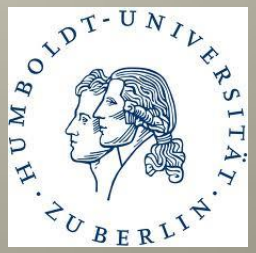
L2 complexity as syntactic modification in a developmental L2 German corpus

Nina Vyatkina
University of Kansas

Marc Reznicek
Humboldt University Berlin



AAAL 2013
March 17, 2013
Dallas, TX

Approach

- *study of L2 complexity*
  - the range of forms that surface in language production and the degree of sophistication of such forms (Ortega, 2003)
- *developmental profiling*
  - unique matrices of various linguistic forms characterizing specific interlanguage stages (Clahsen, 1985; Pienemann et al., 1988)
- *learner corpus analysis*
  - large electronic collections of learner texts with strict design criteria relating to their provenance (Granger, 2002)

The L2 corpus: KanDeL

- **Kan**sas **De**velopmental **L**earner corpus:
- *ab initio* learners of German at a large public US university
- L1 English
- curricular writing tasks
- mean text length: ca. 130 words
- collected every 3 to 5 weeks over 4 semesters
- => rich and dense, not a 'mega-corpus':
  - cohort 1: 66142 words, 16-40 essays per time point
  - cohort 2: 25336 words, 12 longitudinal participants

Completed studies

- Quasi-longitudinal data: the 1st KanDeL cohort
  - linear increase:
    - global complexity (sentence length, lexical variety)
    - range and variety of selected morphosyntactic forms
      - subordination
  - no significant increase of clause length
- Individual longitudinal data (2 learners):
  - linear increase in global complexity
  - diverging paths for specific complexity
    - learner 1: clausal; learner 2: phrasal
      - Vyatkina (2012, 2013, in press)

This study

- explore development of syntactic complexity at finer levels of granularity
- syntactic modification: elements not obligatory for the verb argument structure

(Hirschmann et al., in press)

- 2nd KanDeL cohort: longitudinal (12 learners)

Analytical constructs

2 main types of syntactic complexity:

(Biber et al., 2011; Byrnes et al., 2010; Halliday & Martin, 1993)

1. "dynamic" style:
   - verb-related resources, clausal elaboration
     - » oral registers, narrative texts, everyday private discourse
     - » lower levels of L2 proficiency
2. "synoptic" style:
   - noun-related resources, phrasal elaboration
     - » literate registers, expository texts, secondary public discourse
     - » higher levels of L2 proficiency

Research questions

- How does syntactic modification change in beginning learners' of German writing over time?

  -> How do frequencies of (1) verb modifiers and (2) noun modifiers change over time?

  -> Hypotheses:

  - (1) are more frequent than (2)
  - (1) decrease and (2) increase over time

  -> What other factors influence frequencies of syntactic modifiers in addition to time?

Multi-layer corpus annotation

- Following guidelines for the FALKO corpora (Reznicek et al., 2012):
  - parts-of-speech, lemma, syntactic dependencies
  - error corrections: Target Hypotheses (TH):

- text: *Nächste Woche, ich fahre mit das Flugzeug.*
- TH1: *Nächste Woche fahre ich mit dem Flugzeug.*
- TH2: *Nächste Woche fliege ich mit dem Flugzeug.*
- Engl. *Next week, I am going by plane.*

Target features

- text: *Nächste Woche, ich fahre mit das Flugzeug.*
- TH1: *Nächste Woche fahre ich mit dem Flugzeug.*
- TH2: *Nächste Woche fliege ich mit dem Flugzeug.*
- Engl. *Next week, I am going by plane.*

- automated syntactic dependency parsing based on TH1
- manually post-corrected
- frequencies of verb modifiers and noun modifiers tallied separately

