

C-selection is necessary for clausal embedding: evidence from questions with declarative syntax

Daniar Kasenov (NYU)

SYNC2024

1 Introduction

- One of the basic-est questions: why do certain phrases have the distribution that they do
- For example: why does the adjective *proud* take a prepositional phrase instead of a nominal phrase?
- These sorts of questions often receive a syntactic answer (arbitrary co-occurrence properties)
- However: a rather large tradition of analysing distribution semantically
 - NPIs: appear in downward-entailing contexts (Ladusaw 1979 and subsequent work)
 - Connected exceptives: possible with universal quantifiers, impossible with existential quantifiers (von Stechow 1993)
 - Strong determiner phrases: impossible in existentials (Definiteness Effect: Barwise & Cooper 1981; though see Preminger 2018 for a dissenting view)
 - And others...
- While cases like NPIs appear to be undoubtedly semantic, while lexical-selection cases appear to be undoubtedly syntactic, many phenomena lie in between.
- A hot&classic topic: clausal embedding (Grimshaw 1979; Pesetsky 1982; Odijk 1997; Moulton 2009; Uegaki 2015; Elliott 2020; Uegaki 2022 among others)
- For example, recent work argues that distribution of (non-)finite clauses cannot be reduced to semantic composition (see review in Wurmbrand 2024)
- Basic observation: clause-embedding predicates place restrictions on whether their embedded clauses can be declarative/interrogative.
 - (1) a. Nick thinks {that/*whether} Mencius drinks bourbon.
b. Nick knows {that/whether} Mencius drinks bourbon.
c. Nick asks {*that/whether} Mencius drinks bourbon.
- How would you account for that?
 - Strategy I is c-selection: each verb is lexically specified with a feature like [\bullet C \bullet], [\bullet C[+Q] \bullet], [\bullet C[−Q] \bullet].
 - Strategy II: s-selection. Syntax generates everything but certain combinations fail to be well-formed semantically

- ★ A v. simple solution would be a type-mismatch approach (if declaratives are of type $\langle s, t \rangle$ while interrogatives are of type $\langle \langle s, t \rangle, t \rangle$)
- ★ However: type-mismatch analyses cannot accomodate predicates like *know* (so-called responsive predicates; Lahiri 2002) which take both types of clauses without resorting to homophony
- ★ Another note: I take every mixed approach to fall into strategy I (since I aim to claim that c-selection is necessary to account for the data landscape, not that it is sufficient)
- A prominent type of an argument against c-selection approaches is the ‘lost generalization’ sort of argument
 - (2) Argument against c-selection in clausal embedding

There are generalizations about question-embedding predicates which can only be accounted for in an s-selection account.
- These generalizations include (Uegaki 2022: section 1.6.2.1.):
 - Veridical predicates (i.e., ones that imply that their complement is true) describing the subject’s doxastic or epistemic state (e.g., know, remember) always select for both declarative and interrogative complements.
 - Non-veridical predicates expressing preferences (e.g., want, hope) never select for interrogative complements (e.g., *Mary hopes who left).
 - Neg-raising predicates (e.g., think, believe) never select for interrogative complements (e.g., *Mary thinks who left).
 - Among predicates that take interrogative complements, those implying belief in the existence of a true answer to the issue expressed by their complement but ignorance as to which of the answers is true (e.g., wonder, investigate) never take declarative complements.
- Note: this argument only goes through if there is no other way to account for such generalizations
- Another note: I am not arguing with a strawman
 - (3) “Furthermore, the semantic explanations of selectional restrictions presented here will eliminate the need for syntactic subcategorization mechanisms.” (Uegaki 2022: p.205)
- This talk: an argument against pure s-selection approaches
 - (4) Wrong prediction of s-selection approaches

If a clause can act as a matrix question, it can act as an embedded question.
- A core empirical domain: questions with declarative syntax (see also Bobaljik & Wurmbrand 2015)
 - (5) a. John did WHAT? (wh-question)
 - b. Russian
 mama pošla gul’jat’?
 mom went walk
 ‘Did mum go to take a walk?’ (polar question)

- Generalization: questions with declarative syntax cannot be embedded

(6) a. *I asked John did WHAT?

b. Russian

**ja sprosil mama posla guljat'?*

i asked mom went walk

'I asked whether mum went to take a walk?'

