The semantics of conversion nouns and *-ing* nominalizations:
A quantitative and theoretical perspective

Rochelle Lieber & Ingo Plag

University of New Hampshire & Heinrich-Heine-Universität Düsseldorf

Introduction

Two major problems in derivational semantics

One form - many meanings
One affix may give rise to different meanings (e.g. EVENTIVE vs. REFERENTIAL)

One meaning - many forms
One concept may be expressed by more than one form
(affix competition, *-ing,* conversion*, -ion, -ment, -age*)

An overview of this paper

- The general problem concerns competition between *-ing* nominalizations and *conversion* nouns.

- There are many claims about their semantics in the literature, but little solid empirical evidence.

- We look at quantification (mass vs. count), eventuality (eventive vs. referential), aspectual class (accomplishment, achievement, state, …) in both types of nominalizations using a large data base (N=4282, COCA).

- We provide a quantitative analysis and then interpret the results in the light of extant morphological theories.

A sampling of claims in the literature

- Morphological form and count versus mass
  Consensus in the literature (e.g. Brinton 1995, 1998, Grimshaw 1990, Langacker 1991, etc.):
  -ing tends towards mass quantification and *conversion* towards count.

- Aspectual class of base verb and count versus mass
  Mourelatos (1978): state and activity verbs → mass quantified; accomplishment and achievement → count quantified

More claims

- Aspectual class of base verb and form of nominalization:
  - Brinton (1995): State verbs prefer *conversion*; activity verbs prefer *-ing* nominalizations.
  - Alexiadou (2001), Borer (2013): accomplishments and achievements disprefer *-ing* nominalizations (but see Iordachioaia & Werner 2019)
- Morphological form and eventivity:
  - Grimshaw (1990) and others: *-ing* nominalizations are interpreted as complex events; *conversion* nouns cannot be eventive.

Roughly, the expected association:

- *Conversion* ~ count quantification ~ referential reading

- *-ing* nominalization ~ mass quantification ~ eventive reading

Andreou & Lieber (2020)

- There is no necessary correlation between morphological form, eventivity, quantification, and aspectual class.

- Different morphological forms do not show strict preferences for particular aspectual classes.

- Any morphological form can in principle express any combination of quantification, eventivity, and aspectual class of base verb.

What Andreou & Lieber 2020 do not do:

- Provide a quantitative analysis to see whether there are any tendencies towards the associations predicted in prior literature.

- Investigate the role of the individual verb.

- Consider the data from two directions:

Semasiological: what does a form mean?

Onomasiological: which form do I choose to express an intended meaning

(A&L concentrate on the semasiological side.)

Methodology

- Subset of Andreou & Lieber's (2020) database from COCA/BNC of *-ing* and *conversion* nominalizations of 106 verbs:
  57,000 → 4282 singular nouns

- Hand-coded by both authors for

  - eventivity

  - quantification

  - aspectual class

- Parallel independent coding for eventivity

Positive evidence for count versus mass

Count

- A determiner *a/an* precedes the noun
- ordinals or cardinals
- modifiers like *final, single, last*
- quantifiers like *each, every*

Mass

- Singular nominal with no determiner
- quantifiers like *much, some*

Examples

-*ing* count (evidence: indefinite article)

BNC 1992: …tiny phosphorescent sparks around its hands, small ripples in the stone beneath its feet, a gentle breeze around its head, a sudden dampness and DRIPPING of water from the stones of the walls around it.

conversion mass (evidence: singular form, no determiner)

Independent School 2006: Overall, then, the single greatest obstacle to implementing curricular CHANGE and, over time, establishing a culture that values continuous reflection and improvement in a school, is the general predisposition of educators to resist change itself.

Positive evidence for eventive versus referential reading

Eventive (a few examples – there are others):

- modification by temporal adjectives like frequent or repeated

- tensed verb in context that suggests eventive reading

- paraphrases like 'the activity of ...' or 'an instance of'

Referential (a few examples – there are others):

- coreference with a concrete noun present in context

- paraphrases like 'a kind of ...', 'an entity or product', 'the thing or stuff Ved', etc.

Examples

Eventive reading: manner adjective or adverb suggests eventivity

- *Shawshank Redemption 1994*: Norton keeps his eyes on Andy, looking for a wrong glance or nervous BLINK.

Referential reading: coreference with a clearly concrete noun in syntactic context

- *Mag. Inc. 1995*: "Today," says Holman, "employees in the cleaning department, for example, know not to work on the easiest-to-clean CASTING or on the one that happens to be on the top of the pile."