I. Prepositional phrases (PPs)

- Multidimensional profiling of the PP as an "ambicategorical element" (Hilpert, 2010)
  - He **painted** a tree on the wall (VP modifier)
  - **The painting** on the wall is new (NP modifier)
- Factors:
  - participant (12)
  - time point (17)
  - genre (essay / letter/ summary)
  - writing conditions (timed / untimed)
  - participant gender
  - topic

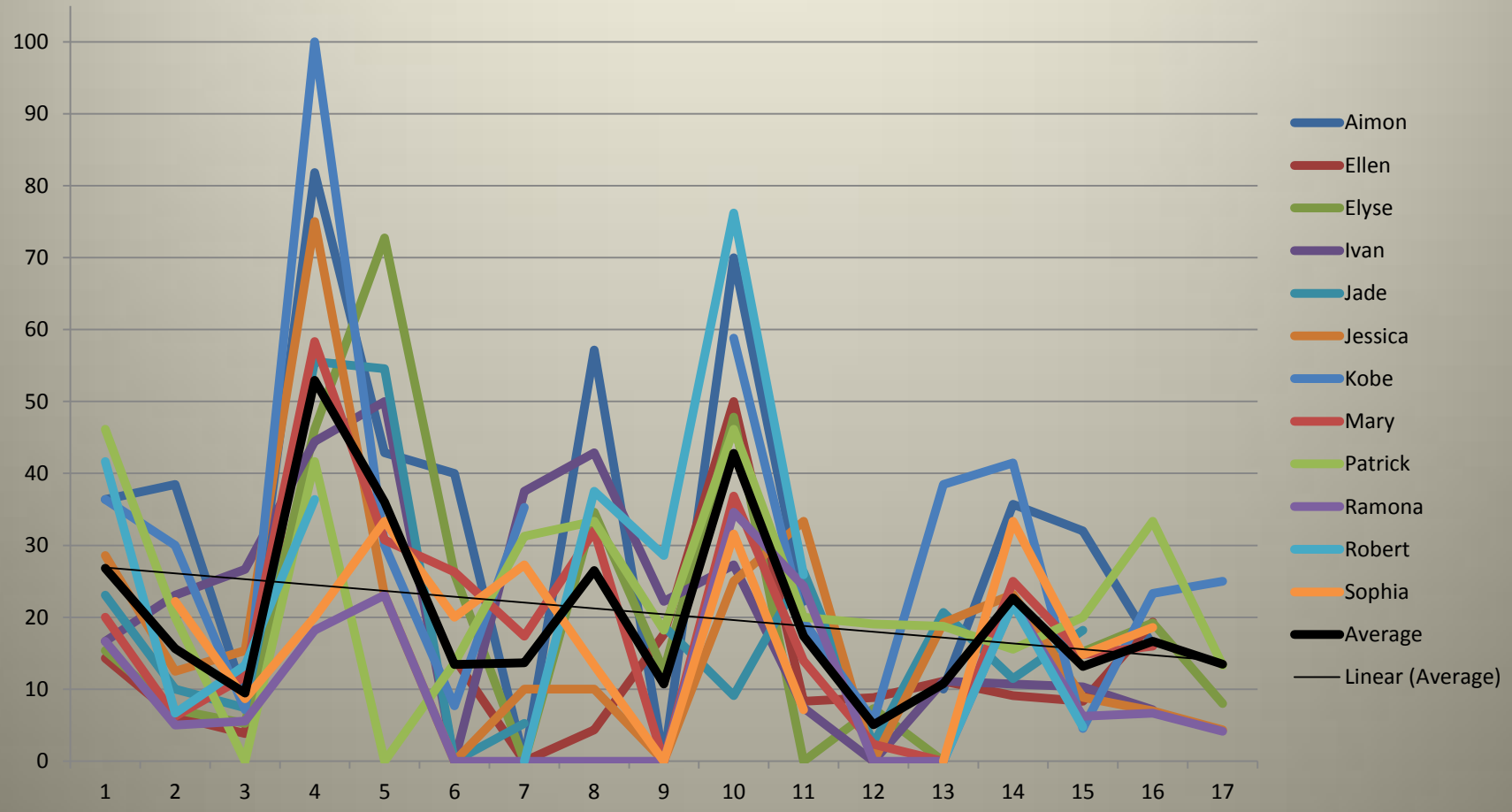
Multilevel Modeling (MLM)