- Bobaljik & Wurmbrand 2015: this necessitates a c-selectional analysis
- In this talk, I:
 - re-iterate Bobaljik and Wurmbrand's points
 - provide additional polar question data from Russian
 - discuss possible s-selection-compliant answers to the problem

2 Logic behind contemporary s-selectional approaches

- Prominent approach to semantics of questions (Hamblin 1976): questions as sets of propositions

(7) $\llbracket \text{Who came?} \rrbracket = \{\lambda w. \text{Nick came in } w, \lambda w. \text{Mencius came in } w, \lambda w. \text{Curtis came in } w, \dots\}$

- Initial puzzle: responsive predicates

(8) I learned {who came / that Sue came}

- How can *learn* combine both with a proposition and a set of propositions?
 - Q-to-P reduction approaches posit an *ANS* operator (from a question to its answer)
 - P-to-Q reduction posit a *TRIV* operator (from p to $\{p\}$ singleton set)
 - Let's assume P-to-Q reduction for clarity (we take Uegaki's system as a foil)
- In s-selectional approaches, the question is: why are some predicates 'picky'?
- Uegaki 2022 pursues an L-triviality approach (see Del Pinal 2022 for a recent take on logicity of language):
 - predicates incompatible with embedded questions (anti-rogative) are trivial when embedding a non-singleton set of propositions
 - predicates incompatible with embedded declaratives (rogative) are trivial when embedding a singleton set of propositions
- Therefore:
 - whatever denotes a non-singleton set of propositions should be able to occur with rogative predicates
 - whatever denotes a singleton set of propositions should be able to occur with anti-rogative predicates

- (9) Wrong prediction of s-selection approaches

If a clause can act as a matrix question, it can act as an embedded question.

- (10) Wrong prediction of s-selection approaches (v.2)

If a clause behaves semantically as an assertion, it can act as an embedded non-interrogative clause.

3 Questions with declarative syntax: the data

- Bobaljik & Wurmbrand 2015 put forth the following generalization

(11) If a language has wh-movement (to Spec,CP), then wh-movement is obligatory in indirect questions.

- Sidenote: some have argued that optional wh-movement is focus-movement (see Cheng 1997 on Indonesian and Faure & Palasis 2021 on Colloquial French). I abstract away from that. See also Tan & Shen 2024 for a discussion of a possible counterexample from Singapore English.

- Here's some data to illustrate

(12) German

- Und wie diese Teilhaber erreichen wir?*
and how these partners reach we
'And how can we reach these partners?'
- **Stark hat gefragt wie diese Teilhaber erreichen wir.*
Stark has asked how these partners reach we
Int.: 'Stark asked how we can reach these partners.'
- Und diese Teilhaber erreichen wir wie?*
and these partners reach we how
'And we can reach these partners how?'
- **Stark hat gefragt diese Teilhaber erreichen wir wie?*
Stark has asked these partners reach we how
Int.: 'Stark asked we can reach these partners how?'

(13) Brazilian Portuguese

- Quem você viu?*
who you saw
'Who did you see?'
- O Pedro pergunta quem você viu?*
The Pedro asked who you saw
'Pedro asked who you saw.'

- c. *Você viu quem?*
you saw who
'Who did you see?'
- d. **O Pedro pergunto você viu quem?*
The Pedro asked you saw who
Int.: 'Pedro asked who you saw.'

(14) Russian

- a. *s kem buxaet Nik?*
with whom drink Nick
'Whom does Nick drink with?'
- b. *Nik buxaet s kem?*
Nick drink with whom
'Whom does Nick drink with?'
- c. *Mentsius sprosil s kem buxaet Nik?*
Mencius asked with whom drink Nick
'Mencius asked whom does Nick drink with?'
- d. **Mentsius sprosil Nik buxaet s kem?*
Mencius asked Nick drink with whom
Int.: 'Mencius asked whom does Nick drink with?'

- Similar phenomenon is found with polar questions with declarative syntax (sometimes also referred to as 'rising declaratives'). See also Rudin 2019 for arguments that apparent counterexamples involve quotation.

(15) Polar questions with interrogative syntax

- a. Did Olivia win an Oscar?
- b. Russian (Shushurin 2024; Esipova & Korotkova 2024)
Govorite =li vy po-russki?
speak =Q you russian
'Do you speak Russian?'

(16) Polar questions with declarative syntax

- a. Olivia won an Oscar?
- b. Russian (Shushurin 2024; Esipova & Korotkova 2024)
vy govornite po-russki?
you speak russian
'Do you speak Russian?'

