

Questions and Information Structure in Urdu/Hindi

Miriam Butt
University of Konstanz
miriam.butt@uni-konstanz.de

LFG14
University of Michigan, Ann Arbor
July 2014

1 Introduction

Overall goal: Understand word order variation in questions

- not via syntax-specific movement triggers
- but via an integration of i(nformation)-structural concerns
(*information packaging*; Chafe (1976), Vallduví (1992), Krifka (2008))

Information Structure and Questions:

- Word order in Hindi/Urdu is known to be correlated with i-structure (e.g., Gambhir 1981, Kidwai 2000, Butt and King 1996, 1997).
- Basic Idea: leverage this knowledge towards understanding pragmatic effects found with word order variation of wh-elements (or “k-elements”) in Hindi/Urdu.

Phenomena: Focus particularly on constituent questions and polar *kya* ‘what’.

Analysis Ingredients:

- **Constituent Questions:** Word Order variation indicates strategies for *Common Ground Management* in the sense of Krifka (2008).
- **Polar *kya* ‘what’:**
 - in initial position it functions as a question marker/particle
 - word order variation partitions a clause into a focus part and a background part (as proposed by Krifka (1992) in a structured meaning approach to focus), whereby the background part is not available for questioning.

2 Some Language Background

2.1 Word Order

- The default word order in Urdu/Hindi is **SOV**.
- Major constituents can scramble.
- **Topics** generally initial, **focus** immediately preverbal.
- Genitives can also scramble (like quantifier float) — know no good analysis.
- Unlike in Japanese and Korean, for example, material can appear after the verb.

2.2 Pro-Drop

- Any argument can be dropped (no licensing via agreement or other morphological or syntactic means).
- Pro-drop licensed by context: the center of discourse can be dropped (and objects can only be dropped if subjects are also dropped), see Prasad (2000, 2003) for details.

3 Question Formation

3.1 Polar Questions — Take 1

Polar questions in Urdu/Hindi generally take the syntactic form of declaratives. Intonation is used to distinguish between declaratives and interrogatives.

- (1) a. anu=ne uma=ko kitab d-i
 Anu.F=Erg Uma.F=Dat book.F.Sg.Nom give-Perf.F.Sg
 ‘Anu give a/the book to Uma. declarative
- b. anu=ne uma=ko kitab d-i?
 Anu.F=Erg Uma.F=Dat book.F.Sg.Nom give-Perf.F.Sg
 ‘Did Anu give a/the book to Uma? polar question

3.2 Constituent Questions

Urdu/Hindi has traditionally been characterized as a **wh-in-situ** language.

- (2) a. sita=ne d^hyan=se **ram=ko** dek^h-a t^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^hyan=se **kis=ko** dek^h-a t^h-a?
 Sita.F=Erg carefully who.Obl=Acc see-Perf.M.Sg be.Past-M.Sg
 ‘Who had Sita looked at carefully?’

But: **default** position for questions is actually the **preverbal** focus position (for topic/focus analyses of Urdu/Hindi, see Gambhir 1981, Butt and King 1996, 1997, Kidwai 2000).

- (3) a. sita=ne d^hyan=se ram=ko dek^h-a t^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 **kaise** ram=ko dek^h-a t^h-a?
 Sita.F=Erg how Ram.M=Acc see-Perf.M.Sg be.Past-M.Sg
 ‘How had Sita managed to see Ram?’
 (expresses some degree of wonder)
- c. sita=ne ram=ko **kaise** dek^h-a t^h-a?
 Sita.F=Erg Ram.M=Acc how see-Perf.M.Sg be.Past-M.Sg
 ‘How had Sita looked at Ram?’ (default order for a how-question)
- (4) a. sita=ne d^hyan=se **kis=ko** dek^h-a t^h-a?
 Sita.F=Erg carefully who.Obl=Acc see-Perf.M.Sg be.Past-M.Sg
 ‘Who had Sita looked at carefully?’
- b. ram=ko d^hyan=se **kis=ne** dek^h-a t^h-a?
 Ram.M=Acc carefully who.Obl=Erg see-Perf.M.Sg be.Past-M.Sg
 ‘Who had looked at Ram carefully?’

- Besides the default position, wh-constituents can appear anywhere in the clause:
 1. they have the same kind of scrambling possibilities as normal NPs (Manetta 2012).
 2. **But:** word order variation appears to go hand-in-hand with a difference in interpretational possibilities. (cf. (3b) vs. (3c)).
 3. This has not been the subject of much discussion in the literature (see Bhatt 2003 for a comprehensive overview over the state-of-the-art then).

3.3 Movement and Scope

In the literature to date:

- The left periphery has received the most overall attention.
- In particular, the literature has focused on why, given a general assumption of covert *wh*-movement for Urdu/Hindi, *wh*-words in embedded clauses cannot take matrix scope (e.g., Mahajan 1990, Srivastav 1991, Dayal 1994, 1996, 2014, Lahiri 2002, Bhatt and Dayal 2007, Manetta 2010, 2012).

- (5) ravi jan-ta t^h-a
 Ravi.M.Nom know-Impf.M.Sg be.Past-M.Sg
 [ke sita=ne d^hyan=se **kis=ko** dek^h-a t^h-a]
 that Sita.F=Erg carefully who.Obl=Acc see-Perf.M.Sg be.Past-M.Sg
 ‘Ravi used to know who Sita had looked at carefully’
 *‘Who did Ravi used to know Sita had looked at carefully?’

Why the focus on this?

- In LF-based approaches, the *wh*-word is assumed to have to move to a position where it can act as an operator (usually SpecCP) — classic example: English overt *wh*-movement.
- Since the *wh*-word in Urdu/Hindi can stay in situ, LF-based approaches have to assume that the *wh*-word undergoes **covert movement** to the appropriate operator position.
- However, if covert movement is in principle possible, then what prevents it from applying in (5)?

I will not go into the various (movement or copy-theory) solutions proposed in the literature, see Mycock (2006) for an LFG approach.

Empirically, there are two ways of getting matrix scope for embedded *wh*-words:

1. The *wh*-constituent appears in the matrix clause (“movement/extraction”)¹
2. Via the so-called *scope marking* construction ((7)–(8)).

