ON WACKERNAGEL TEN YEARS LATER: PERSONAL PRONOUNS IN INFORMATION STRUCTURE
Anita Peti-Stantić, University of Zagreb
Croatian Science Foundation project no. HRZZ-IP-2016-06-1210

12th Slavic Linguistics Society Annual Meeting, Ljubljana
September 23, 2017
AN OUTLINE OF THE TALK

1. PROJECT MEGAHR/MEGACRO (The Building Blocks of Mental Grammar in Croatian: Constraints of Information Structure)

2. THEORETICAL ACCOUNT: INFORMATION STRUCTURE AND PARALLEL ARCHITECTURE (Ray Jackendoff)

3. DATA AND ANALYSIS: CLITIC CHOICE AND PLACEMENT IN CROATIAN (WO and PROSODIC PHRASING)

4. CONCLUSIONS
GOALS OF THE TALK

1. to give an additional insight into the constraints on information structure and its role in the sentence organization

2. to show the necessity of combining all three modules in an analysis (conceptual, phonological and syntactic)

3. to confront the simplicity of classical sentences usually used for linguistic analysis and complexity of all other sentences attested in corpus and in real communication
MAIN RESEARCH QUESTION

- How is Wackernagel’s rule relevant in Croatian?
  - Does it predict ‘neutral’ clitic/clitic cluster position?

- How syntax and prosody interact in Croatian?
  - Which are the constraints at the interface of syntactic and prosodic representation?
THE BUILDING BLOCKS OF MENTAL GRAMMAR IN CROATIAN: CONSTRAINTS OF INFORMATION STRUCTURE

- based on corpus and psycholinguistic research
- we look into how speakers process morphosyntactic and semantic agreement of concrete and/or abstract lexical items in coordinated constructions and elliptical conditions, as well as how they process sentence information structure with regard to complex clitic cluster positioning
- AIM: to determine a set of necessary, universal and language-specific constraints representing the requirements that govern the information structure in Croatian, both on the morphosyntactic and the semantic level
  - we are predominantly interested in local (linear) and global (hierarchical) processing constraints within the mental grammar of Croatian
- 4 years (started March 10, 2017)
THE BUILDING BLOCKS OF MENTAL GRAMMAR IN CROATIAN: CONSTRAINTS OF INFORMATION STRUCTURE

+ METHODOLOGICAL FRAMING: psycholinguistics and corpus linguistics
+ (1) to establish a correlation between concreteness and imageability for a set of high and low frequency lexical items (3000 words) and grammatical items (constructions, full forms and clitics)
+ (2) to establish the relationship between concreteness and agreement (in gender, ellipsis and clitic cluster positioning)
+ (3) to establish language specific constraints/patterns for building effective word and grammar webs within the educational system
+ THEORETICAL GOALS:
  + to enhance existing models of language computational modeling
  + to establish connections between global and local building blocks of the mental grammar of Croatian
THE BUILDING BLOCKS OF MENTAL GRAMMAR IN CROATIAN: CONSTRAINTS OF INFORMATION STRUCTURE

- by now 3000 lexems from HrWaC corpus
  - high and low frequency (nouns, verbs and adjectives)
  - concretness and abstractness
  - relative frequency
  - 900 participants
  - data collected and quantitatively analyzed

- next step – to test same lexemes for imageability and AoA
COLLABORATORS AND EXPERTS
linguistics, psycholinguistics, developmental psychology, natural language processing, computational modeling, neurolinguistic modeling
**INFORMATION STRUCTURE**

- IS is NOT just pragmatics
- IS is the meeting point of the phonological, syntactic and semantic representations
- the conceptual structure of the sentence does not depend on the semantics of individual elements or on the syntax of individual constructions, **but on the interaction between interface constraints**
INFORMATION STRUCTURE AND PARALLEL ARCHITECTURE

- Phonological formation rules
  - Phonological structures
  - Interface
- Syntactic formation rules
  - Syntactic structures
  - Interface
- Conceptual formation rules
  - Conceptual structures
  - Interface

The diagram illustrates the relationship between phonological, syntactic, and conceptual structures, showing how rules translate into structures and interfaces.
PRONOUNS IN CROATIAN

- so called long or full and short or (en)clitic forms
- GEN mene/me, tebe/te, njega.ga etc.
- DAT meni/mi, tebi/ti, njemu/mu etc.
- ACC mene/me, tebe/te, njega.ga etc.