(17) Embedding data

- a. Yorgos wondered whether Olivia won an Oscar.
- b. *Yorgos wondered Olivia won an Oscar? (bad as proper embedding).
- c. Russian (Shushurin 2024; Esipova & Korotkova 2024)
Masha sprashivaet / somnevaetsia govorite =li vy po-russki.
Masha asks doubts speak Q you Russian
‘Masha asks/doubts whether you speak Russian.’
- d. Russian (Shushurin 2024; Esipova & Korotkova 2024)
**Masha sprashivaet / somnevaetsia vy govorite po-russki.*
Masha asks doubts you speak Russian
‘Masha asks/doubts whether you speak Russian.’

- If distribution of embedded questions is only determined by their semantics, why can’t all clauses that act as questions in root contexts act as questions in embedded contexts?
- Possible answers
 - Syntactic size (cf. Bhatt & Dayal 2020): whatever allows a syntactically-declarative clause to have interrogative semantics-pragmatics is ‘too high’ to be embedded (c-selection!)
 - Esipova & Korotkova 2024: whatever allows a syntactically-declarative clause to have interrogative semantics-pragmatics require a subsequent conversational move (“conventionally encodes pressure to react”).
 - However, I see nothing pragmatically illicit about claiming that X wonders about a question Q while putting pressure on the interlocutor to resolve Q (to be completely fair, Esipova & Korotkova 2024 do not present a thorough account, so this line of thought is yet to fully emerge)
- Note: this argumentation assumes that questions with declarative syntax have interrogative semantics (denote sets of propositions)
- This assumption can be dropped

4 What if questions with declarative syntax denote propositions, not sets of propositions?

- One can argue that the argumentation only goes through if it is indeed the case that questions with declarative syntax acquire question-like pragmatic behavior in the same way as proper questions do.
- Biezma 2020 argues that this is not the case:
 - Her system is built upon the idea that questions with declarative syntax are assertions that do not resolve QuD in any way but that *enforce* the QuD

(18) What did Nick buy

Assertion: $\{\text{bought}(\mathbf{n}, x) \mid x \in D_e\}$ (set of proposition)

Introduces QuD ‘what did Nick buy’

(19) Nick bought [WHAT]_F

Assertion: $\exists p \in \{\text{bought}(\mathbf{n}, x) \mid x \in D_e\} p(w_0)$ (a proposition)

Requires QuD ‘what did Nick buy’

- This accounts for a number of the pragmatic properties of such questions (their infelicity out-of-the-blue, etc.). See Biezma’s paper for details
- Does this account for the embeddability data, though?
- Note however, that the sword possesses a second edge

(20) Wrong prediction of s-selection approaches (v.2)

If a clause behaves semantically as an assertion, it can act as an embedded non-interrogative clause.

- This is not borne out! Questions with declarative syntax are a root phenomenon: they are not embedded by anything

(21) a. *Nick said (that) Mencius drank what?

b. Russian (with a rising intonation at the end)

**Masha dumaet vy govorite po-russki.*

Masha thinks doubts you

- The unavailable interpretation would be something like “John thinks that there is something that Mencius drank” / “Masha thinks that you speak or do not speak Russian” + raising a QuD
- So, something along the lines of “John thinks that there is something that Mencius drank. What did Mencius drink, if anything?” / “Masha thinks that you speak or do not speak Russian. Do you?”
- Note: both examples above are able to have a matrix-question interpretation. A possible s-selectional answer would be the following: (i) questions with declarative syntax are able to occur under declarative clause-embedding predicates; (ii) their semantics is such that the interpretation is invariably matrix

(22) \exists [Nick said that Mencius drank [WHAT]_F].

$\llbracket \text{what} \rrbracket = \{x \mid x \in D_e\}$

Interpretation: $\exists p \in \{\text{Nick said that Mencius drank } x \mid x \in D_e\} p(w_0)$

- Such a response would require the alternatives activated by the wh-word to necessarily percolate upwards, to the root node.
 - How can that be achieved?
 - The operator that flattens the set of propositions to a single proposition should be unavailable in an embedded context
 - So, it is ‘too high’. C-selection!