- (6) **kya**_i tum jan-te ho [ki us=ne —_i ki-ya]?
 what you.Pl.Nom know-Impf.M.Pl be.Pres.2.Pl [that 3.Sg.Obl=Erg do-Perf.M.Sg
 ‘What did you know that he did?’ (Srivastav 1991:766) **wh-extraction**

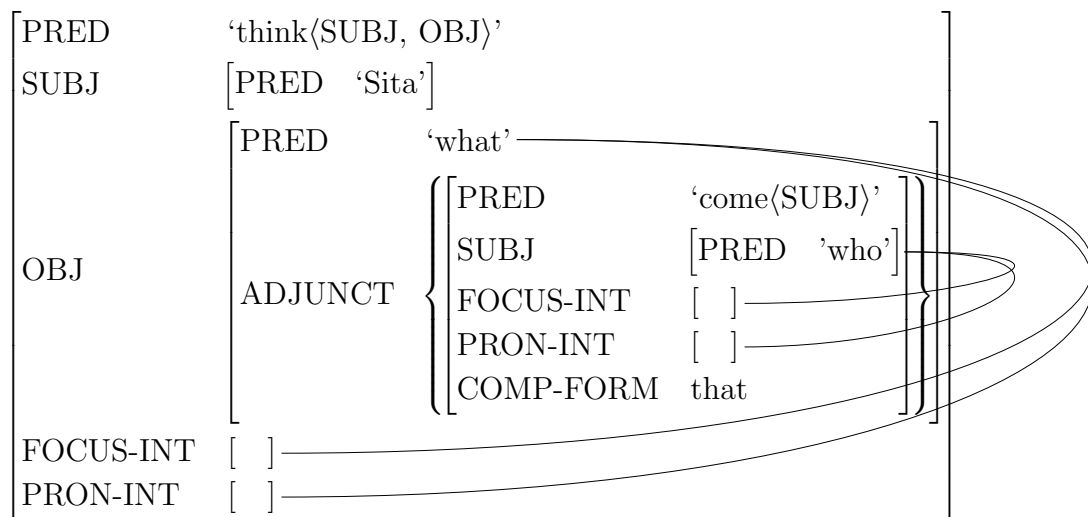
- (7) sita **kya** soc-ti hai [ki **kon** ja-ye-ga]?
 Sita.F.Nom what think-Impf.F.Sg is that who go-3.Sg-Fut-M.Sg
 Who does Sita think will go? **thematic kya**
 (Lit.: What does Sita think, that who will go?)

¹The acceptability of long-distance extraction as in (6) has been somewhat under debate. The first reported instances go back to Gurtu (1985); however, many speakers found them unacceptable. It has since been established that they are acceptable under certain intonational contours (Dayal 2014), reflecting the information structural sensitivity that is generally associated with scrambling in Urdu/Hindi). NB: discussions exist which categorically deny the existence of long distance *wh*-extraction in Hindi and adduce this as a crucial factor in understanding the nature of cyclicity in *wh*-scope (e.g., Stepanov and Stateva 2006).

(8) ravi **kya** jan-ta hε
 Ravi.M.Nom what.Nom know-Impf.M.Sg be.Pres.3.Sg
 [ke sita **kis=ko** pasand kar-ti hε]?
 that Sita.F who.Obl=Acc liking do-Impf.F.Sg be.Pres.3.Sg
 ‘Who does Ravi know Sita likes?’
 Lit.: ‘What does Ravi know, who does Sita like?’ **thematic *kya***

- In the scope marking construction, a *kya* ‘what’ is introduced in the matrix clause.
- The embedded *wh*-element remains in situ.
- Dayal has argued for an indirect dependency account by which the embedded *that*-clause is anaphorically related to the *kya* in the matrix clause, which in turn is not seen as an expletive, but as a “thematic” ‘what’, i.e., a full *wh*-element (Srivastav 1991, Dayal 1994, 1996, 2014).
- I follow Dayal’s overall indirect dependency analysis, but render the ‘that’-clause as an adjunct modifying the thematic *kya*.

Scope Marking Construction, f-structure for (7)



- The f-structure follows the analysis established for questions within ParGram (cf. <http://typo.uni-konstanz.de/redmine/projects/pargram/wiki/Questions>).

3.4 Interim Summary

Questions in Urdu/Hindi have the following properties:

- wh-in-situ
- wh-scope out of embedded clauses either via overt extraction or thematic *kya* (scope marking)
- default position for focus is immediately preverbal, so wh-constituents are also naturally found there (rather than just in-situ).

3.5 Word Order Variation in *Wh*-questions

Recall that the *wh*-constituent can appear anywhere an NP can (Manetta 2012).

In particular, the following examples have been discussed in some detail (Mahajan 1997, Bhatt and Dayal 2007, Manetta 2012).

- (9) a. sita=ne d^hyan=se **kis=ko** dek^h-a t^h-a?
 Sita.F=Erg carefully who.Obl=Acc see-Perf.M.Sg be.Past-M.Sg
 'Who had Sita looked at carefully?' (wh-in-situ/preverbal focus)
- b. sita=ne d^hyan=se dek^h-a t^h-a **kis=ko?**
 Sita.F=Erg carefully see-Perf.M.Sg be.Past-M.Sg who.Obl=Acc
 'Sita had looked at carefully at who?' (wh postverbal)
- c. sita=ne d^hyan=se dek^h-a **kis=ko** t^h-a?
 Sita.F=Erg carefully see-Perf.M.Sg who.Obl=Acc be.Past-M.Sg
 Reading 1: 'Who had Sita looked at carefully?' (wh in verbal complex)
 Reading 2: 'Who had Sita really looked at carefully?'
 (i.e., she had not looked at anybody carefully) (rhetorical question)

3.5.1 Mahajan

(9b) was first discussed by Mahajan (1997) (in a footnote).

- It presents a problem for him because (9b) is not good under a straight forward information-seeking interpretation.
- Bhatt and Dayal (2007) analyze it as an echo question and show that Mahajan's account does not provide an account of this reading.
- Mahajan assumes an anti-symmetric approach to Hindi (Kayne 1994)
- This means that the basic underlying word order for Hindi/Urdu is taken to be SVO (and not SOV).
- This in turn means that Mahajan has to assume massive leftward movement to get a *wh*-word on the right periphery as in (9b).
- The appearance of a *wh*-constituent in the verbal complex as in (9c) (first discussed by Bhatt and Dayal) can only be accounted for by means of unnatural assumptions.