- both serve the same syntactic function
- (en)clitics are prosodically defective, need a prosodic host on their left side
- clitic placement – according to Wackernagel rule
  - second position
  - INTERPLAY between prosodic and syntactic second position
PRONOUNS IN CROATIAN

- **mene**
  - morphology
  - unrestricted in terms of phrasal phonology and almost unrestricted in WO

- **me**
  - restricted in terms of phrasal phonology and WO – Wackernagel’s law, but...
  - **HYPOTHESIS:** Wackernagel’s law is not a constraint on a clausal, but on a phrasal level – this authorizes clitics to move in a linear structure towards the end of a sentence
LONG/FULL FORMS

+ 1.1 Meni studentica donosi knjige. (nobody would say this (without a good reason))
   + ‘Student brings books to me.’
   + (Mental Grammar) repair: Meni studentica donosi knjige, a ne tebi.

+ 1.2 Studentica meni donosi knjige. (again, nobody...)
   + (MG) repair: Studentica meni donosi knjige, a ne tebi.

+ 1.3 Studentica donosi meni knjige. (again, nobody...)
   + (MG) repair: Studentica donosi meni knjige, a ne tebi.

+ 1.4 Studentica donosi knjige meni. (again, nobody...)
   + (MG) repair: Studentica donosi knjige meni, a ne tebi.

+ contrastive focus needed for an interpretation (not necessarily explicit)

+ WHY?

+ Long forms are unrestricted in terms of phrasal phonology.

+ With the long form, all the ‘information weight’ is already in the choice of the form itself.

+ Because of the information structure.

+ When asked about the grammaticality or ‘naturalness’, native speakers grant their preference to sentences 1.1 and 1.4, while 1.2 and 1.3 receive very low ratings (1.2 being judged somewhat better)
LONG/FULL AND CLITIC FORMS

2.1. Studentica *meni* donosi knjige.  
+ focused

2.2. Dobra studentica *meni* donosi knjige.  
+ focused

2.3. *Dobra* *meni* studentica donosi knjige.  

2.4. Dobra studentica donosi *meni* knjige.  
+ focused

CONCLUSION: so called long/full pronominal forms always bring in the focus in a sentence/utterance

3.1. Studentica *mi* donosi knjige.  
+ syntactic and prosodic positions match

3.2. Dobra studentica *mi* donosi knjige.  
+ syntactically generated position, but can be treated as a position where syntactic and prosodic positions match

3.3. Dobra *mi* studentica donosi knjige.  
+ prosodically generated position

3.4. Dobra studentica donosi *mi* knjige.  
+ position generated at the prosody-syntax interface

How?
+ complex first syntactic constituent, relatively heavy
+ you start counting for prosodical purposes all over again from the verb on
CLITICS FROM THE POINT OF INFORMATION STRUCTURE

+ Wackernagel says nothing about the information structure and ‘neutrality’

+ 3.1. Studentica mi donosi knjige.
  + [S[NP[N studentica]] [VP {CL mi} [V donosi] [NP [N knjige]]]]
  + (U (IP studentica mi donosi knjige))

+ 3.2. Dobra studentica mi donosi knjige.
  + [S[NP[AP dobra [N studentica]]] [VP {CL mi} [V donosi] [NP [N knjige]]]]
  + (U (IP dobra studentica mi donosi knjige))

+ 3.3. Dobra mi studentica donosi knjige.
  + [S[NP[AP dobra [N studentica]]] [VP [V {CL mi} donosi] [NP [N knjige]]]]
  + (U (IP dobra mi studentica (PP donosi knjige)))

+ 3.4. Dobra studentica donosi mi knjige.
  + [S[NP[AP dobra [N studentica]]] [VP [V donosi] {CL mi} [NP [N knjige]]]]
  + (U (IP (PP dobra studentica) (PP donosi mi knjige)))

+ CONCLUSION: when prosody and syntax do not match, i.e. when prosodically constraint, relevant for specific IS reading/interpretation
4.1. Studentica socijalne psihologije *mi* donosi brojne nove i zanimljive knjige.