- tracks both *group and individual change*
- permits *missing data and varying sample sizes*
- allows for the analysis of *multiple factors*
- deals with *both categorical and continuous* variables
- treats *time as a continuous* variable
- accounts for their *interactions* as well as *residual variation* and *random effects*.

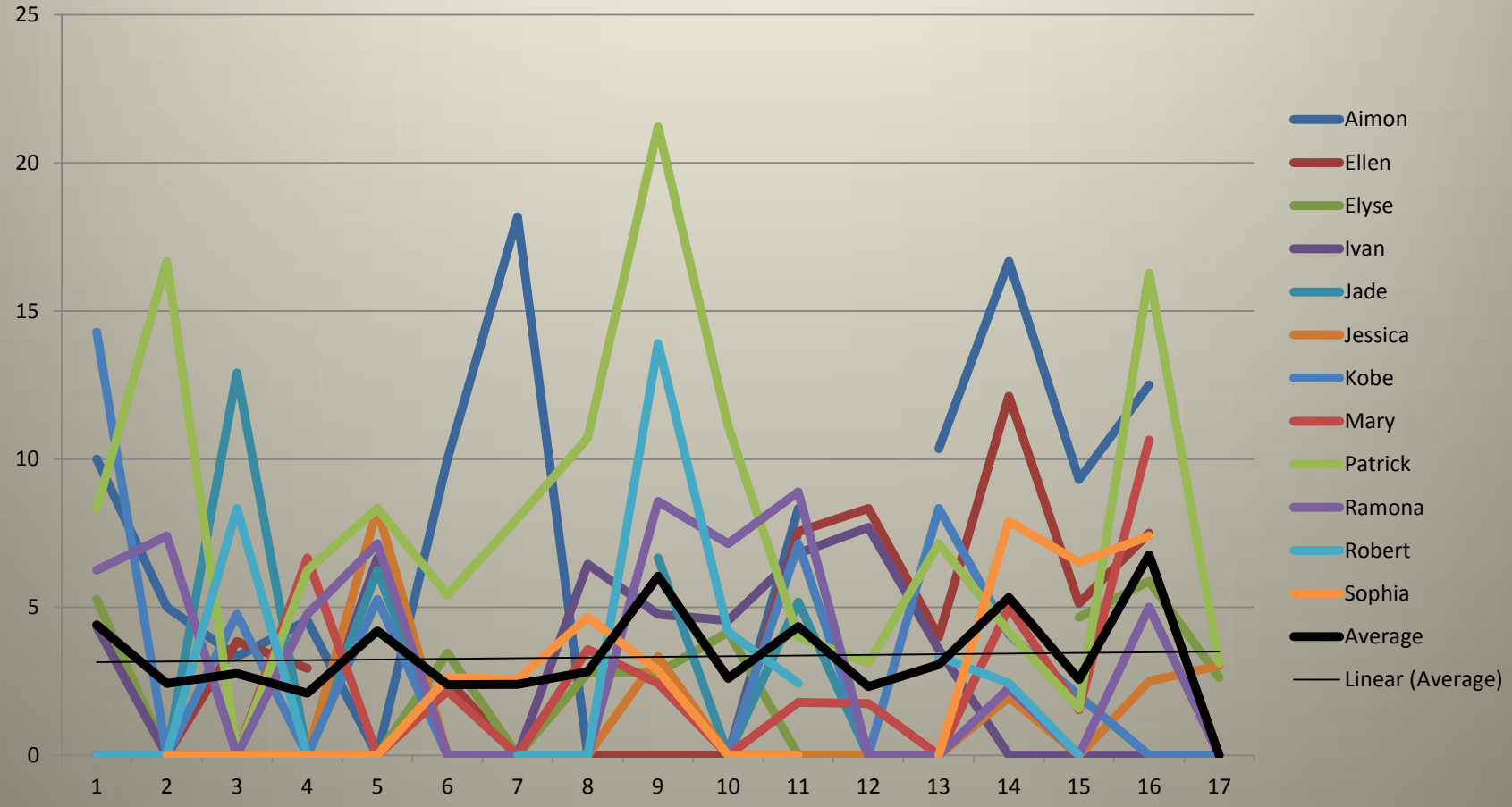
Results I: PP as VP or NP modifier

- No significant time effect
- PPs more frequently modify VPs than NPs
    1. in the timed condition (p=.015)
    2. in male students‘ writing (p=0.006)
    3. in summaries and letters in comparison with essays (p= 0.0008)
- A significant random effect for topic
- No random effect for individuals
- Overall, male students modify more than female students

Individual data: % PP per VP



Individual data: % PP per NP



PP results summary

- H1 confirmed:
  - beginning learners of German use PPs more frequently as VP modifiers than as NP modifiers in their writing (19.46 vs. 3.36 per phrase)
- H2: no significant interaction between the two PP modifying functions over time
  - the two functions should be considered separately along with other VP and NP modifiers
  - other contributing factors: genre, topic, time limit, gender

