

Variation in Bangla Complementizer Order at the Syntax-Prosody Interface

Brian Hsu (University of Southern California)

1 The Bangla *je* puzzle

- The complementizer *je* optionally introduces finite embedded clauses.
- If the embedded CP is **postverbal**, *je* is obligatorily initial in its clause.

(1) Jon jane [(**je**) ma kal rate oSudh kheyechē]
 John knows [that mother last.night medicine ate]
 'John knows that mother took medicine last night.'

- If the embedded CP is **preverbal**, *je* can only be pronounced in a non-initial position (Dasgupta 1980; Bhattacharya 2001).

(2) Jon [ma (**je**) kal rate oSudh kheyechē] jane
 John [mother that last.night medicine ate] knows
 'John knows that mother took medicine last night'

(3) [oSudh (**je**) ma kal rate kheyechē] Jon jane
 [medicine that mother last.night ate] John knows
 'John knows that mother took medicine last night'

- In preverbal CPs, *je* is not restricted to second position, and can appear after any number of phrases, under certain interpretational restrictions (§2).

(4) John [kal rate (**je**) oSudh (**je**) ma (**je**) kheyechē] jane
 John [last.night that medicine that mother that ate] knows
 'John knows that mother took medicine last night.'

- The positions of embedded clauses depend on their interpretations (Simpson & Bhattacharya 2003):

- Postverbal CPs: informationally neutral. [initial *je*]
- Medial CPs: emphatic/corrective focus. [non-initial *je*]
- Sentence-initial CPs: topicalized information. [non-initial *je*]

2 Topic, Focus, and non-initial *je*

- Phrases preceding *je* are typically topicalized or contrastively focused.
- Diagnostics for topic-hood:

- Definite DPs (indicated by [NP [Num-Cl] order) are good topics.
- Indefinite and quantified DPs are ineligible as topics, but can be contrastively focused.

- **Postverbal CPs:** Topics can be raised to a position to the right of *je*.

(5) Jon bol-lo [je dadubhai *chatro du-to-ke* dekheche]
 John say-pst that grandfather student 2-cl-acc saw
 'John said that grandfather saw the two students'

(6) Jon bol-lo [je *chatro du-to-ke* dadubhai dekheche]
 John say-pst that student 2-cl-acc grandfather saw
 'John said that grandfather saw the two students'

- Indefinite or quantified objects can only be fronted if they are contrastively focused.

(7) Jon bol-lo [je #*kau-ke/KAU-KE* dadubhai dekheche ni]
 John say-pst that anyone-acc grandfather see-perf neg]
 'John said that grandfather didn't see ANYONE.'

- **Preverbal CPs:** The position to the left of *je* is typically reserved for topics. Quantified phrases occur in this position only under contrastive focus **Error!**
Reference source not found.

(8) Jon [*chatro du-to-ke je* dadubhai dekheche] bol-lo
 John student 2-cl-acc that grandfather see-perf say-pst
 'John said that grandfather saw the two students'

(9) Jon [#*kau-ke/KAU-KE je* dadubhai dekheche ni] bol-lo
 John anyone-acc that grandfather see-perf neg say-pst

- Both (multiple) topics and focus can co-occur before *je*, with topics crucially ordered before focus.

(10) Jon [chatro du-to -ke DADUBHAI je dekh-eche] bol-lo
 John student 2-cl -acc GRANDFATHER that see-perf say-pst
 'John said that GRANDFATHER saw the two students'