Aspectual class/features

aspectual class	dynamic	durative	implied endpoint	example
state	-	+	-	*know, love*
activity	+	+	-	*push, float*
accomplishment	+	+	+	*cook, cover*
achievement	+	-	+	*arrive, find*
semelfactive	+	-	-	*blink, knock*

What we eliminated

- Quantification: Cases where there was no positive evidence for quantification

- Eventivity: Cases where there was no positive evidence for eventivity, or where eventivity was unclear or ambiguous, or where raters disagreed on eventivity

- Aspectual class: Cases where verbal bases could be interpreted as more than one aspectual class depending on context

Datasets for each sub-study

- Quantification: N=2379 (only tokens with unambiguous evidence for mass or count reading)

- Eventivity: N=3286 (only unanimous ratings and unambiguous tokens)
  78% agreement (4 values, Cohen's Kappa for 2 Raters: 0.584, z=48.5, p=0). Contradictory ratings: 8%

- Aspectual class/feature: N=3405 (only unambiguous verbs)

- Pooled data (all semantic dimensions): Intersection of 1-3, N=1421 (150 types based on 84 verbs)

Results
Univariate analyses

Relation between two variables at a time

Quantification
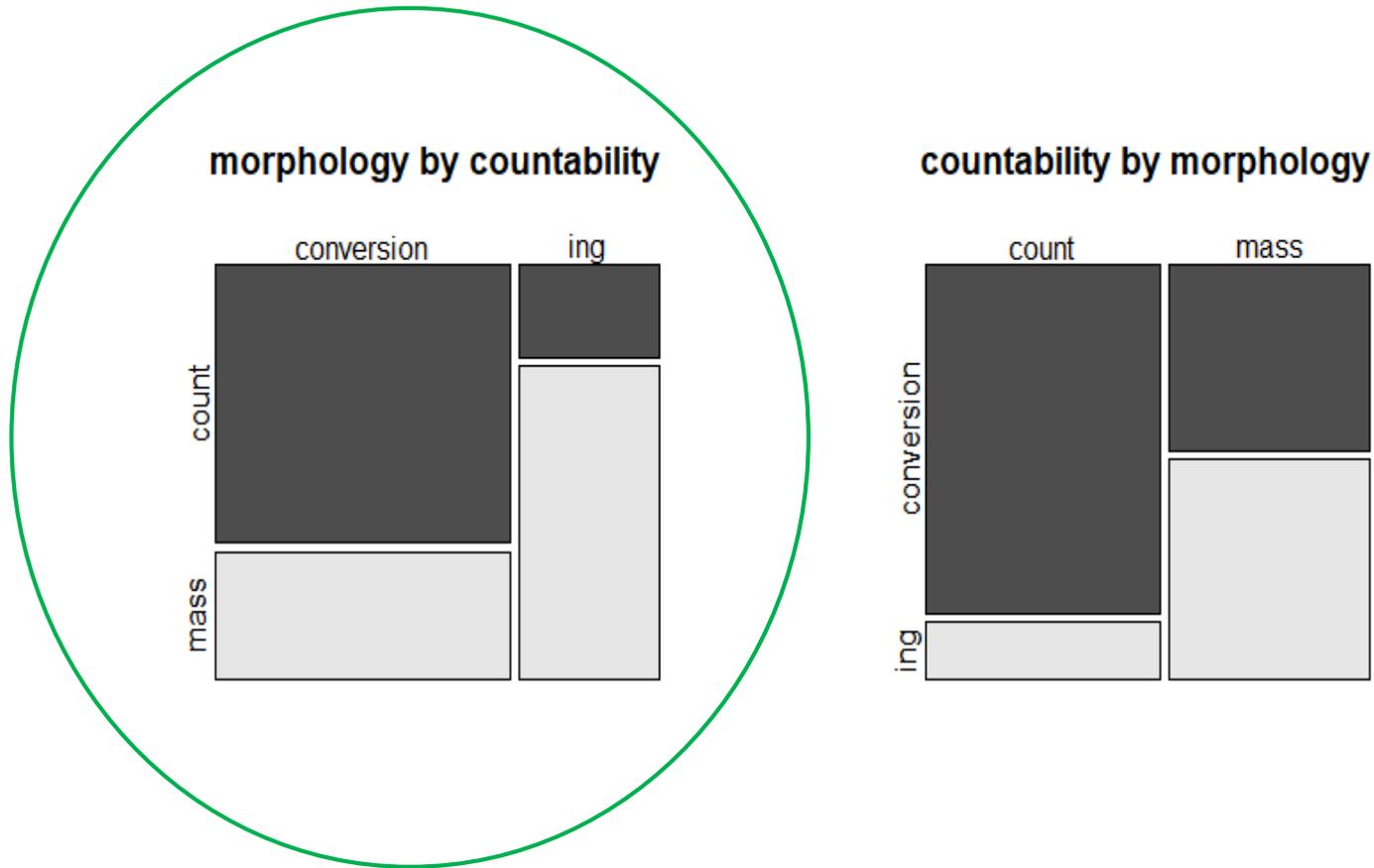
Figure 1: Morphology and quantification, N=2379
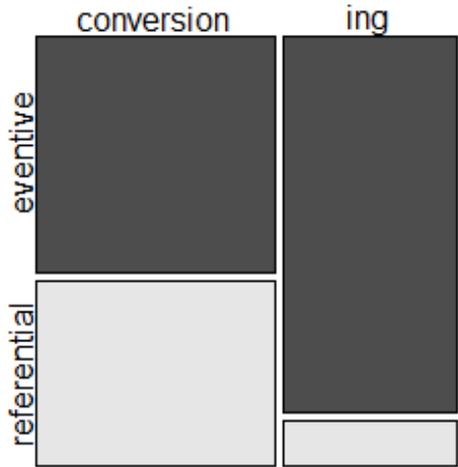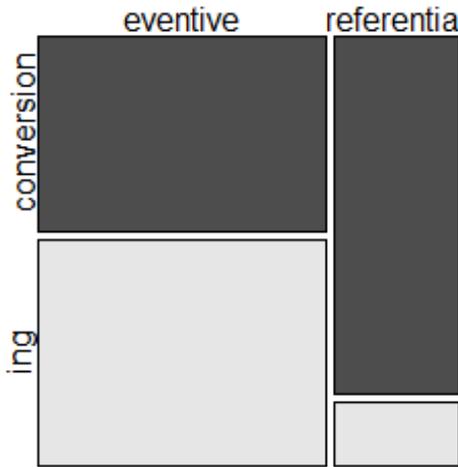
Eventuality