4.1.a [S [NP [N studentica] [AP GEN socijalne psihologije]] [VP {CL mi} [V donosi] [NP [AP brojne nove i zanimljive knjige]]]]

4.1.b (U (IP studentica socijalne psihologije mi donosi brojne nove i zanimljive knjige))

4.2. Dobra studentica socijalne psihologije *mi* donosi brojne nove i zanimljive knjige.

4.2.a [S [NP [AP dobra [N studentica]] [AP GEN socijalne psihologije]] [VP {CL mi} [V donosi] [NP [AP brojne nove i zanimljive knjige]]]]

4.2.b (U (IP dobra studentica socijalne psihologije mi donosi brojne nove i zanimljive knjige))

5.1. Studentica socijalne psihologije donosi *mi* brojne nove i zanimljive knjige.

5.1.b (U (IP (PP studentica socijalne psihologije) (PP donosi mi brojne nove i zanimljive knjige)))

5.2. Dobra studentica socijalne psihologije donosi *mi* brojne nove i zanimljive knjige.

5.2.b (U (IP (PP dobra studentica socijalne psihologije) (PP donosi mi brojne nove i zanimljive knjige)))
THE LONGER, THE MERRIER... BUT COMPLEXITY MATTERS!

CONCLUSIONS:

- multi-word constituents matter
- prosodic phrasing matters - \((u (ip (pp dobra studentica socijalne psihologije) ((pp mi donosi brojne nove i zanimljive knjige)))\) should not be grammatical because of enclitic
WHAT DO WE EXPECT?  
‘NEUTRALITY’.  
BUT...

- based on 3.2. *Dobra studentica mi donosi knjige* neutrality, we expect 6.1. to be neutral too

- 6.1. Studentica socijalne psihologije iz Zagreba *mi* donosi brojne nove i zanimljive knjige.
  - ‘The student of social psychology brings me many new and interesting books from Zagreb/FROM ZAGREB.’

- potential syntactic structures

  6.1. a [S[N,P[N, studentica][AP,GEN socijalne psihologije][PP, iz Zagreba][VP, {CL, mi}][V, donosi][NP, brojne nove i zanimljive knjige]]]

  6.1. b [S[N,P[N, studentica][AP,GEN socijalne psihologije][VP, PP, iz Zagreba]{CL, mi}][V, donosi][NP, brojne nove i zanimljive knjige]]]

- potential prosodic structures

  6.1. c (IP, studentica socijalne psihologije iz Zagreba) (IP, mi donosi brojne nove i zanimljive knjige))

  6.1. d (IP, studentica socijalne psihologije) (IP, iz Zagreba mi donosi brojne nove i zanimljive knjige))
WHAT DO WE GET?

+ **CONSTRAINTS ON INFORMATION STRUCTURE**: interplay between syntax, linearization/prosody and the length of the phrase
  + Nespor and Vogel’s (1986) hypothesis of prosodic structure or sentential prosodic alignment
  + not all potential structures get to be realized
  + relevant complexity of NP (Genitive *socijalne psihologije*, PP *iz Zagreba*)

+ 6.1. Studentica socijalne psihologije iz Zagreba *mi* donosi brojne nove i zanimljive knjige.
  + ‘The student of social psychology brings me many new and interesting books FROM ZAGREB.’