3.5.2 Bhatt and Dayal

Instead, Bhatt and Dayal (2007) argue for an analysis in terms of **Rightward Remnant Movement**.

- Hindi/Urdu is back to being SOV in their account.
- Leftward movement of DPs is allowed to specifier or adjoined positions.
- Verbs can optionally move to an aspectual head.
- Rightward movement is restricted to verbal projections (VPs).
- Question formation in Hindi/Urdu involves covert movement at LF — but the wh-item can only move if it is in the relevant domain.

The postverbal wh-constituent in (9b) is analyzed as an **echo question**.

- **Explanation for Echo-Question Effect:**

The postverbal wh-constituent cannot be interpreted as a normal question because it is trapped in a remnant VP, which acts as an island.

The **wh-constituent within the verbal complex** in (9c) is a normal question.

There is no rightward movement involved and it is not trapped in an island.

Deriving [Subj V DO-wh Aux] (based on Bhatt and Dayal 2007:295)

- (10) a. Start with: [S [DO-wh V] Aux]
- b. Leftward scrambling of DO-wh \rightarrow
 [S [DO-wh_i [t_i V]] Aux]
- c. Leftward scrambling (topicalization) of V \rightarrow
 [S [[t_i V]_j [DO-wh_i t_j]] Aux]
- d. Covert wh-movement \rightarrow
 [DO-wh_i [S [[t_i V]_j [t'_i t_j]] Aux]]

Points of Concern:

- Some of the movements appear unmotivated. I.e., “short distance left-ward topicalization of the verb” (p. 295) used to get [Subj V DO-wh Aux].
- Despite using topicalization as a reason for movement, no awareness that i-structural information may play a role in general and should be integrated into the analysis.

3.5.3 Manetta

Manetta (2012) looks at a larger range of constructions than Bhatt and Dayal and argues:

- The Remnant-VP approach does not cover enough empirical ground.
- Instead, one should do the work exclusively via scrambling.
- Scrambling can be both leftwards and rightwards.
- Reasons for scrambling:
 - probe-goal relationships (i.e., certain features motivate scrambling)
 - Some features that can trigger scrambling: EPP, Q(uestion), *wh*, E(cho)

Information Structure Relevance

- Manetta refers to some existing work on information structure as well as scrambling in Urdu/Hindi (Gambhir 1981, Butt and King 1996, Dayal 2003) and concludes:
 - Topic/Focus is the result of leftward scrambling
 - Backgrounded/Old Information is the result of rightward scrambling
- Integrates this into her analysis via features on lexical items.
 - An echo reading is taken to be connected to old information status of the *wh*-word.
 - So *wh*-word in postverbal position carries an E(cho) feature.

4 Information Structure — Take 1

The following is generally agreed upon for Hindi/Urdu (Gambhir 1981, Kidwai 2000):

- Clause initial position is associated with topics.
- The immediately preverbal position is associated with focus.
- The discourse particle *to* marks preceding constituents as topics (though I think it is probably better analyzed as a contrastive topic in the sense of Krifka (2008)).
- The discourse particle *hi* serves to emphasize/focus preceding constituents.

4.1 Butt and King

- Butt and King (1996, 1997) developed a four-way i(nformation)-structure system in terms of [\pm prom(inent)] and [\pm new].
- Partitioning of a clause into four main information structural components: topic, focus, background and completive information.
- Proposal based on insights from Vallduví (1992), Choi (1996) and É. Kiss (1995).
- (Taken up by Mycock (2006, 2013)).

| Type | Definition | Position |
|------------------------|---------------------|-------------------------|
| Topic | [−New,+Prom(inent)] | Clause-Initial |
| Focus | [+New,+Prom(inent)] | Immediately Preverbal |
| Background | [−New,−Prom(inent)] | Postverbal |
| Completive Information | [+New,−Prom(inent)] | Between Topic and Focus |

- (11) [nadya]_T (to) [ab^hi]_{CI} [təfi]_{CI} [bazar=se]_F xarid
 Nadya.F.Nom indeed just now toffee.F.Nom market.M=from buy
 rah-i t^h-i [mere=liye]_B
 stay-Perf.F.Sg be.Past-F.Sg I.Gen.Obl=for
 ‘Nadya was just buying toffee at the market for me.’

4.2 More on Clausal Position and Information Structure

- Butt and King’s system was based on a small corpus study of Bollywood movie dialogs.
- The system is simple and in need of expansion (e.g., need to deal with embedded clauses, also cf. Mycock’s (2006) brief discussion).
- In fact, Gambhir (1981) showed that the sentence-final position has a multitude of functions:
 - de-emphasis (mostly involves pronominals)
 - afterthought
 - added emphasis on new information
 - TV/radio announcement style, new information sentence final
 - “heavy” items
 - to create suspense
- In light of Bhatt and Dayal (2007), an echo question reading also needs to be added.
- Gambhir (1981) also shows that the clause initial position is not always a topic, but can also be used for *scene setting*.

– I think this is *frame setting* in Krifka’s (2008) sense:

- (12) A. How is John?
 B. {Healthwise}, he is [FINE]_F.

4.3 Clausal Position of Wh-Words

4.3.1 Postverbal *wh*

Recall that the echo question reading of (13) has been analyzed as being tied to the postverbal position of the *wh*-word.

- (13) sita=ne d^hyan=se dek^h-a t^h-a **kis=ko?**
 Sita.F=Erg carefully see-Perf.M.Sg be.Past-M.Sg who.Obl=Acc
 ‘Sita had looked at carefully at who?’ (wh postverbal=echo)

Echo questions are analyzed as involving a type of focus (Artstein 2002, Truckenbrodt 2012).

- The echo reading noted by Bhatt and Dayal (2007) for postverbal *wh* would thus appear not to be able to follow from the background/old information connection, but from the category of “added emphasis/new information”.

4.3.2 *Wh* within the verbal complex

So far there is no good account of the occurrence of the *wh*-constituent within the verbal complex, as in (14).