Figure 2: Morphology and eventivity, N=3286
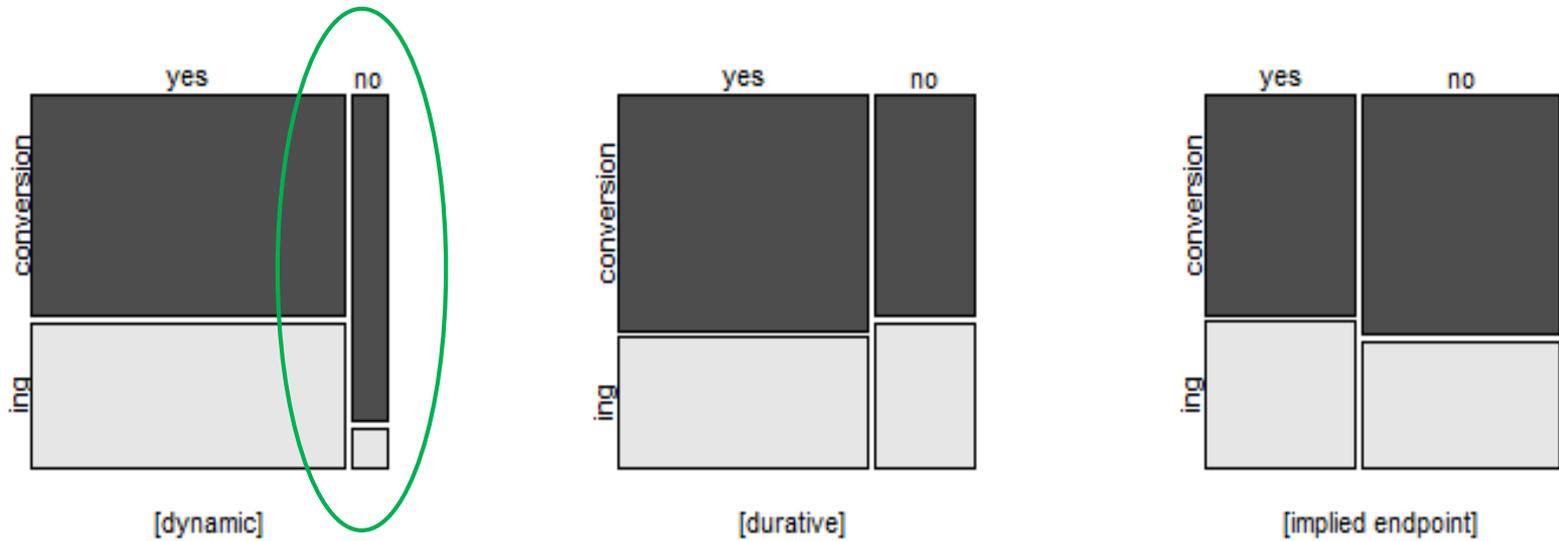
Aspectual features: Morphology

Figure 6: Aspectual features by morphology (N=3405)
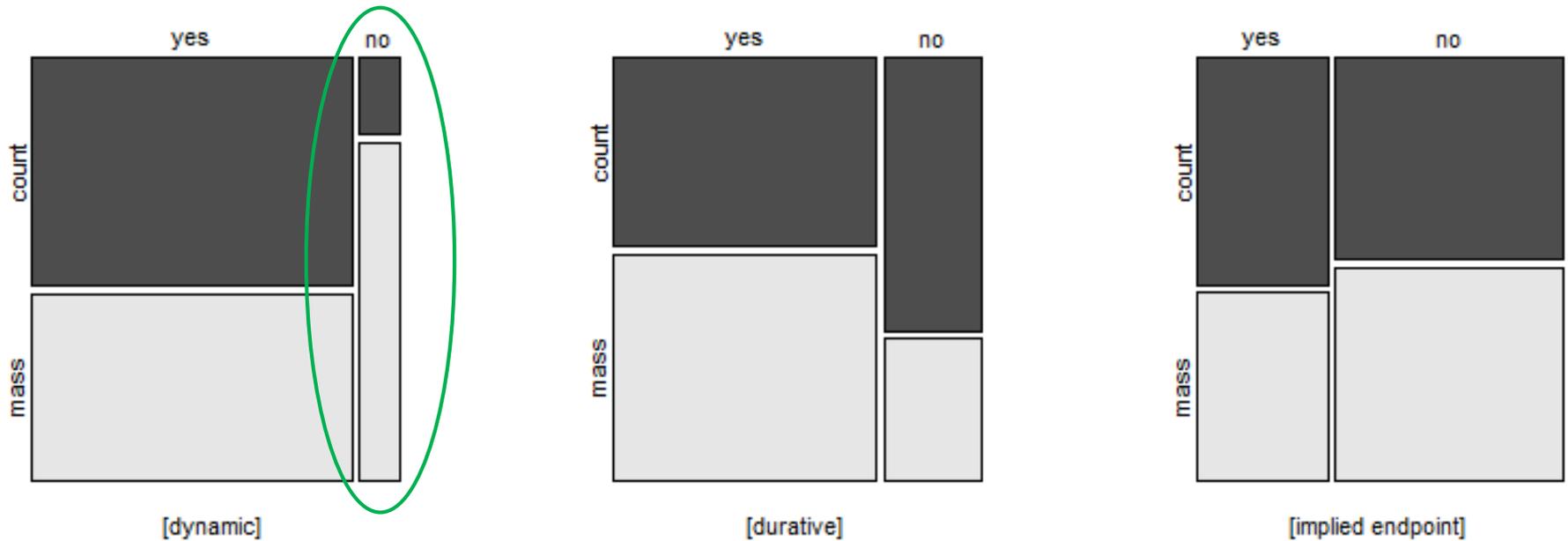
Aspectual features: Quantification

