Morphological Marking of Constituent Questions.
A Case for Nonlocal Amalgamation

HPSG 2021
Olga Zamaraeva

Department of Linguistics, University of Washington
July 2021
Introduction summary

- **Data:** Constituent (wh-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Implemented system of HPSG grammars using one “core”
  - Restricted version of formalism, esp. wrt lists

---

1. Copestake 2000
2. Bouma et al. 2001; Ginzburg and Sag 2000
Introduction summary

- **Data:** Constituent (*wh-* ) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Implemented system of HPSG grammars using one “core”
  - Restricted version of formalism, esp. wrt lists
- **Theory:** Nonlocal amalgamation
  - Heads “append” arguments’ nonlocal features

---

1 Copestake 2000
2 Bouma et al. 2001; Ginzburg and Sag 2000
Introduction summary

▶ **Data:** Constituent *(wh-)* questions cross-linguistically
▶ **Project:** The Grammar Matrix
  ▶ Implemented system of HPSG grammars using one “core”
  ▶ Restricted version of formalism, esp. wrt lists ¹
▶ **Theory:** Nonlocal amalgamation ²
  ▶ Heads “append” arguments’ nonlocal features
▶ **Problem:** NA complicates the analysis of multiple question fronting

---

¹ Copestake 2000
² Bouma et al. 2001; Ginzburg and Sag 2000
Introduction summary

- **Data:** Constituent (wh-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Implemented system of HPSG grammars using one “core”
  - Restricted version of formalism, esp. wrt lists
- **Theory:** Nonlocal amalgamation
  - 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. Copestake 2000
2. Bouma et al. 2001; Ginzburg and Sag 2000
Introduction summary

- **Data:** Constituent (wh-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Implemented system of HPSG grammars using one “core”
  - Restricted version of formalism, esp. wrt lists
- **Theory:** Nonlocal amalgamation
  - 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!
- **Conclusion:** Choose between formalism restrictions and sharing the core?

---

1 Copestake 2000
2 Bouma et al. 2001; Ginzburg and Sag 2000
Introduction summary

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

- **Project**: The Grammar Matrix
  - Implemented system of HPSG grammars using one “core”
  - Restricted version of formalism, esp. wrt lists

- **Theory**: Nonlocal amalgamation
  - 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!

- **Conclusion**: Choose between formalism restrictions and sharing the core?
  - ...or, reanalyze fronting with flexible word order?

---

1. Copestake 2000
2. Bouma et al. 2001; Ginzburg and Sag 2000
Introduction summary

- **Data:** Constituent (wh-) questions cross-linguistically
- **Project:** The Grammar Matrix
  - Implemented system of HPSG grammars using one “core”
  - Restricted version of formalism, esp. wrt lists
- **Theory:** Nonlocal amalgamation
  - 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!
- **Conclusion:** Choose between formalism restrictions and sharing the core?
  - ...or, reanalyze fronting with flexible word order?
  - ...or/and, revisit arguments/adjuncts distinction

---

1. Copestake 2000
2. 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)³

(2) eeva iche -¿a -m?
    what see -FUT.Q -1SG.Q
    ‘What will I see?’
    (Negidal [neg]; Tungusik)⁴

---

³ Constructed by a native speaker of Russian.
⁴ 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]\(^5\)

(4) ačaq=qa dudu’k who=CONTENT.3SG sing
‘Who is singing?’ (Makah [myh]; Wakashan)\(^6\)

- **Goal:** Have a system of analyses for a range of phenomena such as above
  - All grammars share the same core

---

\(^5\) Constructed by a native speaker of Russian.
\(^6\) Davidson 2002
The Grammar Matrix

- Meta-grammar engineering framework\(^7\)
- Input: Typological specification, lexicon, morphological rules
- Output: Implemented HPSG grammar fragment
  - Parse and generate sentences
  - Output syntactic and semantic representations
- Many syntactic phenomena are supported\(^8\)
  - Most recently: \textit{wh}-questions\(^9\)

\(^{7}\) https://matrix.ling.washington.edu/customize/matrix.cgi
\(^{8}\) Zamaraeva, Howell, et al. 2019; Howell and Zamaraeva 2018; Saleem 2010; Song 2014; Nielsen 2018; Drellishak and Bender 2005; Crowgey 2013; Bender and Flickinger 2005; Zamaraeva 2021
\(^{9}\) Zamaraeva 2021; Zamaraeva and Emerson 2020
A restricted version of HPSG\textsuperscript{10}

Unification the only native operation

\begin{itemize}
  \item i.e. no shuffle operator, no linearization
  \item Number and order of daughters are fixed (lists have fixed, bounded length)
  \item List append has to be explicitly encoded\textsuperscript{11}
\end{itemize}

\[
\begin{array}{c}
  \text{append-list} \\
  \text{LIST} \\
  \text{APPEND} \\
\end{array}
\]

\[
\begin{array}{c}
  \text{0list} \\
  \text{list} \\
  \text{APPEND-RESULT 0}
\end{array}
\]

\textsuperscript{10} Copestake 2000

\textsuperscript{11} Copestake 2000; Zamaraeva and Emerson 2020; Emerson 2017, 2019
List-valued features in DELPH-IN HPSG

- **Valence** (SUBJ, COMPS, adjuncts (MOD))
  - No DEPS list combining arguments and adjuncts

- **Semantics** (RELS, CONT, ICONS)

- **Nonlocal** (SLASH, QUE, REL)
  - QUE necessary for *wh*-question semantics and for pied piping; SLASH for any kind of fronting/dislocation
SLASH and QUE: Nonlocal dependencies