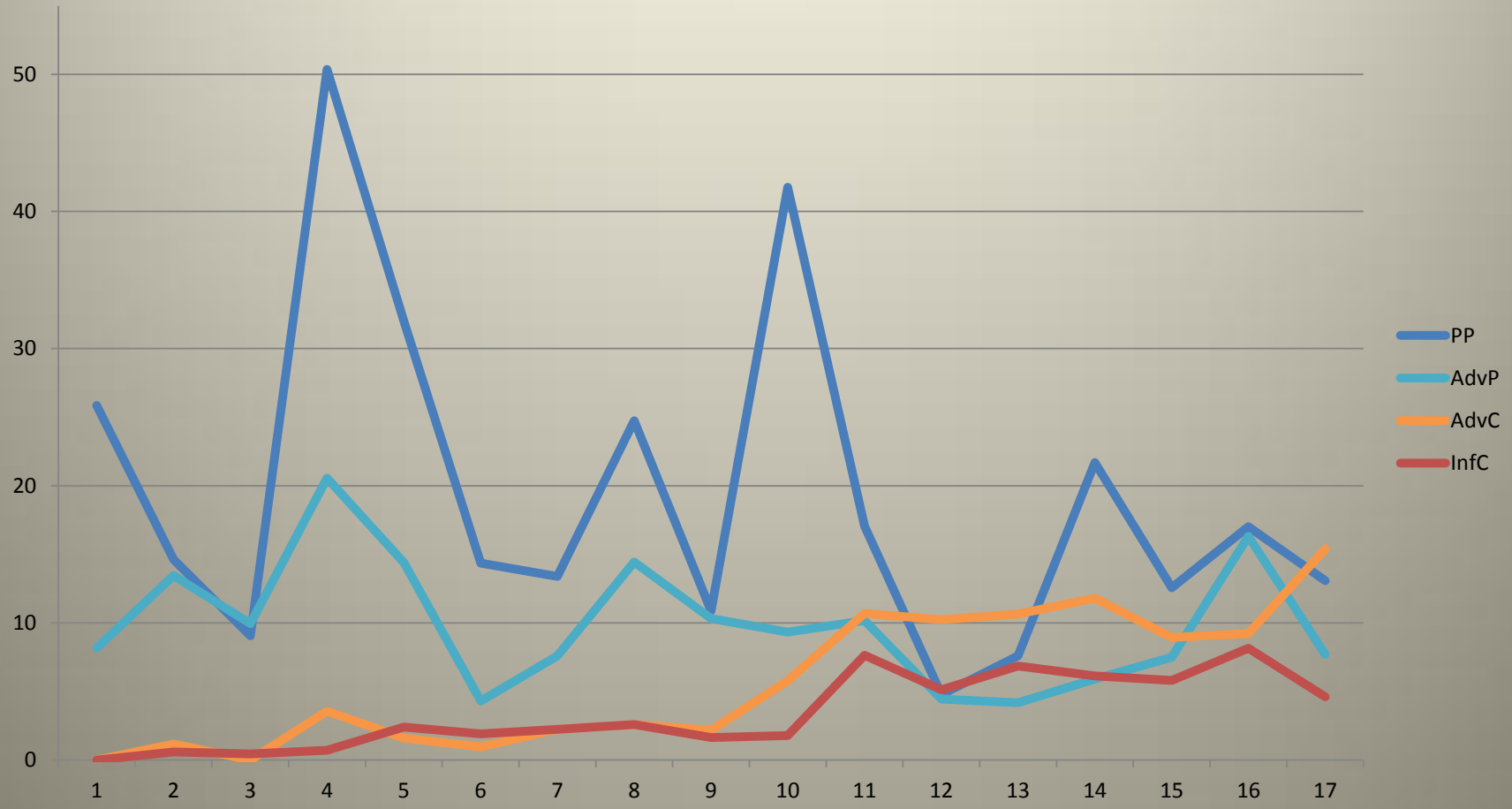
II. Other VP and NP modifiers

- VP modifiers:
  - adverbs (AdvP)
  - PP
  - adverbial clauses (AdvC)
  - infinitive clauses (InfC)
- NP modifiers:
  - premodifiers: attributive adjectives (AdjP)
  - postmodifiers:
    - nouns in the genitive case indicating possessor (NP)
    - PP
    - relative clauses (RelC)
- Exploratory analysis: group data; linear regression