Figure 7: Aspectual features by quantification ('yes' indicates a positive value for the feature, 'no' a negative value for the feature, N=1934)
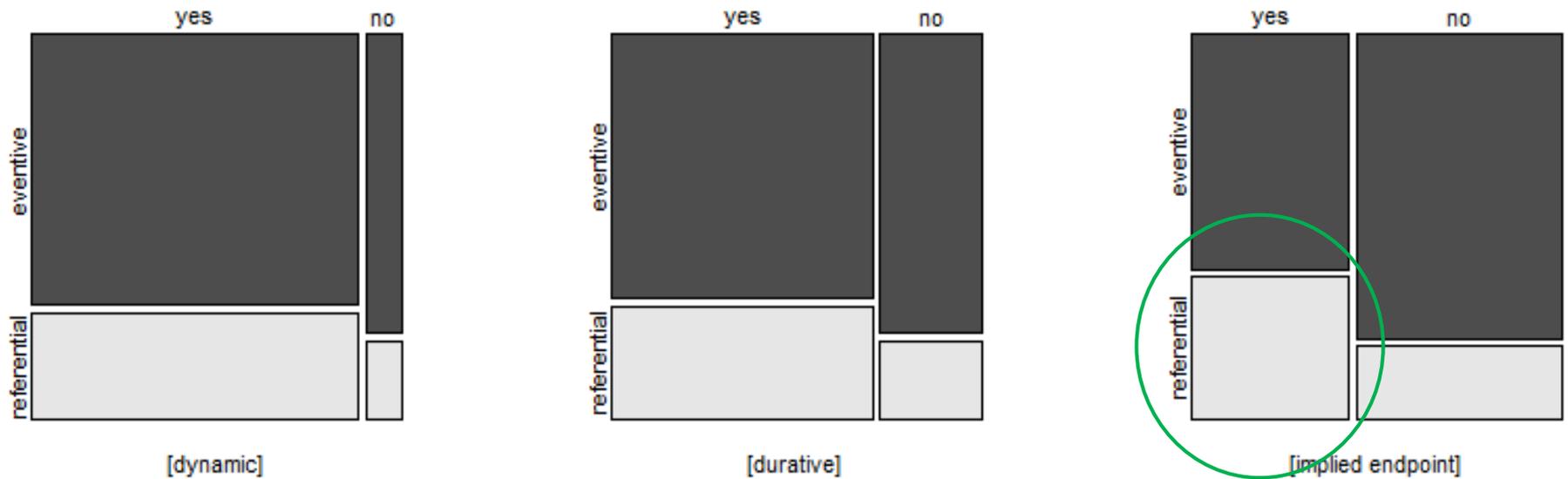
Aspectual features: Eventivity

Figure 8: Aspectual features by eventivity (N = 2568).