- (14) sita=ne d^hyan=se dek^h-a **kis=ko** t^h-a?
 Sita.F=Erg carefully see-Perf.M.Sg who.Obl=Acc be.Past-M.Sg
 Reading 1: ‘Who had Sita looked at carefully?’ (wh in verbal
 Reading 2: ‘Who had Sita really looked at carefully?’
 (i.e., she had not looked at anybody carefully) (rhetorical question)

- Manetta’s account amounts to lexical stipulation.
- Bhatt and Dayal’s account involves short distance topicalization, which is not independently motivated.
- There is some kind of added emphasis on the immediately postverbal element, allowing for a rhetorical question reading — this needs to be explained.

4.3.3 Other Word Orders

Recall that these are not the only possible positions for *wh*-elements.

- They can occur anywhere that NP-constituents can.
- Additionally, they can also occur within the verbal complex, where NP-constituents cannot (as far as I know).

Some “real-life” examples from a Bollywood song

- Movie: *Kuch, Kuch Hota Hai* ‘Things Happen’
- Song: *Koi Mil Gaya* ‘Found someone’
- See Appendix for the overall context of the song.

- (15) a. **kyū** mē kʰo ga-ya hū
 why I.Nom lost go-Perf.M.Sg be.Pres.1.Sg
 ‘Why have I gotten lost?’
- b. pʰir ye becēni hε **kyū**
 then this.Nom restlessness.F.Sg.Nom be.Pres.3.Sg why
 ‘Then why is there this restlessness?’
- (16) a. nam ūs=ka hε **kya**
 name.M.Sg.Nom Pron.3.Sg.Obl=Gen.M.Sg be.Pres.3.Sg what
 ‘what her name is.’
- b. tum=ne ban-a-ya hε
 you.Sg.Fam=Erg be.made-Caus-Perf.M.Sg be.Pres.3.Sg
kya apn-a ye hal
 what self-M.Sg this state
 ‘What state have you gotten yourself into?’

Note: (16) contains discontinuous NPs. The default word orders are given in (17).

- (17) a. [ūs=ka nam] **kya** hε
 Pron.3.Sg.Obl=Gen.M.Sg name.M.Sg.Nom what be.Pres.3.Sg
 ‘what her name is.’
- b. tum=ne ye apn-a hal
 you.Sg.Fam=Erg this self-M.Sg state
kya ban-a-ya hε
 what be.made-Caus-Perf.M.Sg be.Pres.3.Sg
 ‘What state have you gotten yourself into?’

I do not at present have an analysis for all the possible clausal positions of *wh*-elements.

- Want to pursue the direction presented in the next section — leverage established connection between information structure and clausal position to explain pragmatic effects of word order variation of *wh*-constituents.
- Begin by focusing on clause-final and immediately postverbal *wh*.

5 Information Structure — Take 2

Proposal: Understand the word order variation found with *wh*-constituents not in terms syntactic movement triggers (or alternative pronouncements of various copies of a *wh*-element; Manetta (2013)) or (covert) LF-movement, but in terms of *information packaging*.

5.1 Krifka on Information Structure

Krifka (2008) proposes to cut through the plethora of different types of topic/focus posited by providing a combinatory interaction between semantics and pragmatics.

Among other things, he argues that understanding information structure in terms of features such as $[\pm\text{new}]$ is not useful.

- The use of these features is overly simplistic.
- The features do not yield the right semantic/pragmatic effects.

Instead:

- Information Structure consists of two major parts:
 1. Common Ground Content: truth conditionally relevant information
 2. Common Ground Management: pragmatics, packaging of information to fulfill communicative needs/structure the discourse in a certain way.
- The use of word order variation for information packaging falls under Common Ground Management (CG Management).
- Focus and Topic are both understood as being interpreted with respect to alternative sets (Rooth’s Alternative Focus Semantics) out of which the relevant ones are identified as part of the communication.
- “File Card Semantics” à la Heim (and Discourse Representation Theory) are used to manage information coming in from a discourse.

The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG content. (Krifka 2008:(39))

- With respect to both topic and focus, the notion of *givenness* is centrally important — essentially, this is what is (or what is assumed to be) already in the Common Ground.

5.2 Mycock's LFG Perspective

Mycock (2006) investigates *wh*-constituent questions from a typological perspective.

- Proposal of a parallel architecture in which phonology/prosody and morphosyntax can contribute to the overall interpretation of an utterance on a equal footing.
- Posits that all question phrases must occupy a position consistent with being syntactically focused.
- Integrates information structure into her overall treatment of *wh*-constituent questions.

Semantics

Mycock adopts Ginzburg and Sag's (2000) propositional abstract semantics.

- *Wh*-words introduce a parameter which is to be filled in.
- Interrogative vs. non-interrogative focus is distinguished explicitly via a feature (though this actually seems to be redundant).
- Interrogative scope is effected via a meaning constructor [**interrog-scope**].
- This meaning constructor can be introduced via the syntax (annotation on c-structure rules) or via the prosody.
- Interpreted within glue semantics.

Information Structure

- *wh*-constituents are always considered to be in focus (in agreement with Krifka)
- Butt and King's four-way distinction between topic, focus, background and completive information is adopted.
- Example (18) illustrates a typical i-structure analysis.

(18) [What]_{FOCUS} did [CHARLIE]_{FOCUS} eat?

i-structure

$$\left[\begin{array}{l} \text{FOCUS} \\ \text{BACK.INF} \end{array} \left[\begin{array}{ll} \text{interrog} & \{x\} \\ \text{non-interrog} & \{['\text{Charlie}']\} \\ & \{['\text{eat}']\} \end{array} \right] \right]$$

5.3 Taking Stock

- Mycock does not deal with non-canonical interpretational effects produced by word order variation.
- Krifka does not deal (explicitly) with questions or with the effects of information structure in SOV languages like Urdu/Hindi.

Proposal: Combine Mycock’s overall approach to questions with Krifka’s view on information structure for an understanding of the pragmatic effects associated with word order variation of *wh*-constituents in Urdu/Hindi.

So the analysis for (19) would then be as shown below.