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

› SLASH creates LDD with the verb argument

› QUE creates LDD with the wh-word

› non-wh words have empty QUE

› (Perhaps a better name: WH)

---

12 Pollard and Sag 1994
13 Ginzburg and Sag 2000
14 Nielsen 2018
Nonlocal amalgamation

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

\[
\begin{align*}
\text{basic-two-arg-lex-item} & \\
\text{ARG-ST} & \begin{cases}
\text{NON-LOCAL} & \text{SLASH } 1 \\
\text{REL} & 2 \\
\text{QUE} & 3 \\
\end{cases}
\end{align*}
\]
\[
\begin{align*}
\text{SYNSEM|NON-LOCAL} & \begin{cases}
\text{SLASH|APPEND} & \langle 1, 4 \rangle \\
\text{REL|APPEND} & \langle 2, 5 \rangle \\
\text{QUE|APPEND} & \langle 3, 6 \rangle \\
\end{cases}
\end{align*}
\]

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,\(^{16}\) for easy-adjectives

---

\(^{16}\) Flickinger 2000, 2011
Multiple question fronting in DELPH-IN HPSG

- 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

  \[
  \text{head-mod-phrase} \\
  \quad | \\
  \text{extracted-adj-phrase}
  \]

  \[
  \text{extracted-adj-first} \quad \text{extracted-adj-middle} \quad \text{extracted-adj-last}
  \]

---

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 and should be removed
  - If you extract explicitly, append NONLOCAL explicitly to avoid extra rules
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 and should be removed

- 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 and should be removed
  - 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 and should be removed
  - 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 and should be removed
  - 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  iče-ǯee-v
     track-ACC see-FUT-1SG
‘I will see the tracks.’ [neg]\textsuperscript{18}

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

(8)  eevə  iče-ǯə-m?
     what see-FUT.Q-1SG.Q
‘What will I see?’ [neg]

(9)  ?ačaq=qaːś  dudu’k
     who=CONTENT.3SG  sing
‘Who is singing?’ [myh]\textsuperscript{19}

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

\textsuperscript{18} Hölzl 2018
\textsuperscript{19} Davidson 2002
Morphologically marked interrogatives: Typology

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

```
lexical rule

(i) indicative   (ii) interrogative

(a) polar       (b) constituent   (c) both
```
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
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 HPSG)

\[
\begin{bmatrix}
\text{lex-rule} \\
\text{INFLECTED} \\
\text{infl-satisfied}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{indicative-lex-rule} \\
\text{SYNSEM|SF} \\
\text{prop}
\end{bmatrix}
\begin{bmatrix}
\text{interrogative-lex-rule} \\
\text{SYNSEM|SF} \\
\text{ques}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\text{polar-lex-rule} \\
\text{wh-subj-lex-rule} \\
\text{wh-obj-lex-rule}
\end{bmatrix}
\]
But, the *wh-obj-lex-rule* will apply spuriously!

- ...in languages where there is only one morpheme to mark any *wh*-question
- Cannot constrain it’s *SUBJ* to be empty (saturated)
- ...would violate the assumption that lexical rules apply before phrasal
Analysis without nonlocal amalgamation: (a) vs (b)

\[
\begin{align*}
\text{non-wh-cons} & : [ \text{synsem} \\
\text{FIRST} & : \text{NON-LOCAL.QUE.LIST} \langle \rangle \\
\text{REST} & : \text{non-wh-list} \\
\text{polar-lex-rule} & : \text{SYNSEM|LOCAL|CAT|VAL} \\
\text{wh-subj-lex-rule} & : \text{SYNSEM|LOCAL|CAT|VAL|SUBJ} \\
\text{wh-obj-lex-rule} & : \text{SYNSEM|LOCAL|CAT|VAL}
\end{align*}
\]
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:

\[
\begin{align*}
\text{polar-lex-rule} & \\
\text{SYNSEM}\vert\text{SF} & \quad \text{ques} \\
\text{DTR}\vert\text{SYNSEM}\vert\text{NON-LOCAL}\vert\text{QUE}\vert\text{LIST} & \quad \langle \rangle
\end{align*}
\]

\[
\begin{align*}
\text{wh-lex-rule} & \\
\text{SYNSEM}\vert\text{SF} & \quad \text{ques} \\
\text{DTR}\vert\text{SYNSEM}\vert\text{NON-LOCAL}\vert\text{QUE}\vert\text{LIST} & \quad \text{cons}
\end{align*}
\]
Conclusion

▶ Presented an analysis of morphological marking in DELPH-IN HPSG
  ▶ Implemented as part of the Grammar Matrix\textsuperscript{20}
  ▶ Implementation tested on Makah [myh] (Wakashan) and pseudolangages\textsuperscript{21}
▶ In DELPH-IN HPSG, treatment of morphological marking and fronting of questions\textsuperscript{22} seem to be in competition
▶ Nonlocal amalgamation\textsuperscript{23} seems particularly important for morphological marking
  ▶ Analysis is easy both conceptually and in terms of implementation
▶ It complicates multiple fronting with flexible word order but perhaps this indicates more work on word order is required?
▶ ...or revisiting the arguments/adjuncts distinction is in order?\textsuperscript{24}

\textsuperscript{20} Bender, Flickinger, and Oepen 2002; Bender, Drellishak, et al. 2010
\textsuperscript{21} Zamaraeva 2021
\textsuperscript{22} Zamaraeva and Emerson 2020
\textsuperscript{23} Bouma et al. 2001
\textsuperscript{24} Przepiórkowski 2016
References


References