Variation within and across base verbs and their nominalizations

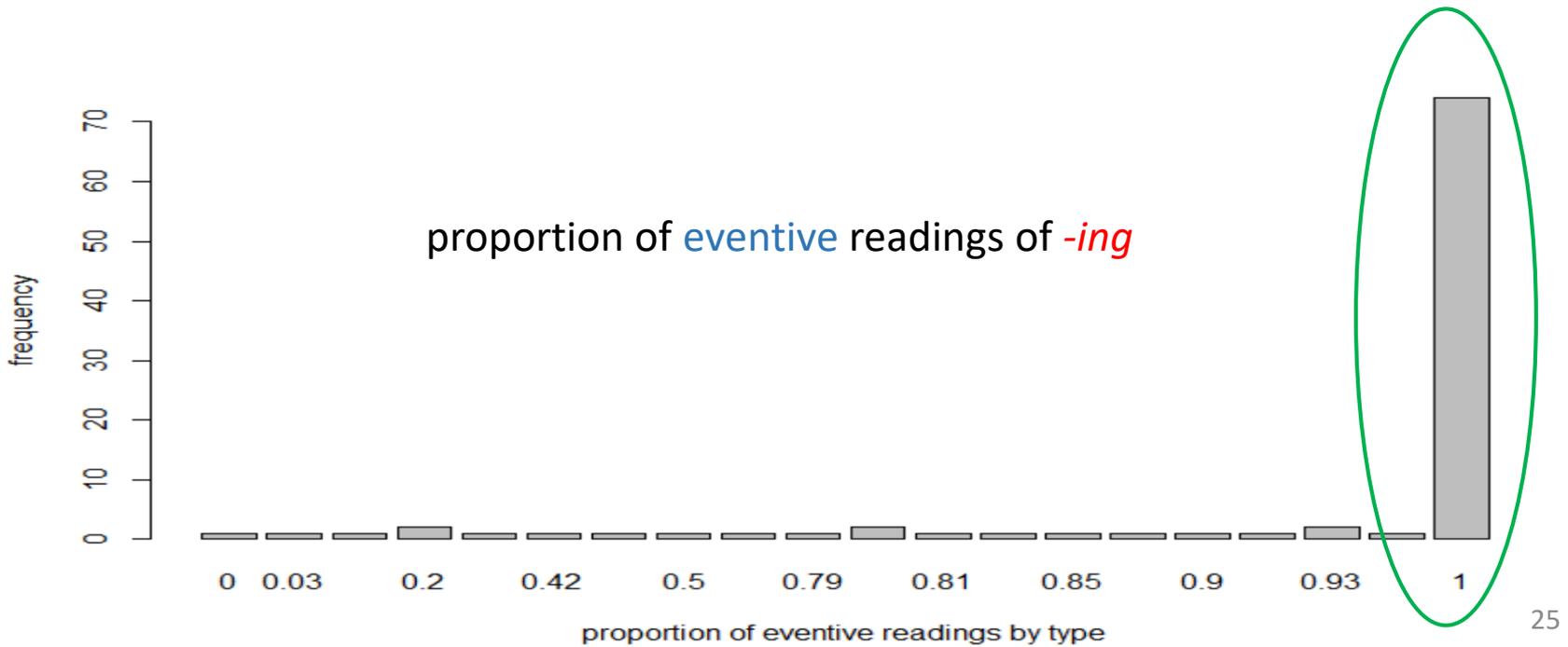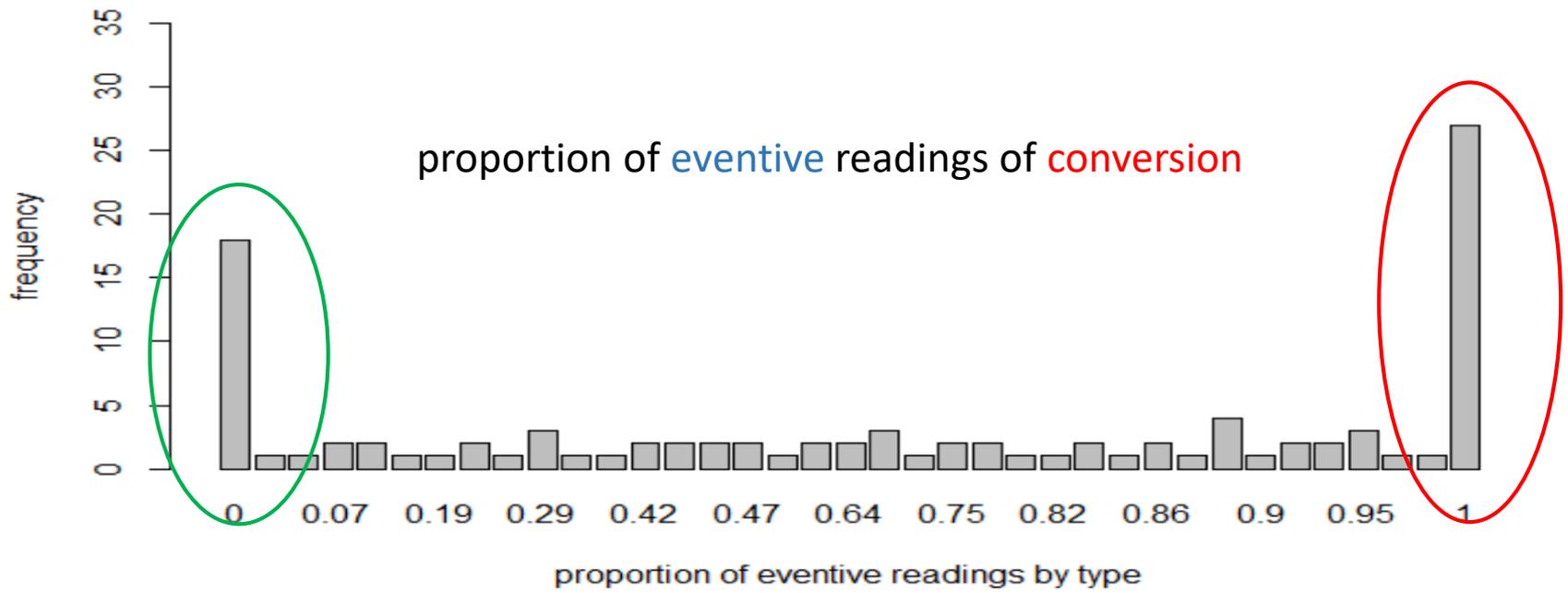Table 2: Token counts of nominalizations for three base verbs