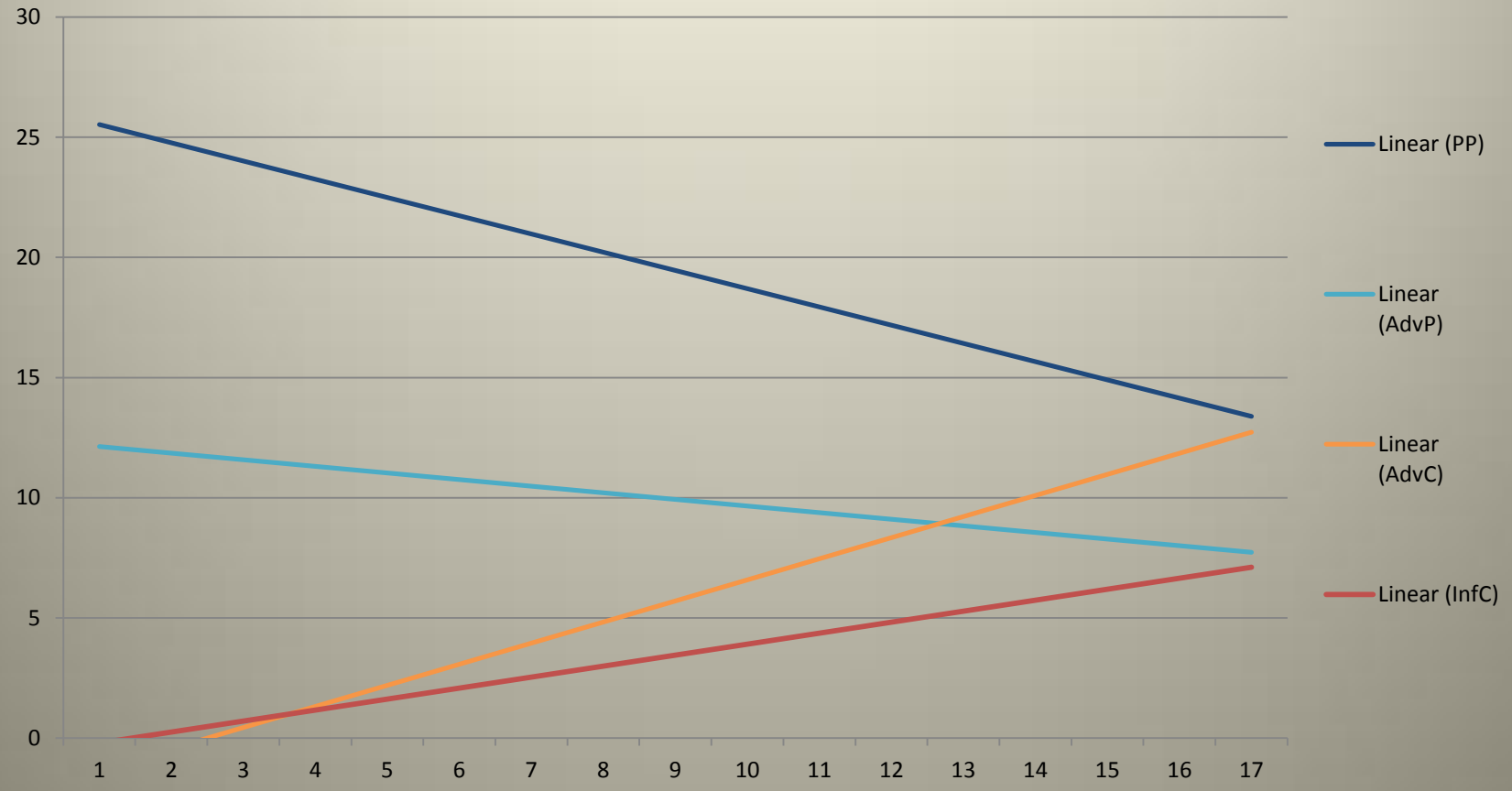
Results II: NP and VP modifiers

- Beginning learners of German use more VP modifiers than NP modifiers (9.6 vs. 4.7 per phrase)
- Linear regression results:
  - VP modifiers:
    - phrasal (AdvP and PP): insignificant decrease
    - clausal (AdvC and InfC): increase ($p<0.0001$)
  - NP modifiers:
    - prenominal (AdjP): increase ($p<0.0001$)
    - postnominal:
      - NP and RelC: increase ($p<0.01$)
      - PP: insignificant increase