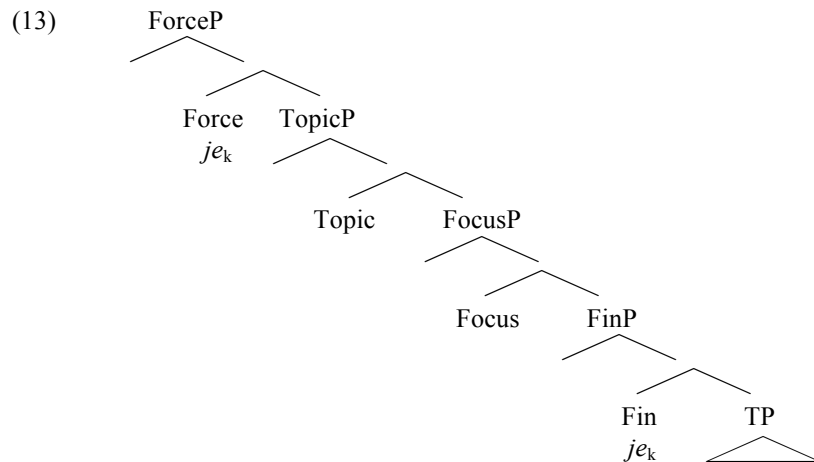
- Subjects preceding *je* are not restricted in interpretation, suggesting that fronting before *je* is not purely due to discourse requirements (Bayer & Dasgupta, *to appear*; Bhattacharya, 2001).

(11) Jon [du-to chatro je es-che] bollo
 John 2-cl student that come-perf said
 'John said that two students came'

(12) Jon [keu je ase-ni] bollo
 John [anyone that come.perf-neg] said
 'John said that no one came'

3 The Bangla left periphery

- The patterning of *je* is derived by a **lower copy spell-out analysis** (Franks 1998; Bošković 2001) in an **extended CP structure** (Rizzi, 1997, among others).



- All embedded clauses share the same base structure (13).
 - *je* is merged in Fin and copied in Force, resulting in its specification as [+finite] and [+declarative]
 - Linear order is determined by which copy is pronounced at PF; **Initial *je* is pronounced in Force, non-initial *je* is pronounced in Fin.**
- Head movement depends on agreement between Fin and Force, and does not need to proceed through all structurally intervening heads (cf. Roberts, 2000).
 - This predicts that *je*'s lower copy can be preceded by multiple topics/focus (ex. (4), (10)).
 - Such patterns are not straightforwardly derived by necessarily local operations, e.g. Morphological Merger (Embick and Noyer 2001) or Prosodic Inversion (Halpern 1992).

4 Prosodic factors in copy spell-out

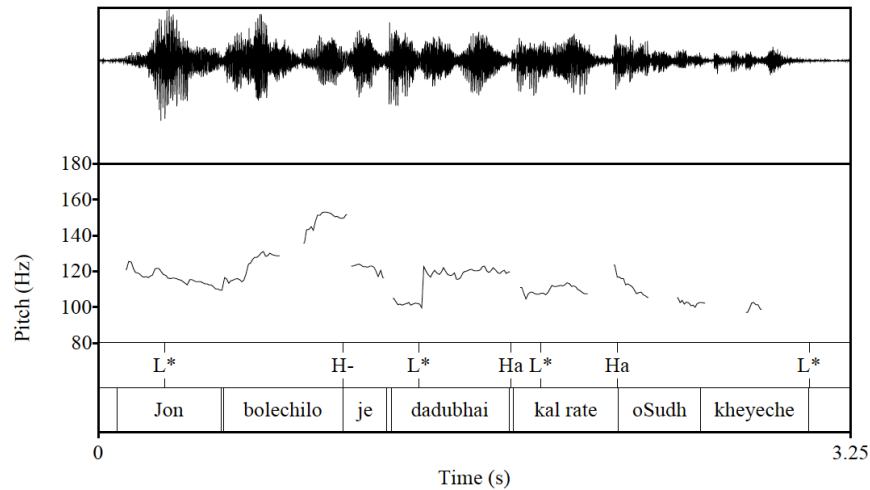
- **Main proposal:** Lower copy spell-out takes place to avoid *je* in initial position of an intonational phrase. Only pre-verbal CPs are separate intonational phrases.