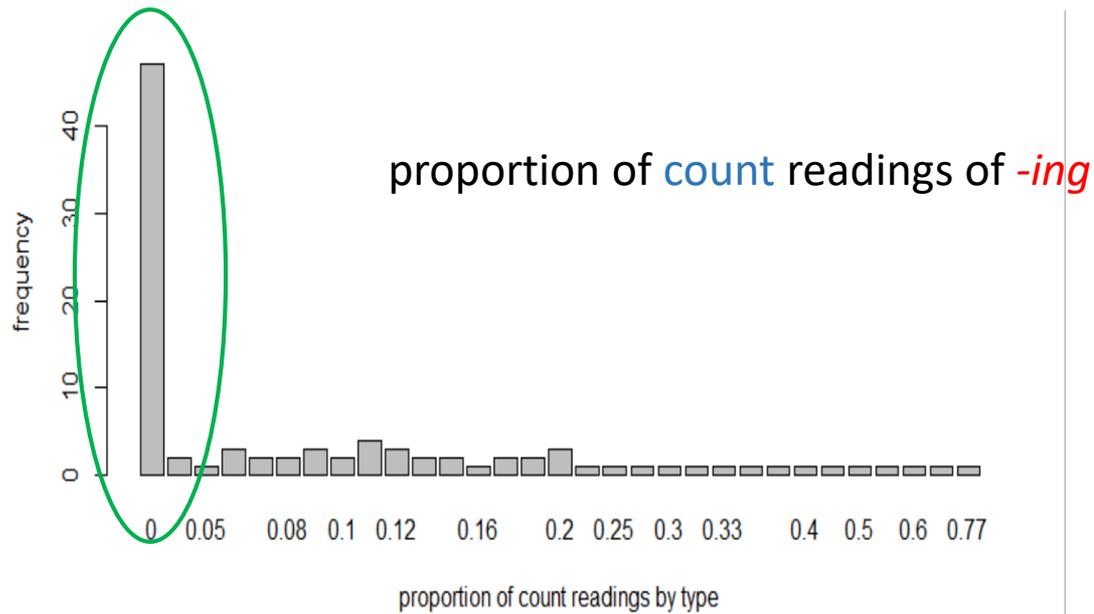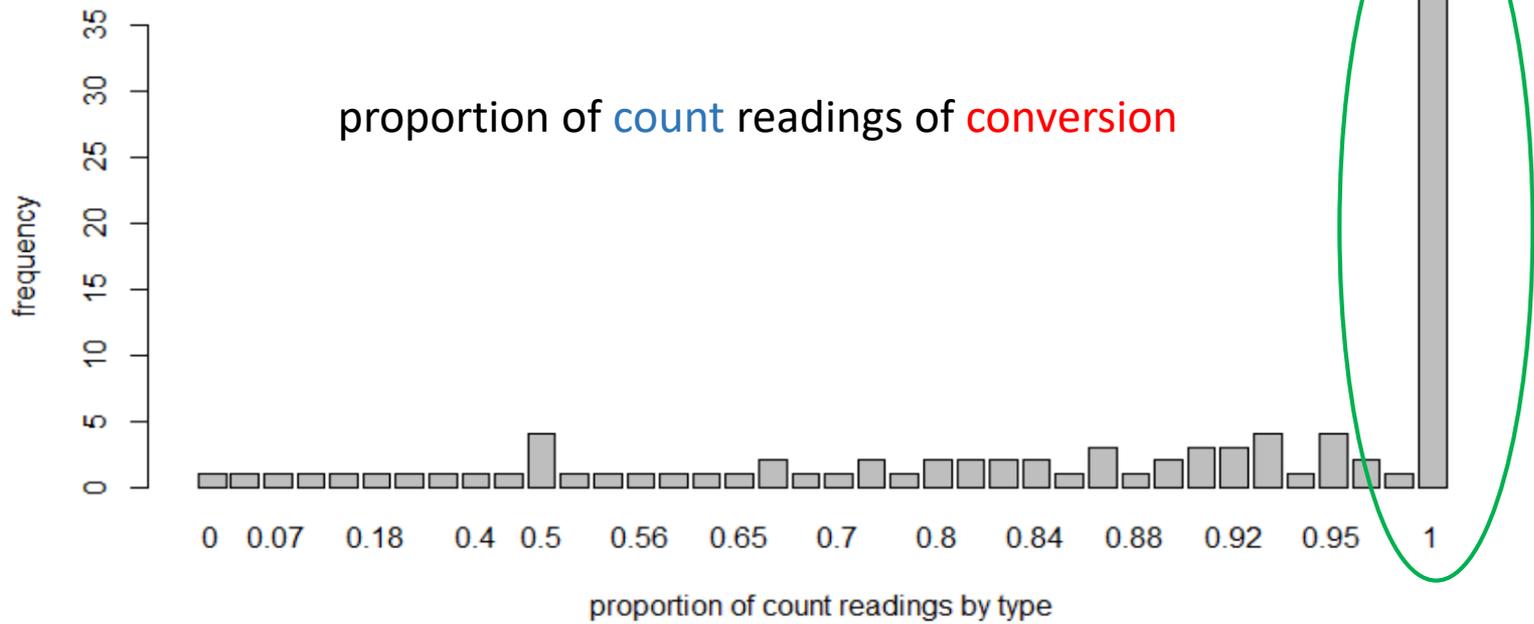|	CUT		DANCE		DESIGN	
	cut	*cutting*	*dance*	*dancing*	*design*	*designing*
eventive						
referential						
proportion eventive readings						