- (19) [nadya]_T (to) [ab^hi]_{CI} [tɔfi]_{CI} [bazar=se]_F xarid
 Nadya.F.Nom indeed just now toffee.F.Nom market.M=from buy
 rah-i t^h-i [mere=liye]_B
 stay-Perf.F.Sg be.Past-F.Sg I.Gen.Obl=for
 ‘Nadya was just buying toffee at the market for me.’

i-structure for (19)

| | |
|------------|---|
| TOPIC | $\left\{ \left[\begin{array}{ll} \text{PRED-FN} & \text{Nadya} \\ \text{TOPIC-TYPE} & \text{frame-setting} \end{array} \right] \right\}$ |
| FOCUS | $\left\{ \left[\begin{array}{ll} \text{PRED-FN} & \text{bazar} \\ \text{FOCUS-TYPE} & \text{default} \end{array} \right] \right\}$ |
| BACKGROUND | $\left\{ \left[\begin{array}{ll} \text{PRED-FN} & \text{bazar} \\ \text{BACK-TYPE} & \text{default} \end{array} \right] \right\}$ |

The values for Topic and Focus are sets because one can have multiple foci and topics. I also posit GIVEN for constituents that the morphosyntax or phonology mark as given.

5.4 Relationship between i-structure and semantics/pragmatics

Note: the approach taken here differs from Dalrymple and Nikolaeva (2011)

- They reject a Krifka type structured meaning approach.
- See information structure as partitioning sentence meaning into i-structure categories.

In contrast

- I follow Krifka in seeing information structure as additionally providing instructions for Common Ground Management.

- The computation of the semantics of an utterance is related to this indirectly — i.e., it is one of many factors taken into account in computing the overall semantics and pragmatics of an utterance.

The overall ramifications/predictions of the differing proposals with respect to information structure need to be investigated in detail.

6 Analysis: immediately postverbal *wh*-constituent

Proposal:

- The immediately postverbal position also is a focus position.
- Cf. the situation in Romance, where two focus positions are also assumed, Zubizarreta (1998), Samek-Lodovici (2005).

(20) sita=ne d^hyan=se dek^h-a **kis=ko** t^h-a?
 Sita.F=Erg carefully see-Perf.M.Sg who.Obl=Acc be.Past-M.Sg
 ‘Who had Sita really looked at carefully?’ (i.e., she had not looked at anybody carefully)

(21) varna nic^he log soc^h-ẽ-g-e pata nahĩ
 otherwise underneath people think-3.Pl-Fut-3.M know not
 ye log us kamre=mẽ kar **kya** rah-e hẽ
 this people that.Obl room.M.Sg.Obl=in do what stay-Perf.M.Pl be.Pres.3.Pl
 itn-i der=se
 this much-F.Sg time.F.Sg=Inst
 ‘Otherwise the downstairs people will think: What are these people doing in that room for such a long time?’
 (Lit. Otherwise the downstairs people will think: Don’t know WHAT these people are doing in that room for such a long time.) From *Socha Na Tha*

Some Initial Evidence:

- The *wh*-constituent cannot appear anywhere else in the verbal complex.
- The *wh*-constituent must be stressed.
- Manetta (2013) notes that clausal negation can occur either immediately preverbally or immediately postverbally (between the main verb and the auxiliaries).
- The Urdu/Hindi *nahĩ* ‘not’ is generally taken to have incorporated the focus particle *hi* ‘only’.

Analysis

- The different positions within a clause that a *wh*-element can appear in are correlated with the pragmatics of CG Management.
- The precise ways in which this occurs remains to be investigated.
- Current Take:
 - The immediately postverbal position is also a structural focus position.
 - However, it is not used to express default focus, but to add additional emphasis of some kind.
 - In terms of Alternative Semantics, this focus position is interpreted as follows: there is **no real alternative** that can be opened up.
 - I.e., the expectation is that there is no answer to the question.
 - This analysis could also potentially explain why no NP constituents can appear in this position — an NP provides at least one possible entity.
 - It could also explain why (22b) is bad.

- (22) a. sita kosi yazal ga-ti rah-ti t^h-i?
 Sita.F.Nom which.M.Sg song.M.Sg.Nom sing-Impf.F.Sg stay-Impf.F.Sg be.Past-F.Sg
 ‘Which song (ghazal) did Sita keep singing?’
- b. *sita ga-ti kosi yazal rah-ti t^h-i?
 Sita.F.Nom sing-Impf.F.Sg which.M.Sg song.M.Sg.Nom stay-Impf.F.Sg be.Past-F.Sg
 ‘Which song (ghazal) did Sita keep singing?’

i-structure for (20)

$$\left[\begin{array}{l} \text{TOPIC} \\ \text{FOCUS} \end{array} \left\{ \left[\begin{array}{ll} \text{PRED-FN} & \text{Sita} \\ \text{TOPIC-TYPE} & \text{default} \end{array} \right] \right\} \right]$$

$$\left[\begin{array}{l} \text{FOCUS} \\ \text{FOCUS} \end{array} \left\{ \left[\begin{array}{ll} \text{PRED-FN} & \text{kon} \\ \text{FOCUS-TYPE} & \text{emph} \end{array} \right] \right\} \right]$$

- At the moment, the FOCUS-TYPE is recorded as EMPH — this feature can trigger the above interpretation.
- But – it also leaves open other possibilities for the pragmatic CG Management.
- That is, the EMPH signals that this question should not be interpreted as a run-of-the-mill information seeking question, but that some extra pragmatics are to be associated with it.
- Cf. Sicoli et al. (2014), who find that pitch is used early on in an utterance to alert hearers to the fact that the question is to be interpreted non-canonically.

7 Polar Questions — Take 2

Recall that:

- Polar Questions generally have the same syntax as declaratives.

(23) a. anu=ne uma=ko kitab d-i
 Anu.F=Erg Uma.F=Dat book.F.Sg.Nom give-Perf.F.Sg
 ‘Anu give a/the book to Uma. declarative

b. anu=ne uma=ko kitab d-i?
 Anu.F=Erg Uma.F=Dat book.F.Sg.Nom give-Perf.F.Sg
 ‘Did Anu give a/the book to Uma? polar question

- Polar Questions can be overtly marked morphosyntactically via *kya* ‘what’.

(24) **kya** koi k^ho gā-ya
 what some lost go-Perf.M.Sg
 ‘Was someone lost?’ (from song *Koi Mil Gaya*)

- This has recently been dubbed as *Polar kya* by Bhatt and Dayal (2014).

(25) **kya** anu=ne uma=ko kitab d-i?
 what Anu.F=Erg Uma.F=Dat book.F.Sg.Nom give-Perf.F.Sg
 ‘Did Anu give a/the book to Uma? polar *kya*
 (Lit.: ‘What, Anu gave the book to Uma?’)