+ realized syntactic structure
  + 6.1.b \[S [NP [N studentica] [AP GEN socialne psihologije] [VP [PP iz Zagreba] {CL mi} [V donosi] [NP [AP brojne nove i zanimljive knjige]]]]

+ realized prosodic structure
  + 6.1.d \((IP studentica socijalne psihologije) (IP iz Zagreba mi donosi brojne nove i zanimljive knjige))
INTERIM CONCLUSION

+ long forms invariably attract information interpretation centered around themselves, while the short forms, through their position in the linear structure, reorganize the information structure of the sentence

+ the informative power of clitic/clitic cluster position depends on the syntactic complexity of multi-word constituents

+ psycholinguistic research needed
  + to test ‘neutral’ and context-bound variants
  + to test mental grammar variations
WHAT DO WE EXPECT?
‘FOCUS’. BUT...

+ based on focus in 3.4. *Dobra studentica donosi mi knjige*, here we expect focusing too

+ 7.1. Studentica socijalne psihologije iz Zagreba donosi *mi* brojne nove i zanimljive knjige.
  + ‘The student of social psychology from Zagreb BRINGS/brings me many new and interesting books.’

+ potential syntactic structures
  + 7.1.a [S [NP [N studentica] [GEN socialne psihologije] [PP iz Zagreba]] [VP [V donosi] {CL mi}] [NP [AP brojne nove i zanimljive knjige]]]
  + 7.1.b [S [NP [N studentica] [GEN socialne psihologije]] [VP [PP iz Zagreba] [V donosi] {CL mi}] [NP [AP brojne nove i zanimljive knjige]]]

+ potential prosodic structures
  + 7.1.c (U (IP studentica socijalne psihologije iz Zagreba) (IP donosi mi brojne nove i zanimljive knjige))
  + 7.1.d (U (IP studentica socijalne psihologije) (IP (PP iz Zagreba) (PP donosi mi brojne nove i zanimljive knjige)))
WHAT DO WE GET?

+ again, CONSTRAINTS ON INFORMATION STRUCTURE: interplay between syntax, linearization/prosody and the length of the phrase
+ not all potential structures get to be realized

+ 7.1. Studentica socijalne psihologije iz Zagreba donosi *mi* brojne nove i zanimljive knjige.
+ ‘The student of social psychology from Zagreb brings me many new and interesting books.’

+ realized syntactic structure
  + 7.1.a \([S \[NP \[N \text{studentica}\] \[GEN \text{socijalne psihologije}\] \[PP \text{iz Zagreba}\]\] \[VP \[V \text{donosi}\}\{CL \text{mi}\}\]\] \[NP \[AP \text{brojne nove i zanimljive knjige}\]\]\

+ realized prosodic structure
  + 7.1.c \((IP \text{studentica socijalne psihologije iz Zagreba}) (IP \text{donosi mi brojne nove i zanimljive knjige})\)
CONCLUSIONS

+ How is Wackernagel’s rule relevant in Croatian?
  + both as a syntactic and prosodic constraint, but also at the interfaces
  + therefore, clitic/clitic cluster is always second in an IP (intonational phrase), not necessarily second in the sentence

+ Does it predict ‘neutral’ clitic/clitic cluster position?
  + not except for syntactically simplest sentences Studentica mi donosi knjige. vs. *Mi studentica donosi knjige. and Studentica donosi mi knjige.
  + repair: Studentica donosi mi knjige, a ne netko drugi. not neutral, activates repair, or specific information structure interpretation, contrastive focus
  + contrastive focus does not have to be explicit, but it’s always in the mental grammar of a speaker and hearer
CONCLUSIONS

How syntax and prosody interact in Croatian?

- My analysis clearly shows that potential shifting of clitic positions in the linear organization of a sentence is not arbitrary and that different positions serve to enable different conceptual readings within the range of possible positions (in Croatian no IP, no position after conjunctions a and i, no position following prepositions, except for Accusative clitic)

Which are the constraints at the interface of syntactic and prosodic representation?

- IS conditioning
- in Croatian, mostly prosodic constraints (in terms of phonological phrasing)
- therefore, I see information structure as a part of grammatical, and not pragmatic (in a traditional sense) sentence organization
HVALA!