4.1 Background: Bangla intonational prosody

- I adopt Khan's (2008) model of Bangla intonational phonology.
- The **Accentual Phrase (AP)** roughly corresponds with every syntactic XP.
 - APs bear a low pitch accent on their first stressed syllable, and a high right boundary tone (L* ... Ha).
- The **Intermediate Phrase (ip)** and **Intonational Phrase** contain several APs and correspond to a larger syntactic units.
 - Both IPs and ips are identified by a variety of **boundary tones** realized on the rightmost AP in their domain.
 - ips are typically embedded clauses or topics, while IPs demarcate parentheticals or full clauses.
 - Full IPs are typically distinguished from ips by **pauses**, and the possibility of a unique 'continuation rise' boundary tone (HLH%).

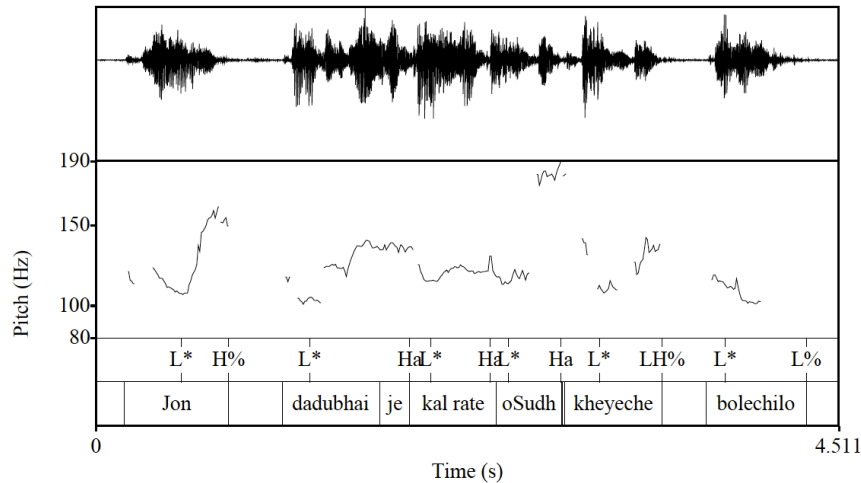
4.2 Phrasing of embedded clauses

- Postverbal embedded clauses are **intermediate phrases**, typically marked by a H- or L- boundary tone, not separated by pauses.



(14) Jon bolechilo [je dadubhai kal rate oSudh kheyechē]
 John say.pst.perf.3 that grandfather last night medicine ate]
 'John said that grandfather took medicine last night'

- Preverbal CPs form **intonational phrases**, typically separated by pauses preceded by a LH%, H%, or HLH% boundary tone.



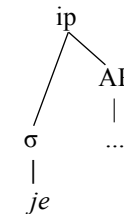
(15) Jon [dadubhai je kal rate oSudh kheyechē] bolechilo
 John grandfather that last night medicine ate] say.pst.perf.3
 'John said that grandfather took medicine last night.'

- **je is always non-initial if its embedded clause is an intonational phrase.**

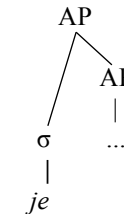
4.3 Prosodic motivation for lower copy spell-out

- The fact that *je* never undergoes vowel lengthening indicates that it is a stray syllable that does not form a PWd, since PWds in Bangla are minimally bimoraic (Fitzpatrick-Cole 1991).
- **Je is not an enclitic** in the initial position of an embedded clause (contra Bayer & Dasgupta, *to appear*; Chacón, 2014);
 - *je* does not bear the main verb's boundary tone (e.g. in (14)), as would be expected under AP enclisis.
 - A representation of postverbal *je* as an AP-enclitic structure is ruled out by the presence of an ip boundary.
- The AP-level phrasing of postverbal, clause-initial *je* is indeterminate between a free clitic, affixal proclitic, or internal proclitic structure, since there is no reliable diagnostic for the left edge of an AP.