Traditional grammars only mention the clause initial position for polar *kya* (e.g., Glassman 1977, Platts 1884).

However, Bhatt and Dayal (2014) report a relatively free distribution of polar *kya*.

(26) (**kya**) anu=ne (**kya**) uma=ko (**kya**) kitab (**kya**) d-i?
 what Anu.F=Erg what Uma.F=Dat what book.F.Sg.Nom what give-Perf.F.Sg
 ‘Did Anu give a/the book to Uma? polar *kya*

- Bhatt and Dayal suggest that the different possible positions result from topicalization.
- They adduce evidence for the topicalization analysis from interactions with weak indefinites, idiomatic objects and gapping.
- Thus again, a clear connection is drawn between the position of the question element and information structure
- Bhatt and Dayal (2014) seek to understand polar *kya* as a speech act operator in the sense of Krifka (2014).

Alternative Proposal:

- Polar *kya* has two functions.
 1. In clause initial position it marks the clause as a polar question.
 2. In clause medial position it serves to partition the meaning of a clause.
 - Material to the left of the polar *kya* is not available for questioning — they are considered to be given.
 - Material to the right of the clause (in particular to the immediate right) is available for questioning, see (27) (data due to Rajesh Bhatt).

- (27) a. ram=ne sita=ko **kya** kitab d-i ya āguṭʰi?
 Ram.M=Erg Sita.F=Dat what book.F.Nom give-Perf.F.Sg or ring.F.Nom
 ‘Did Ram give a book or a ring to Sita?’
- b. ram=ne **kya** sita=ko kitab d-i ya amra=ko/*ravi=ne?
 Ram.M=Erg what Sita.F=Dat book.F.Nom give-Perf.F.Sg or Amra.F=Dat/Ravi.M=Erg
 ‘Did Ram give a book to Sita or Amra?/Did Ram or Ravi give a book to Sita?’

This appears to be similar to the partitioning of the meaning of a clause into a focus part and a background/given part within the **Structured Meaning** approach to focus Krifka (2008, 1992).

The background in (27) would be everything to the left of *kya* — this is part of what is presupposed/given and not available for focus (or for questioning in (27)).

One could therefore see polar *kya* as a type of **focus-sensitive operator** that determines which parts of a clause are backgrounded/presupposed and which are open for further discussion.

- Note: polar *kya* cannot appear immediately postverbally

- (28) *sita=ne dʰyan=se ram=ko dekʰ-a **kya** tʰ-a?
 Sita.F=Erg carefully see-Perf.M.Sg who.Obl=Acc what be.Past-M.Sg
 ‘Did Sita carefully look at Ram?’

- This restriction makes sense under the analysis that the immediately postverbal position is a focus position in which lexically contentful constituents make sense, but not focus-sensitive operators (which create their own focus domains).

$$\begin{array}{c} \mathbf{i\text{-}structure\ for\ (27a)} \\ \left[\begin{array}{l} \text{TOPIC} \\ \text{FOCUS} \\ \text{GIVEN} \end{array} \left\{ \begin{array}{l} \left[\begin{array}{l} \text{PRED-FN} \quad \text{Ram} \\ \text{TOPIC-TYPE} \quad \text{default} \end{array} \right] \\ \left[\begin{array}{l} \text{PRED-FN} \quad \text{kitab} \\ \text{FOCUS-TYPE} \quad \text{default} \end{array} \right] \\ \left[\begin{array}{l} \text{PRED-FN} \quad \tilde{\text{a}}\text{gu}\dot{\text{t}}^{\text{h}}\text{i} \\ \text{FOCUS-TYPE} \quad \text{default} \end{array} \right] \\ \left[\begin{array}{l} \text{PRED-FN} \quad \text{Ram} \\ \text{PRED-FN} \quad \text{Sita} \end{array} \right] \end{array} \right\} \right] \end{array}$$

- The partitioning of a clause into a given part vs. a focus part is accomplished via annotations in the c-structure introduced by the lexical entry of polar *kya*.
- This is analogous to the introduction of the meaning constructor [**interrog-scope**] by Mycock.
- Via f-precedence everything to the left of polar *kya* is GIVEN at i-structure.

8 Summary

There are various uses of Hindi/Urdu *kya* ‘what’:

- standard *wh*-item meaning ‘what’
- thematic (scope marking) *kya*
- polar *kya*

Understood word order variation in terms of information packaging needs.

- **Constituent Questions:**

- Word Order variation indicates strategies for *Common Ground Management* in the sense of Krifka (2008).
- Proposal of a second structural focus position immediately to the right of the verb — expresses the expectation that not alternative set of answers can be constructed (= there is actually no answer).

- **Polar *kya*:**

- in initial position it functions as a question marker/particle
- word order variation partitions a clause into a focus part and a background (given) part (as proposed by Krifka (1992) in a structured meaning approach to focus), whereby the given part is not available for questioning.

Acknowledgements

I would like to acknowledge Rajesh Bhatt, Veneeta Dayal for help with the data and for conducting many interesting discussions on the theoretical aspects of questions in Urdu/Hindi with me. I would further like to thank Ghulam Raza and Ashwini Vaidya for further help with the data and the audience of the TripleA 1 Workshop at the University of Tübingen, in particular, Manfred Krifka, for valuable feedback. My colleagues Maribel Romero and Maria Biezma are proving invaluable with respect to navigating me through the forest of the semantics of questions. Finally, I would like to thank my brother Yassin for putting on a CD at just the right moment.

