Morphological Marking of Constituent Questions. A Case for Nonlocal Amalgamation

DELPH-IN 2021

Olga Zamaraeva

Department of Linguistics, University of Washington

July 20 2021

Introduction summary

- **Data:** Constituent (*wh*-) questions cross-linguistically

[1] Bouma et al. 2001; Ginzburg and Sag 2000

Introduction summary

- **Data:** Constituent (*wh*-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Grammars share "core", including definition of *list*

[1] Bouma et al. 2001; Ginzburg and Sag 2000

Introduction summary

- **Data:** Constituent (*wh*-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Grammars share "core", including definition of *list*
- **Theory:** Nonlocal amalgamation[1](NA; aka lexical threading)
  - Heads "append" arguments' nonlocal features

[1] Bouma et al. 2001; Ginzburg and Sag 2000

Introduction summary

- **Data:** Constituent (*wh*-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Grammars share "core", including definition of *list*
- **Theory:** Nonlocal amalgamation[1](NA; aka lexical threading)
  - Heads "append" arguments' nonlocal features
- **Problem:** NA complicates the analysis of multiple question fronting

[1] Bouma et al. 2001; Ginzburg and Sag 2000

Introduction summary

- **Data:** Constituent (*wh*-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Grammars share "core", including definition of *list*
- **Theory:** Nonlocal amalgamation[1](NA; aka lexical threading)
  - Heads "append" arguments' nonlocal features
- **Problem:** NA complicates the analysis of multiple question fronting
- **Problem:** But without NA, the analysis of morphological marking of questions is... questionable!

[1] Bouma et al. 2001; Ginzburg and Sag 2000

Introduction summary

- **Data:** Constituent (*wh*-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Grammars share "core", including definition of *list*
- **Theory:** Nonlocal amalgamation[1](NA; aka lexical threading)
  - Heads "append" arguments' nonlocal features
- **Problem:** NA complicates the analysis of multiple question fronting
- **Problem:** But without NA, the analysis of morphological marking of questions is... questionable!
- **Question:** What can be changed/improved?

1 Bouma et al. 2001; Ginzburg and Sag 2000

Data: Constituent questions

- Questions about *who* did *what to whom where*, etc.
- Different marking strategies across languages, including:
  - Question phrase fronting
  - Morphological marking

(1) Gde kto chto
where who.NOM what.ACC
vidit?
see.3SG
'Who sees what where?'
(Russian [rus]; IE)[2]

(2) eeva iche -ǯa -m?
what see -FUT.Q -1SG.Q
'What will I see?'
(Negidal [neg]; Tungusik)[3]

2 Constructed by a native speaker of Russian.
3 Hölzl 2018

Data: Constituent questions

- Fronting can be long distance
- Morphological marking can be distinct in polar vs. wh-

(3) Gde kto chto my
where who.NOM what.ACC 1PL.NOM
vyjasnili vidit?
find.out.PL.PAST see.3SG
'Who did we find out sees what where?' [rus][4]

(4) ačaq=qa dudu'k
who=CONTENT.3SG sing
'Who is singing?' (Makah [myh]; Wakashan)[5]

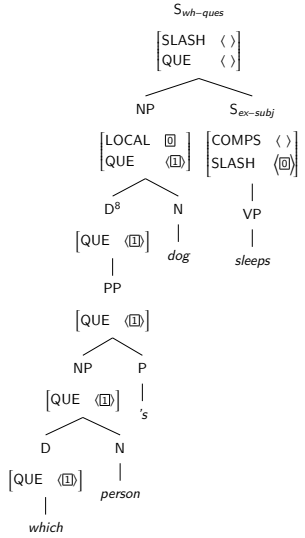
- **Goal:** Have a system of analyses for a range of phenomena such as above
  - All grammars share have list-valued features implemented the same way
  - ...SLASH, QUE

4 Constructed by a native speaker of Russian.
5 Davidson 2002

SLASH and QUE: Nonlocal dependencies

(5) Which person's (son's) dog (do you think) sleeps? [eng]

- SLASH creates LDD between the verb and its argument[6]
- QUE creates LDD within the (complex) argument NP
  - non-*wh* words have empty QUE
  - (Perhaps a better name: WH[7])



6 Pollard and Sag 1994

7 Ginzburg and Sag 2000

Nonlocal amalgamation[9]