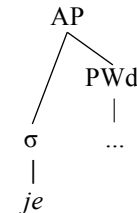
(16) *Free clitic*



(17) *Affixal clitic*



(18) *Internal clitic*



- Crucially, all possible representations of clause-initial *je* violate STRONGSTART(σ) (Selkirk 2011; Elfner 2012).
- **Proposal:** STRONGSTART violations are additionally sensitive to the presence of larger prosodic boundaries (cf. Flack, 2009). In Bangla, they are banned at an intonational phrase boundary.
- Prosodic constraints interact with syntactic constraints on linearization to determine the pronunciation of copies (cf. Elfner, 2012; López, 2009).
 - In an OT grammar (Prince and Smolensky 1993) lower copy spell-out results from the ranking STRONGSTART(σ /IP) >> HIGHESTCOPY.¹

- (19) STRONGSTART(σ /IP)
 Assign a violation mark for every prosodic constituent whose leftmost daughter constituent is of type σ and is lower in the Prosodic Hierarchy than its sister constituent immediately to the right, where σ is at the left edge of an intonational phrase.
- (20) HIGHESTCOPY
 Pronounce the only the highest copy of a movement chain.
- (21) *Lower copy spell-out prevents violation of STRONGSTART(σ /IP)*

[_{ForceP} <i>je</i> [XP [_{FinP} <i>je</i> ...]]] V	STRONGSTART (σ /IP)	HIGHEST COPY
(IP (_{ip} <i>je</i> (_{AP} XP) ...) <i>free clitic</i>	*!	
(IP (_{ip} (_{AP} <i>je</i> (_{AP} XP)) ...) <i>affixal proclitic</i>	*!	
(IP (_{ip} (_{AP} <i>je</i> XP) ...) <i>internal proclitic</i>	*!	
[☞] (IP (_{ip} (_{AP} (_{AP} XP) <i>je</i>) ...) <i>lower copy, affixal enclitic</i>		*

6 Conclusion

- Variation in Bangla complementizer placement is derived during copy spell-out, driven by prosodic factors.

¹ To prevent lower copy spell-out from applying in postverbal embedded clauses, STRONGSTART(σ /ip) is ranked below HIGHESTCOPY.

- The data expand the known typologies of second position effects and bans on weak elements at prominent prosodic edges.
- Close analysis of syntax-prosody correspondence sheds light on puzzling variation in syntax.

Selected References

- Bayer, Josef, and Probal Dasgupta. *To appear*. Emphatic Topicalization and the structure of the left periphery : Evidence from German and Bangla. *Syntax*.
- Bhattacharya, Tanmoy. 2001. Peripheral and clause-internal complementizers in Bangla : A case for remnant movement. *Proceedings of WECOL 2000* 12: 1–14.
- Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Oxford: Elsevier.
- Chacón, Dustin Alfonso. 2014. *The antisymmetric analysis of Bangla complementation*. Ms., University of Maryland, College Park.
- Dasgupta, Probal. 1980. Questions and relative and complement clauses in a Bangla grammar. New York University.
- Elfner, Emily. 2012. *Syntax-Prosody Interactions in Irish*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Embick, David, and Rolf Noyer. 2001. Movement Operations after Syntax. *Linguistic Inquiry* 32: 555–595.
- Fitzpatrick-Cole, Jennifer. 1991. The minimal word in Bengali. *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*: 157–170.
- Flack, Kathryn. 2009. Constraints on onsets and codas of words and phrases. *Phonology* 26: 269–302.
- Franks, Steven. 1998. Clitics in Slavic. In *Comparative Slavic Morphosyntax Workshop*, 1–96. Indiana University, Bloomington.
- Halpern, Aaron. 1992. Topics in the placement and morphology of clitics. Stanford University.
- Khan, Sameer ud Dowla. 2008. *Intonational Phonology and Focus Prosody of Bengali*. Ph.D. dissertation, University of California, Los Angeles.
- López, Luis. 2009. Ranking the Linear Correspondence Axiom. *Linguistic Inquiry* 40: 239–276.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, ed. Liliane Haegeman, 281–337. Dordrecht: Kluwer.
- Roberts, Ian G. 2000. Head Movement. In *Handbook of Syntactic Theory*, ed. Mark Baltin and Chris Collins, 113–147. Oxford: Blackwell.
- Selkirk, Elisabeth. 2011. The Syntax-Phonology Interface. In *The Handbook of Phonological Theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu, 2nd ed., 1–66. Oxford: Blackwell.