>[kja]</td>
<td>[toh]</td>
<td>[fa]</td>
<td>[di]</td>
<td>[ja]</td>
<td>[t$^h$a]</td>
</tr>
<tr>
<td>INDEX</td>
<td>...</td>
<td>S$_7$</td>
<td>S$_8$</td>
<td>S$_9$</td>
<td>S$_{10}$</td>
<td>S$_{11}$</td>
<td>S$_{12}$</td>
</tr>
</tbody>
</table>
kya as a polar question (ToBI $\sim = H^*$)

Lexicon

<table>
<thead>
<tr>
<th>CONCEPT</th>
<th>S-FORM</th>
<th>P-FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUESTION</td>
<td>kya</td>
<td>Q</td>
</tr>
<tr>
<td>‘GIFT’</td>
<td>tofa</td>
<td>N</td>
</tr>
</tbody>
</table>

ToBI

| ... | ... | ... | ... | ... | ... | ... |

DUR.

| ... | 0,17 | 0,25 | 0,15 | 0,12 | 0,14 | 0,17 |

$F_0$

| ... | 185  | 176  | 213  | 264  | 255  | 250  |

VALUE

| ... | [kja] | [tɔh] | [fa] | [di] | [ja] | [tʰa] |

INDEX

| ... | $S_7$ | $S_8$ | $S_9$ | $S_{10}$ | $S_{11}$ | $S_{12}$ | ... |
For semantic/pragmatic analysis, the following information needs to have become available via the syntax-prosody interface:

1. The fact that it is a polar question
   QUESTION-TYPE polar

2. The fact that there was a *kya*
   INT-FORM kya

3. What material the polar *kya* can be associated with:
   - Material to its right, in particular the constituent on its immediate right (via f-precedence and right sister).
   - Material that is prosodically stressed (via a Metarulemacro that checks for each constituent whether it was stressed via the prosody-syntax interface).
   - Preferences for association (via integration of OT-style constraints).
Prosody directly indicates meaning: no overt syntactic element implicated (cf. Steedman 2014).

A general type of solution to this irrelevance of syntax has been to postulate some kind of null element or operator in the syntax.

For example, Bhatt and Dayal (2014) for Urdu/Hindi polar kya, following Han and Romero (2004).

\[ cp \text{ Null-Yes/No-Operator [ IP ]} \]

We do not need to postulate a null question operator in the syntax.

Rather than “reconstructing” the effect of prosody in the syntax via empty elements so that semantic interpretation can proceed correctly, we integrate the relevant prosodic information directly.
Summary and Conclusions

- Analysis of various uses of *kya* ‘what’ in Urdu/Hindi.
- See *kya* ‘what’ as a lexically (massively) **underspecified** item. → Allows for various uses in the Urdu/Hindi grammar

- This talk:
  - Polar *kya* vs. constituent question (specifier) *kya*.
  - Resolution of ambiguity via prosody.
  - Semantics/Pragmatics analysis based on Biezma et al. (2018).

- But much more needs to be done!
  - Interaction with Alternative Questions
  - Crosslinguistic Comparison (Turkish, Sinhala)
  - Word order variation in constituent questions
Thanks!

Very many thanks go to Rajesh Bhatt and Veneeta Dayal for the original inspiration and some further discussions, Ghulam Raza for help with the data, suggestions, general pointers and interesting discussions and to Bettina Braun, Regine Eckardt, Gillian Ramchand, Craige Roberts and Maribel Romero for helping us to come to grips with the phenomena and to María Biezma for in-depth cooperation. Mary Dalrymple has been helping us with f-precedence (implementation vs. theory). Many thanks go to Habiba, who has been one of our main informants.
References I


Butt, Miriam and Tracy Holloway King. 1997. Null elements in discourse structure. Written to be part of a volume that never materialized.


References II


References III


The Transfer of Structure ... from syntax to prosody

- where \( S_{\text{min}} \) refers to the *first* syllable within the scope of a node
- where \( S_{\text{max}} \) refers to the *last* syllable within the scope of a node, for example: 

  \[
  \left( \hat{\tau}(T(*)) S_{\text{max}} \text{ Prosodic Phrasing} \right) = \rightbracket_i
  \]

→ In the case of constituent *kya*, \( Q \) would be annotated with:

  \[
  \left( \hat{\tau}(T(*)) S \text{ ToBI} \right) = H^*
  \]