- Idea: Head's NONLOCAL is the union of the daughters' NONLOCALs
- Motivation:
  - Fewer extraction rules required (in theory)
  - *easy*-adjectives: simply stipulate the argument has a gap (nonempty SLASH)
  - LDD can be encoded locally throughout the derivation (e.g. Chamorro)

$$\left[\begin{array}{l} \textit{basic-two-arg-lex-item} \\ \text{ARG-ST} \quad \left\langle \left[\text{NON-LOCAL} \left[\begin{array}{ll}\text{SLASH} & \boxed{1} \\ \text{REL} & \boxed{2} \\ \text{QUE} & \boxed{3}\end{array}\right]\right], \left[\text{NON-LOCAL} \left[\begin{array}{ll}\text{SLASH} & \boxed{4} \\ \text{REL} & \boxed{5} \\ \text{QUE} & \boxed{6}\end{array}\right]\right] \right\rangle \\ \text{SYNSEM|NON-LOCAL} \left[\begin{array}{ll}\text{SLASH|APPEND} & \langle \boxed{1}, \boxed{4} \rangle \\ \text{REL|APPEND} & \langle \boxed{2}, \boxed{5} \rangle \\ \text{QUE|APPEND} & \langle \boxed{3}, \boxed{6} \rangle\end{array}\right] \end{array}\right]$$

9 Bouma et al. 2001

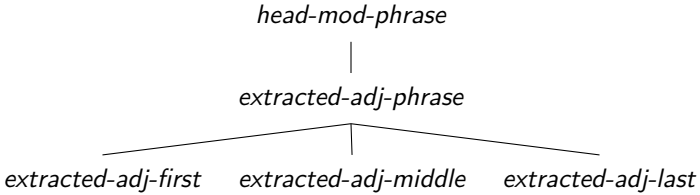
Extraction rules in DELPH-IN

- Extraction rules may not be needed for English but they probably are needed cross-linguistically
  - E.g. valence-changing morphology
- Bouma et al.'s analysis relies on DEPS (arguments and adjuncts together)
  - Not adopted in DELPH-IN; e.g. counting adjuncts is hard
- **Bottom line**: DELPH-IN maintains extraction rules
  - ...but NA is used in e.g. the English Resource Grammar,[10] for *easy*-adjectives

[10] Flickinger 2000, 2011

Multiple question fronting in DELPH-IN JRF[11]

- With the **combination** of DELPH-IN lists and NA:
- Extraction rules merely specify some list is nonempty
  - They do not extend or combine SLASH sets/lists
  - Need to say: An adjunct is extracted before/after/between the arguments
- Implementing multiple question phrase fronting with flexible word order thus necessitates **even more** extraction rules



11 Zamaraeva and Emerson 2020

Summary of introduction

- Goal: Have a system of analyses (the Grammar Matrix) covering multiple question phrase fronting **as well** as other phenomena
  - ...cross-linguistically, way beyond just English or just IE languages

Summary of introduction

- Goal: Have a system of analyses (the Grammar Matrix) covering multiple question phrase fronting **as well** as other phenomena
  - ...cross-linguistically, way beyond just English or just IE languages
- Zamaraeva and Emerson (2020) argued that NA complicates the system
  - If you extract explicitly, append NONLOCAL explicitly to avoid extra rules

Summary of introduction

- Goal: Have a system of analyses (the Grammar Matrix) covering multiple question phrase fronting **as well** as other phenomena
  - ...cross-linguistically, way beyond just English or just IE languages
- Zamaraeva and Emerson (2020) argued that NA complicates the system
  - If you extract explicitly, append NONLOCAL explicitly to avoid extra rules
- This talk: A counterpoint:

Summary of introduction

- Goal: Have a system of analyses (the Grammar Matrix) covering multiple question phrase fronting **as well** as other phenomena
  - ...cross-linguistically, way beyond just English or just IE languages
- Zamaraeva and Emerson (2020) argued that NA complicates the system
  - If you extract explicitly, append NONLOCAL explicitly to avoid extra rules
- This talk: A counterpoint:
  - Morphological marking of interrogative constructions

Summary of introduction

- Goal: Have a system of analyses (the Grammar Matrix) covering multiple question phrase fronting **as well** as other phenomena
  - ...cross-linguistically, way beyond just English or just IE languages
- Zamaraeva and Emerson (2020) argued that NA complicates the system
  - If you extract explicitly, append NONLOCAL explicitly to avoid extra rules
- This talk: A counterpoint:
  - Morphological marking of interrogative constructions
  - ...Much simpler with NA!