Variation within and across base verbs and their nominalizations

Table 2: Token counts of nominalizations for three base verbs

	CUT		DANCE		DESIGN	
	cut	*cutting*	*dance*	*dancing*	*design*	*designing*
eventive	3	5	5	11	0	9
referential	10	7	4	2	23	0
proportion eventive readings	*0.23*	*0.42*	*0.55*	*0.85*	*0*	*1*

How are these proportions distributed across all verbs?

proportion of eventive readings of conversion

proportion of eventive readings of *-ing*

proportion of count readings of conversion

proportion of count readings of -ing

Results
Multivariate analyses

The effect of many predictor variables
at the same time
on an outcome

Conditional inference trees

- Predict an outcome on the basis of all other variables ('predictors')

Three models

- MORPHOLOGY ~ QUANTIFICATION * EVENTIVITY * [± DYNAMIC] * [± DURATIVE] * [± ENDPOINT]

- QUANTIFICATION ~ MORPHOLOGY * EVENTIVTIY * [± DYNAMIC] * [± DURATIVE] * [± ENDPOINT]

- EVENTIVITY ~ MORPHOLOGY * QUANTIFICATION * [± DYNAMIC] * [± DURATIVE] * [± ENDPOINT]

- Partition the data into subsets

- Subsets
  - differ significantly in their distribution of the outcome variable
  - have particular constellations of the values of the predictor variables

- Advantage: complex interactions

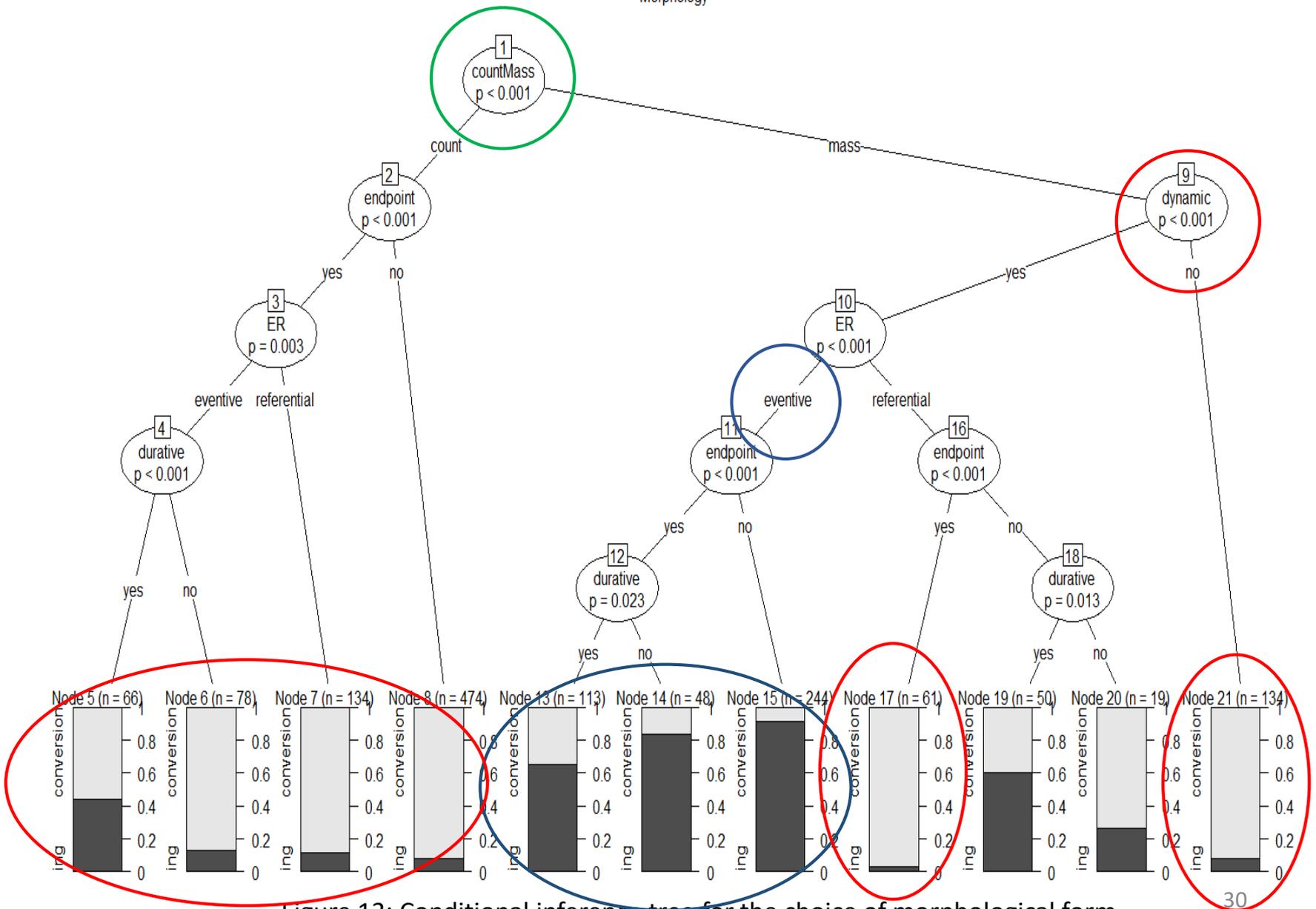- Disadvantage: necessarily excludes VERB as predictor