References

- Artstein, Ron. 2002. A focus semantics for echo questions. In Á. Bende-Farkas and A. Riester (eds.), *Workshop on Information Structure in Context*, pages 98–107, Stuttgart: IMS: University of Stuttgart.
- Bhatt, Rajesh. 2003. Topics in the Syntax of the Modern Indo-Aryan Languages — *wh*-in-situ and *wh*-movement, class Handout, UMASS Amherst.
- Bhatt, Rajesh and Dayal, Veneeta. 2007. Rightward Scrambling as Rightward Movement. *Linguistics Inquiry* 38(2), 287–301.
- Bhatt, Rajesh and Dayal, Veneeta. 2014. Polar *kyaa*, talk presented at the *Workshop on Non-Canonical Questions and Interface Issues*, Konstanz, February 2014.
- Butt, Miriam and King, Tracy H. 1996. Structural Topic and Focus without Movement. In M. Butt and T. H. King (eds.), *Proceedings of the First LFG Conference*, CSLI Publications.
- Butt, Miriam and King, Tracy Holloway. 1997. Null Elements in Discourse Structure, written to be part of a volume that never materialized.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles Li (ed.), *Subject and Topic*, pages 27–55, New York: Academic Press.
- Choi, Hye-Won. 1996. *Scrambling: Optimality-theoretic Interaction between Syntax and Discourse*. Ph.D. thesis, Stanford University.
- Dalrymple, Mary and Nikolaeva, Irina. 2011. *Objects and Information Structure*. Cambridge: Cambridge University Press.
- Dayal, Veneeta. 1994. Scope Marking as Indirect *wh*-dependency. *Natural Language Semantics* 2(2), 137–170.
- Dayal, Veneeta. 1996. *Locality in WH Quantification*. Dordrecht: Kluwer Academic Publishers.

- Dayal, Veneeta. 2003. Bare nominals: Non-specific and contrastive readings under scrambling. In Simin Karimi (ed.), *Word Order and Scrambling*, pages 67–90, Oxford: Blackwell.
- Dayal, Veneeta. 2014. Questions and the vP Phase in Hindi-Urdu, ms. Rutgers University.
- É. Kiss, Katalin. 1995. *Discourse Configurational Languages*. Oxford: Oxford University Press.
- Gambhir, Vijay. 1981. *Syntactic Restrictions and Discourse Functions of Word Order in Standard Hindi*. Ph. D. thesis, University of Pennsylvania, Philadelphia.
- Glassman, Eugene H. 1977. *Spoken Urdu*. Lahore: Nirali Kitabon.
- Gurtu, Madhu. 1985. *Anaphoric Relations in Hindi and English*. Ph. D. thesis, CIEFL, Hyderabad.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: The MIT Press.
- Kidwai, Ayesha. 2000. *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. Oxford: Oxford University Press.
- Krifka, Manfred. 1992. A Compositional Semantics for Multiple Focus Constructions. In Joachim Jacobs (ed.), *Informationsstruktur und Grammatik*, pages 17–53, Opladen: Westdeutscher Verlag.
- Krifka, Manfred. 2008. Basic Notions of Information Structure. *Acta Linguistica Hungarica* 55(3–4), 243–276.
- Krifka, Manfred. 2014. Embedding Speech Acts. In Tom Roeper and Peggy Speas (eds.), *Recursion*.
- Lahiri, Utpal. 2002. On the proper treatment of expletive wh in Hindi. *Lingua* 112, 501–540.
- Mahajan, Anoop. 1990. *The A/A-Bar Distinction and Movement Theory*. Ph. D. thesis, MIT.
- Mahajan, Anoop. 1997. Rightward Scrambling. In Dorothee Beerman, David LeBlanc and Henk van Riemsdijk (eds.), *Rightward Movement*, pages 185–213, Amsterdam: John Benjamins.
- Manetta, Emily. 2010. Wh-expletives in Hindi-Urdu: The vP phase. *Linguistic Inquiry* 41(1), 1–34.
- Manetta, Emily. 2012. Reconsidering Rightward Scrambling: Postverbal Constituents in Hindi-Urdu. *Linguistic Inquiry* 43(1), 43–74.
- Manetta, Emily. 2013. Copy theory in wh-in-situ languages: Sluicing in Hindi-Urdu. *Journal of South Asian Linguistics* 6, 3–22.
- Mycock, Louise. 2006. *The Typology of Constituent Questions: A Lexical-Functional Grammar Analysis of wh-questions*. Ph. D. thesis, University of Manchester.

- Mycock, Louise. 2013. Discourse Functions of Question Words. In M. Butt and T.H. King (eds.), *Proceedings of the LFG13 Conference*, Stanford: CSLI Publications.
- Platts, John T. 1884. *A dictionary of Urdu, classical Hindi, and English*. W. H. Allen and Co.
- Prasad, Rashmi. 2000. A Corpus Study of Zero Pronouns in Hindi: An Account Based on Centering Transition Preferences. In *Proceedings of DAARC2000*, pages 66–71, Lancaster University.
- Prasad, Rashmi. 2003. *Constraints on the Generation of Referring Expressions, with Special Reference to Hindi*. Ph. D. thesis, University of Pennsylvania, Philadelphia.
- Samek-Lodovici, Vieri. 2005. Prosody-syntax interaction in the expression of focus. *Natural Language and Linguistic Theory* 23, 687–755.
- Sicoli, Mark A., Stivers, Tanya, Enfield, Nick J. and Levinson, Stephen C. 2014. Marked initial pitch in questions signals marked communicative function. *Language and Speech* DOI: 10.1177/0023830914529247.
- Srivastav, Veneeta. 1991. Subjacency Effects at LF: The Case of Hindi wh. *Linguistic Inquiry* 22(4), 762–769.
- Stepanov, Arthur and Stateva, Penka. 2006. Successive cyclicity as residual wh-scope marking. *Lingua* 116, 2107–2153.
- Truckenbrodt, Hubert. 2012. An analysis of prosodic F-effects in interrogatives: Prosody, syntax and semantics. *Lingua* 124, 131–175.
- Vallduví, Enric. 1992. *The Informational Component*. New York: Garland Press.
- Zubizarreta, Maria Luisa. 1998. *Prosody, Focus and Word Order*. Cambridge: The MIT Press.

Appendix: Questions in the Wilds of Bollywood

Lyrics for *Koi mil gaya*

From Bollywood Movie *Kuch Kuch Hota Hai* (1998)

- (29) a. koi mil gā-ya
 some.Nom meet go-Perf.M.Sg
 ‘(I’ve) met someone.’ (Lit. (to me) someone is found.)
- b. koi mil gā-ya
- c. koi mil gā-ya

Most natural under a question interpretation, but is a declarative.