Summary of introduction

- Goal: Have a system of analyses (the Grammar Matrix) covering multiple question phrase fronting **as well** as other phenomena
  - ...cross-linguistically, way beyond just English or just IE languages
- Zamaraeva and Emerson (2020) argued that NA complicates the system
  - If you extract explicitly, append NONLOCAL explicitly to avoid extra rules
- This talk: A counterpoint:
  - Morphological marking of interrogative constructions
  - ...Much simpler with NA!
    - ...for a certain typological profile at least

Morphologically marked interrogatives

(6) oǯa-va iche-ǯee-v
track-ACC see-FUT-1SG
'I will see the tracks.' [neg][12]

(7) ii-ǰə-m =i?
enter-FUT.Q-1SG.Q =Q
'Shall I come in?' [neg]

(8) eeva iche-ǯa-m?
what see-FUT.Q-1SG.Q
'What will I see?' [neg]

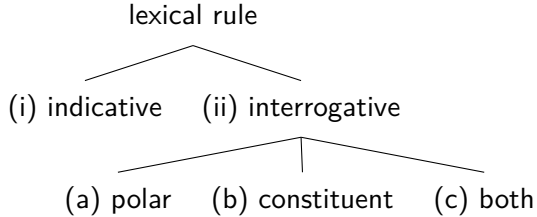
(9) ʔačaq=qaːɬ dudu'k
who=CONTENT.3SG sing
'Who is singing?' [myh][13]

(10) dudu'k='aƛ=qaːk=s
sing=TEMP=POLAR=1SG
'Am I singing?' [myh]

12 Hölzl 2018
13 Davidson 2002

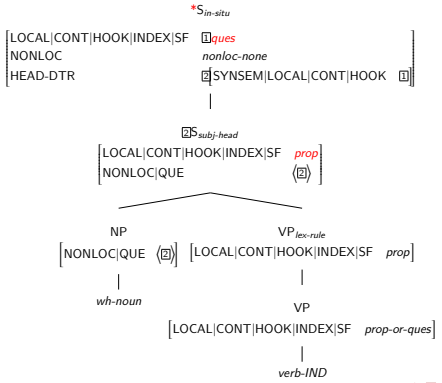
Morphologically marked interrogatives: Typology

- Special paradigm(s) for interrogatives:
  - Polar and constituent questions may have **distinct** paradigms
  - In DELPH-IN JRF:
    - Modeling the (i) vs (ii),(c) distinction is easy with or without NA
    - Modeling (a)–(b) distinction without NA is **not trivial** without NA



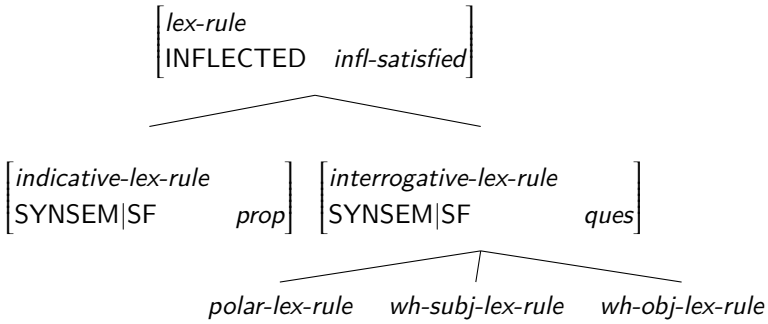
Indicative vs. interrogative, NA does not matter

- Distinction between (i) indicative and (ii) interrog. lex. rules is easy
  - (c) by extension (same as (ii))
- The sentential force SF semantic feature will block any interrogative phrase structure rule



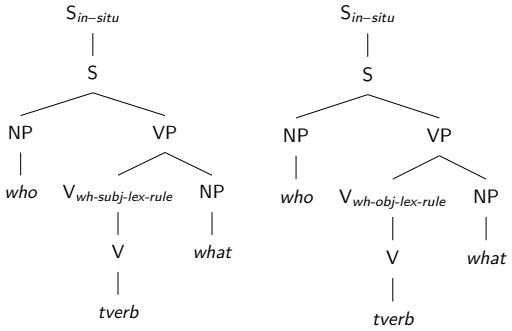
Analysis without nonlocal amalgamation: (a) vs (b)

- Lex. rules for *wh-* (and not polar) questions need to explicitly posit which argument of the head is or isn't *wh*
  - No way to just say: **Some** argument is *wh* (in DELPH-IN JRF)



Analysis without nonlocal amalgamation: (a) vs (b)

- But, the *wh-obj-lex-rule* will apply spuriously!
  - ...in languages where there is only one morpheme to mark any *wh*-question
  - The second argument ends up underspecified (*wh-* or not)
  - Cannot constrain its SUBJ to be empty (saturated)
  - ...would violate the assumption that lexical rules apply before phrasal