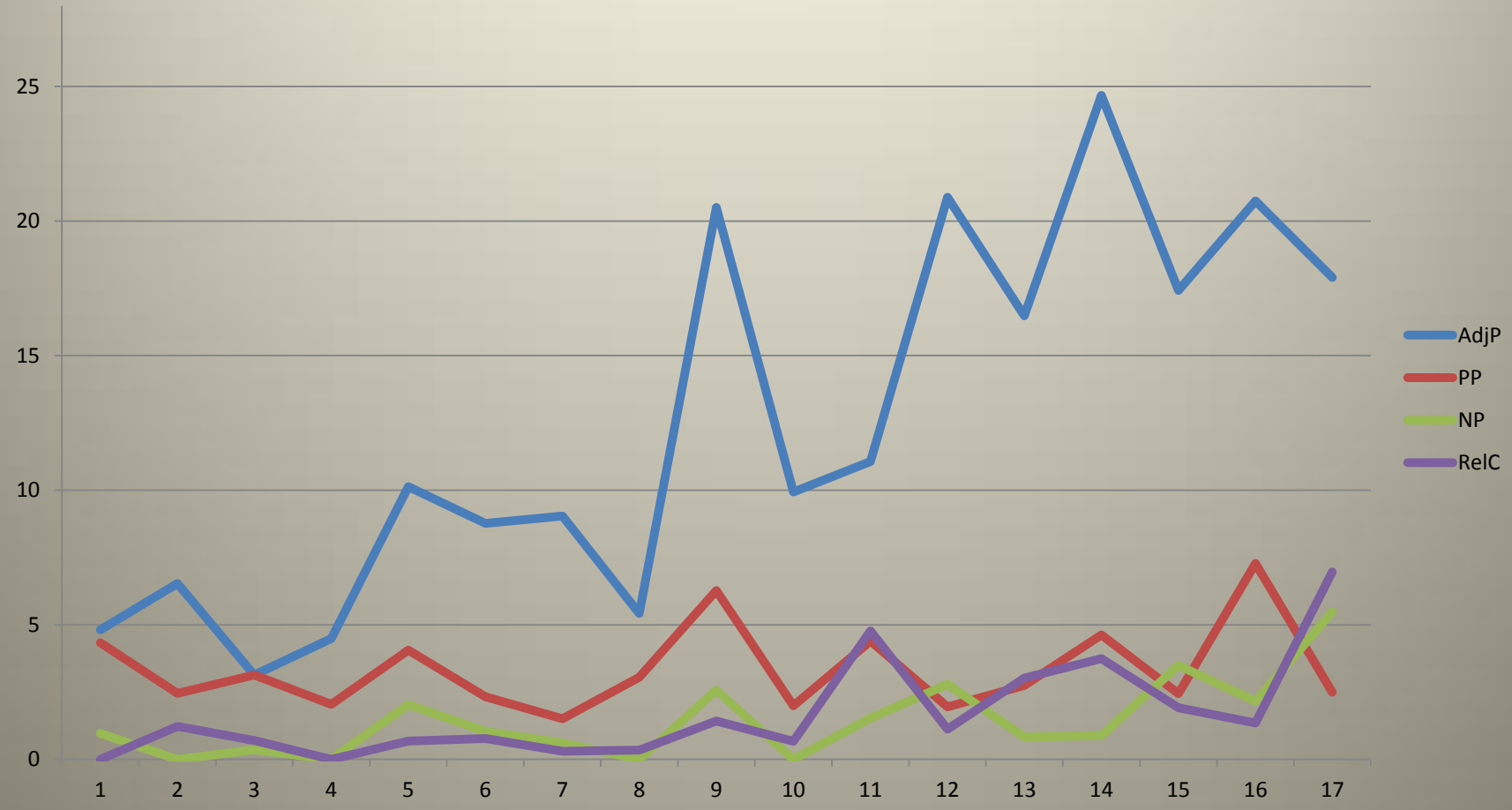
Modifiers per verb phrase



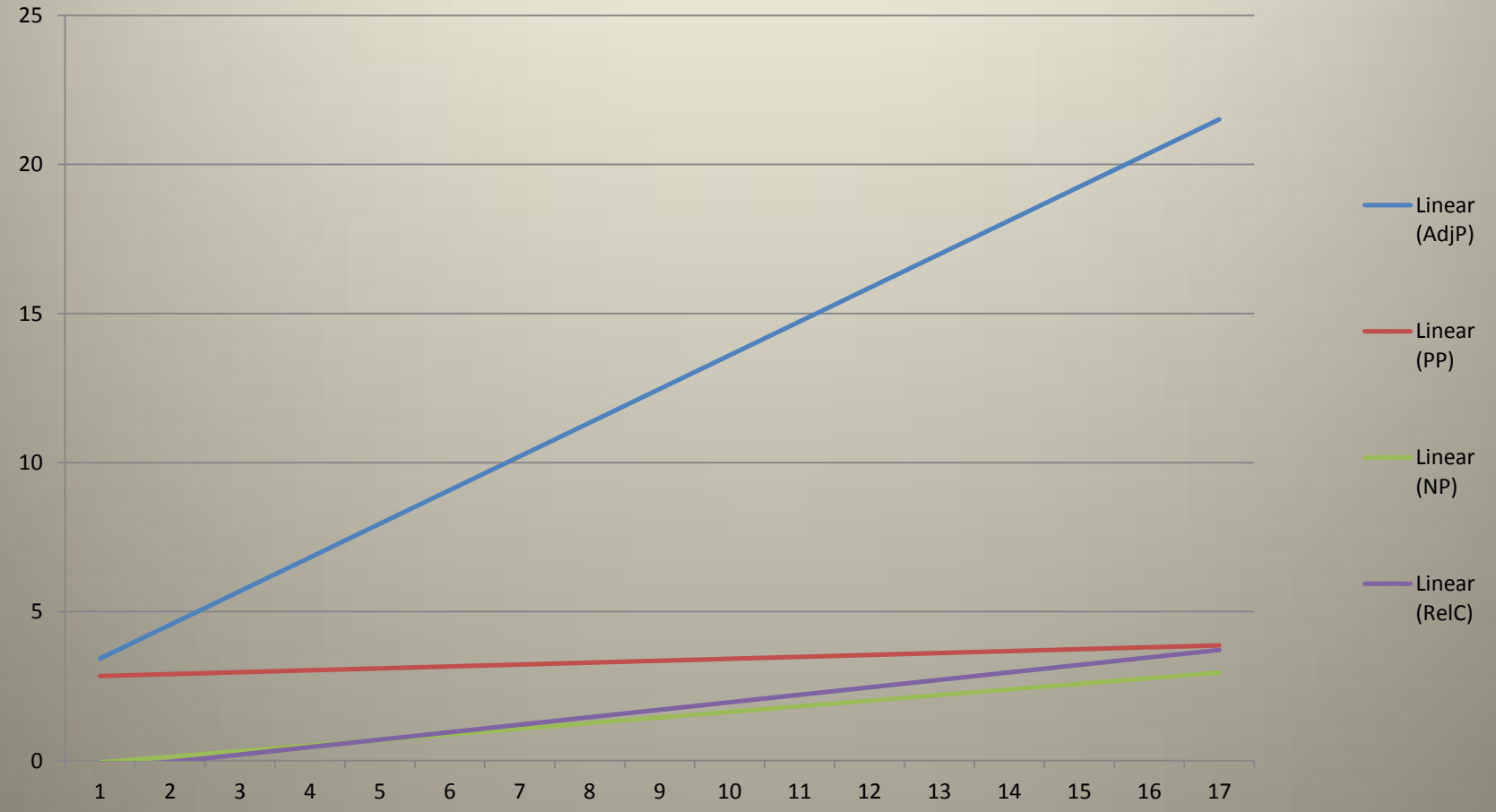
Verb modifiers: trendlines



Modifiers per noun phrase



Noun modifiers: trendlines



Summary

- Syntactic modification emerges at the incipient levels of L2 proficiency
    - Phrasal verb modifiers -> prenominal phrasal modifiers -> clausal verb modifiers & postnominal modifiers
- Complex interaction between syntactic type and function reflected in the developmental profiles

  -> Biber et al. (2011)
- The use of syntactic modifiers is strongly influenced by writing conditions and topics

  -> Golcher & Reznicek (2011)

Future research

- Fine-tune syntactic annotation
- MLM analysis of different syntactic modifiers
- Accuracy analysis
- Longer data collection period
- Qualitative analysis (gender effect?)
- Developmental analysis of lexico-grammatical constructions
- Non-linear methods

Acknowledgments

- Felix Golcher (Humboldt University Berlin) for statistical analysis
- Emily Hackmann and Michael Grünbaum (University of Kansas) for data annotation
- KU Institute for Digital Research in the Humanities, *Digital Humanities* Seed Grant
- *Language Learning* Small Grants Research Program