5 Conclusion

- Our starting point: declarative and interrogative clauses occur with different predicates
- Core question: is this a syntactic or semantic/pragmatic phenomenon
- Focusing on clauses that present a mismatch of syntactic properties and pragmatic behavior, I have argued in favor of syntactic analysis of clause type selection (largely repeating Bobaljik & Wurmbrand 2015).
- An open question: how does one account for semantic generalization re: rogative and anti-rogative predicates?
- I am inclined to think that those are **static** generalizations over the set of verbs that possess a certain formal co-occurrence restrictions (encoded via features)
 - The debate is equivalent to the role of phonological optimization in phonologically-conditioned suppletive allomorphy
 - Recent literature converges on the epiphenomenality of optimization (Embick 2010; Stanton 2021; Rolle 2021; Kalin & Rolle 2023)
 - See recent iterated language learning studies by Uegaki and colleagues for a proposal on how the semantic restrictions on rogative/anti-rogative predicates may come around (Uegaki & Qing 2024)

Acknowledgements

I thank Philip Shushurin, Stepa Mikhailov, Pavel Astafiev, Ivan Kalyakin, and Pavel Rudnev for discussion of these matters.

References

- Barwise, Jon & Robin Cooper. 1981. Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence: resources for processing natural language*, 241–301. Springer.
- Bhatt, Rajesh & Veneeta Dayal. 2020. Polar question particles: Hindi-Urdu kya. *Natural Language & Linguistic Theory* 38. 1115–1144.
- Biezma, María. 2020. Non-informative assertions: the case of non-optional wh-in-situ. *Semantics and Pragmatics* 13. 18–1.
- Bobaljik, Jonathan & Susi Wurmbrand. 2015. Questions with declarative syntax tell us what about selection. *MIT working papers in linguistics* 17. 13–32.
- Cheng, Lisa. 1997. *On the typology of wh-questions*. MIT dissertation.
- Del Pinal, Guillermo. 2022. The logicity of language: contextualism versus semantic minimalism. *Mind* 131(522). 381–427.
- Elliott, Patrick D. 2020. *Elements of clausal embedding*. UCL (University College London) dissertation.
- Embick, David. 2010. *Localism versus globalism in morphology and phonology*. MIT Press.

- Esipova, Masha & Natasha Korotkova. 2024. *To li or not to li*. Talk presented at FDSL 2024.
- Faure, Richard & Katerina Palasis. 2021. Exclusivity! Wh-fronting is not optional wh-movement in Colloquial French. *Natural Language & Linguistic Theory* 39. 57–95.
- von Fintel, Kai. 1993. Exeptive constructions. *Natural language semantics* 1(2). 123–148.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic inquiry* 10(2). 279–326.
- Hamblin, Charles L. 1976. Questions in montague english. In *Montague grammar*, 247–259. Elsevier.
- Kalin, Laura & Nicholas Rolle. 2023. Deconstructing subcategorization: conditions on insertion vs. conditions on position. *Linguistic Inquiry* 55(1). 197–218.
- Ladusaw, William Allen. 1979. *Polarity sensitivity as inherent scope relations*. The University of Texas at Austin dissertation.
- Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford Studies in Theoretical Linguistics.
- Moulton, Keir. 2009. *Natural selection and the syntax of clausal complementation*. University of Massachusetts Amherst dissertation.
- Odiijk, Jan. 1997. C-selection and s-selection. *Linguistic Inquiry*. 365–371.
- Pesetsky, David Michael. 1982. *Paths and categories*. Massachusetts Institute of Technology dissertation.
- Preminger, Omer. 2018. *Puzzle 3: The Definiteness Effect*. <https://omer.lingsite.org/files/NYI-2018-three-puzzles-handout.pdf>.
- Rolle, Nicholas. 2021. Against phonologically-optimizing suppletive allomorphy (posa) in irish, tiene, katu, and konni. *Acta Linguistica Academica* 68(1-2). 103–138.
- Rudin, Deniz. 2019. Embedded rising declaratives and embedded quotation. In *Semantics and linguistic theory*, 1–21.
- Shushurin, Philip. 2024. A head movement analysis of second position clitics: The case of russian polar particle li. *Syntax*.
- Stanton, Juliet. 2021. Allomorph selection precedes phonology: evidence from yindjibarndi. *Natural Language & Linguistic Theory*. 1–36.
- Tan, Rachel & Zheng Shen. 2024. True wh-movement and wh-in situ in one language: Evidence from Colloquial Singapore English. <https://ling.auf.net/lingbuzz/007749>.
- Uegaki, Wataru. 2015. *Interpreting questions under attitudes*. Massachusetts Institute of Technology dissertation.
- Uegaki, Wataru. 2022. *Question-orientedness and the semantics of clausal complementation*. Springer.
- Uegaki, Wataru & Ciyang Qing. 2024. *Semantic triviality leads to ungrammaticality through iterated learning*. A talk at NY Philosophy of Language workshop.
- Wurmbrand, Susanne. 2024. The size of clausal complements. *Annual Review of Linguistics* 10(1). 59–83.