Analysis without nonlocal amalgamation: (a) vs (b)

$$\begin{bmatrix} \textit{non-wh-cons} & \\ \text{FIRST} & \begin{bmatrix} \textit{synsem} \\ \text{NON-LOCAL.QUE.LIST} \quad \langle\,\rangle \end{bmatrix} \\ \text{REST} & \textit{non-wh-list} \end{bmatrix}$$

$$\begin{bmatrix} \textit{polar-lex-rule} & \\ \text{SYNSEM|LOCAL|CAT|VAL} & \begin{bmatrix} \text{SUBJ} & \left\langle \begin{bmatrix} \text{NON-LOCAL|QUE|LIST} \quad \langle\,\rangle \end{bmatrix} \right\rangle \\ \text{COMPS} & \textit{non-wh-list} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textit{wh-subj-lex-rule} \\ \text{SYNSEM|LOCAL|CAT|VAL|SUBJ} \quad \left\langle \begin{bmatrix} \text{NON-LOCAL|QUE|LIST} \quad \textit{cons} \end{bmatrix} \right\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{wh-obj-lex-rule} & \\ \text{SYNSEM|LOCAL|CAT|VAL} & \begin{bmatrix} \text{SUBJ} & \textit{non-wh-list} \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{NON-LOCAL|QUE|LIST} \quad \textit{cons} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Analysis with nonlocal amalgamation

- With NA, **can** say: **some** arg is *wh*!
- It is the same as to say QUE *cons*!
- For (c), just leave QUE underspecified
- No need to think about number or order of args!
- No need to posit any additional types beyond the following two:

$$\begin{bmatrix} \textit{polar-lex-rule} & \\ \text{SYNSEM|SF} & \textit{ques} \\ \text{DTR|SYNSEM|NON-LOCAL|QUE|LIST} & \langle\,\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{wh-lex-rule} & \\ \text{SYNSEM|SF} & \textit{ques} \\ \text{DTR|SYNSEM|NON-LOCAL|QUE|LIST} & \textit{cons} \end{bmatrix}$$

Conclusion

- In DELPH-IN JRF, treatment of morphological marking and fronting of questions[14]seem to be in competition?
- Nonlocal amalgamation[15]seems important for morphological marking
  - Analysis is easy both conceptually and in terms of implementation
- we **probably** do want to have one core for all grammars
- It complicates multiple fronting with flexible word order but perhaps this means more work on word order is required?
- Or maybe languages like Makah are very rare?..

14 Bender, Flickinger, et al. 2002; Bender, Drellishak, et al. 2010
15 Zamaraeva 2021
16 Zamaraeva and Emerson 2020
17 Bouma et al. 2001

A question from a reviewer

- How would a lexical verb be sensitive to QUE?

"The lexical verb can see SLASH elements, but QUE is percolated in a different area of the structures. Since QUE is a nonlocal feature and only local features are shared between filler and gap, the QUE value is not available at the extraction site, i.e. at the verb. Maybe no extraction is involved in the respective languages but if there is, the analysis seems to not work"

References I

Bouma, Gosse, Robert Malouf, and Ivan Sag (2001). "Satisfying Constraints on Extraction and Adjunction". In: *Natural Language & Linguistic Theory* 19.1, pp. 1–65.

Ginzburg, Jonathan and Ivan Sag (2000). *Interrogative investigations*. Stanford: CSLI publications.

Hölzl, Andreas (2018). *A typology of questions in Northeast Asia and beyond: An ecological perspective*. Berlin: Language Science Press.

Davidson, Matthew (2002). "Studies in Southern Wakashan (Nootkan) grammar". PhD thesis. Buffalo, NY: University of New York at Buffalo.

Pollard, Carl and Ivan Sag (1994). *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.

Nielsen, Elizabeth K (2018). "Modeling adnominal possession in the lingo grammar matrix". MA thesis. University of Washington.

Flickinger, Dan (2000). "On building a more efficient grammar by exploiting types". In: *Natural Language Engineering* 6.01, pp. 15–28.

— (2011). "Accuracy v. Robustness in Grammar Engineering". In: *Language from a Cognitive Perspective: Grammar, Usage and Processing*. Ed. by Emily M Bender and Jennifer E. Arnold. Stanford, CA: CSLI Publications, pp. 31–50.

Zamaraeva, Olga and Guy Emerson (2020). "Multiple Question Fronting without Relational Constraints: An analysis of Russian as a basis for cross-linguistic modeling". In: *Proceedings of the 27th International Conference on Head-Driven Phrase Structure Grammar*. Virtual conference, to appear.

Bender, Emily M, Dan Flickinger, and Stephan Oepen (2002). "The Grammar Matrix: An Open-Source Starter-Kit for the Rapid Development of Cross-Linguistically Consistent Broad-Coverage Precision Grammars". In: *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*. Ed. by John Carroll, Nelleke Oostdijk, and Richard Sutcliffe. Taipei, Taiwan, pp. 8–14.

Bender, Emily M, Scott Drellishak, et al. (2010). "Grammar Customization". In: *Research on Language & Computation* 8.1, pp. 23–72. ISSN: 1570-7075.

Zamaraeva, Olga (2021). "A cross-linguistic analysis of constituent questions for the Grammar Matrix". PhD thesis. University of Washington.