Refrain 1:

- (30) a. muj^h=ko **kya** hu-a hai
 I.Obl=Dat what become-Perf.M.Sg be.Pres.3.Sg
 ‘What has happened to me?’
- b. **kyū** mē k^ho gā-ya hū
 why I.Nom lost go-Perf.M.Sg be.Pres.1.Sg
 ‘Why have I gotten lost?’
- c. pagal t^h-a **mē pahle**
 crazy be.Past-M.Sg I.Nom first
 ‘Was I crazy before?’
- d. **ya** ab ho gā-ya hū
 or now be go-Perf.M.Sg be.Pres.1.Sg
 ‘or have I become it now?’
- (31) a. bahaki hē **nigah-ē**
 wandering be.Pres.3.Sg gaze.F-Pl.Nom
 ‘(Your) gaze is wandering’
- b. or bik^hr-e hē **bal**
 and scattered-M.Pl be.Pres.3.Pl hair.M
 ‘and the hair is ruffled’
- c. tūm=ne ban-a-ya hē
 you.Sg.Fam=Erg be.made-Caus-Perf.M.Sg be.Pres.3.Sg
kya apn-a ye hal
 what self-M.Sg this state
 ‘What state have you gotten yourself into?’

crazy clause seems like a declarative, but is in fact part of an alternative question

kya postverbal with another NP

Refrain 2:

- (32) a. koi mīl gā-ya
 b. koi mīl gā-ya
 c. mer-a dīl gā-ya
 my-M.Sg heart.M.Sg.Nom go-Perf.M.Sg
 ‘My heart has gone.’
 d. mer-a dīl gā-ya
 e. **kya** bata-ō yar-ō
 what tell-Subj.1.Sg friend-M.Pl
 ‘what should I tell you, friends?’
 f. kya bata-ō yar-ō
 g. mē **to** hīl gā-ya
 I.Nom Top move go-Perf.M.Sg
 ‘As for me, I’ve been moved (no longer firmly in place).’
 h. mē **to** hīl gā-ya
 i. koi mīl gā-ya, mīl **hi** gā-ya
 some.Nom meet go-Perf.M.Sg meet Emph go-Perf.M.Sg
 ‘(I’ve) found someone, I really have.’
 j. mīl gā-ya, hei mīl **hi** gā-ya
 meet go-Perf.M.Sg hey meet Emph go-Perf.M.Sg
 ‘Found, hey, really found’
 k. mīl gā-ya
 meet go-Perf.M.Sg
 ‘Found.’

(Refrain 1)**(Refrain 2)**

- (33) a. jan ne **kya** ho gā-ya he **muj^h-e**
 know not what be go-Perf.M.Sg be.Pres.3.Sg I.Obl-Dat
 ‘What has happened to me, (I) don’t know.’
 b. divana log kah-ne lag-e
 crazy people.Nom say-Inf.M.Pl begin-M.Pl
 ‘People are beginning to call (me) crazy.’
 c. ja ne kya ho gā-ya he muj^he
 d. divana log kah-ne lag-e

- (34) a. ye divanegi he **kya**
 this.Nom craziness.F.Sg.Nom be.Pres.3.Sg what
 ‘What is this craziness?’
- b. h̄om-ē b^hi to ho pata
 1.Pl-Dat also Top be know
 ‘I should know as well.’
- c. t̄um=ko **kya** ho ḡa-ya
 you.Sg.Fam=Dat what be go-Perf.M.Sg
 ‘What has happened to you?’
- (35) a. arre kal tak m̄oj^h=ko sab hoj t^h-a
 hey yesterday til I.Obl=Dat all conscious be.Past-M.Sg
 ‘Hey, up to yesterday I was conscious of everything.’
- b. d̄il=m̄ē khufi-yō=ka jof t^h-a
 heart.M.Sg=in happiness.F=Obl.Pl=Gen.M.Sg enthusiasm.M.Nom be.Past-M.Sg
 ‘In (my) heart there was a capacity of joy.’
- (36) a. p^hir ye bec̄ēni he **kyū**
 then this.Nom restlessness.F.Sg.Nom be.Pres.3.Sg why
 ‘Then why is there this restlessness?’
- b. p^hir ye betabi he **kyū** Polar kya
 then this.Nom impatience.F.Sg.Nom be.Pres.3.Sg why
 ‘Then why is there this impatience?’
- c. **kya** koi k^ho ḡa-ya
 what some lost go-Perf.M.Sg
 ‘Is someone lost?’

(Refrain 2)

- (37) a. badl ban kar **kon** a ḡa-ya
 cloud.M.Nom be.made having who.Nom come go-Perf.M.Sg
 ‘Who has arrived in the form of a cloud?’
 (Lit. Who has arrived, having become a cloud?)
- b. **kon** he jo d̄il=pe yū c^ha ḡa-ya
 who.Nom be.Pres.3.Sg who heart.M=on like this overcast go-Perf.M.Sg
 ‘Who is it, who has overwhelmed (your) heart like this?’
- (38) a. cah-ū ki bata-ū **m̄ē**
 want-Subj.1.Sg that tell-Subj.1.Sg I.Nom
 ‘I would like that I tell (you)’

- b. p^hir b^hi kah na pa-ũ **mẽ**
 then also say not able-Subj.1.Sg I.Nom
 ‘even then I could not say’
- c. nam us=ka hε **kya**
 name.M.Sg.Nom Pron.3.Sg.Obl=Gen.M.Sg be.Pres.3.Sg what
 ‘what her name is.’
- (39) a. o nam na l-o par kuc **to** kah-o
 oh name.M.Sg.Nom not take-Imp.2.Sg.Fam but something.Nom Top say-Imp.2.Sg.Fam
 ‘So don’t take her name, but do say something.’
- b. halka=sa koi ifara **to** d-o
 weak=kind.M.Sg some sign.M.Sg.Nom Top give-Imp.2.Sg
 ‘Give us some small sign.’
- (40) a. mer-i āk^h-õ=mẽ hε **vo**
 my-Fem eye.Fem-Obl.Pl=in be.Pres.3.Sg Dem.3.Sg.Nom
 ‘She is in my eyes.’
- b. mer-i sās-õ=mẽ hε **vo**
 my-Fem breath.Fem-Obl.Pl=in be.Pres.3.Sg Dem.3.Sg.Nom
 ‘She is in my breath.’
- c. or kah-ũ **tum=se** **kya**
 more say-Subj.1.Sg you.Sg.Fam=Inst what
 ‘what else should I say to you?’

(Refrain 2)

(Refrain 1)

(Refrain 2)