Figure 13: Conditional inference tree for the choice of morphological form

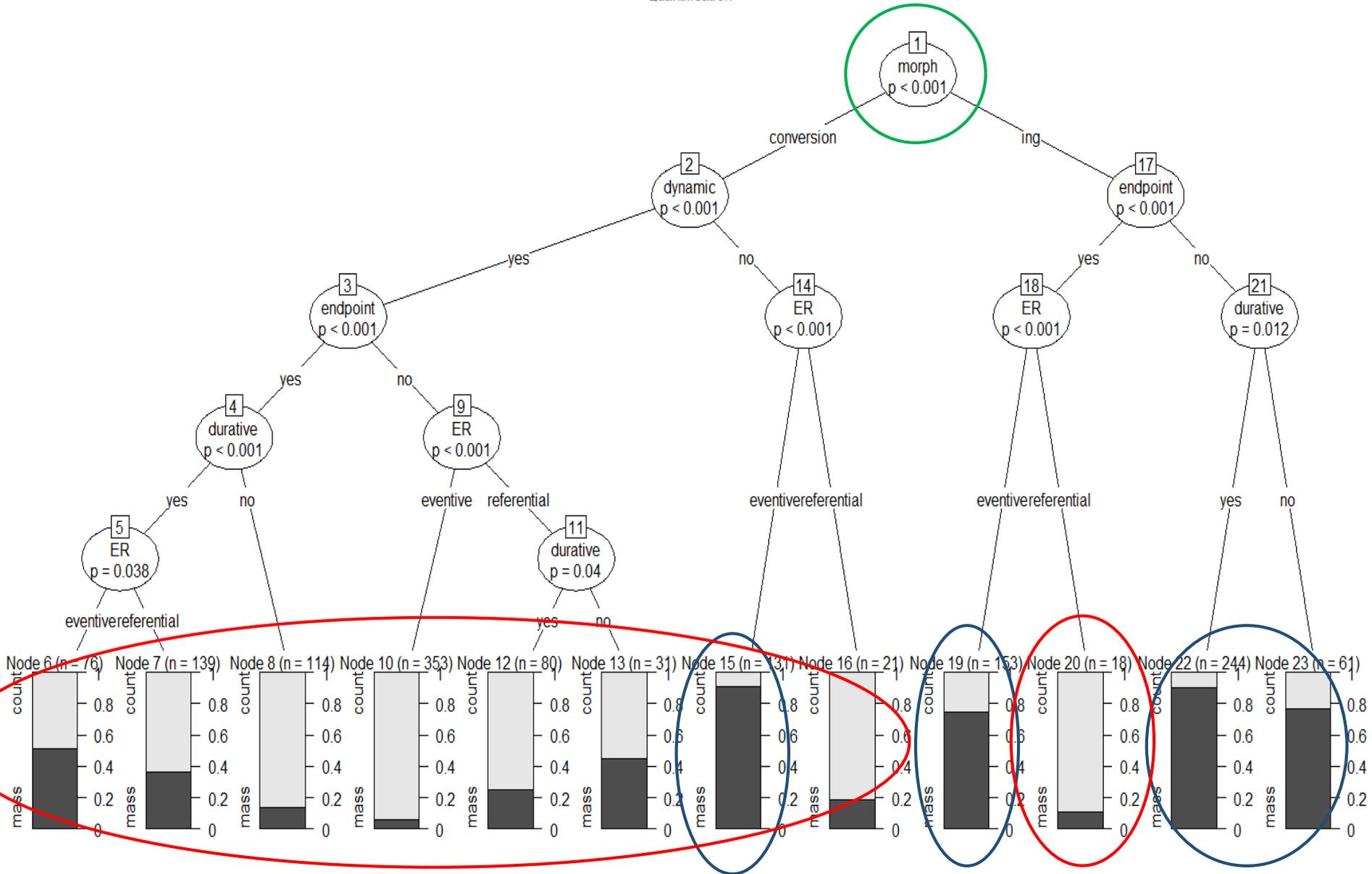Figure 13: Conditional inference tree for the choice of morphological form

Figure 14: Conditional inference tree for the choice of count vs. mass reading

Figure 15: Conditional inference tree for the choice of eventive vs. referential reading
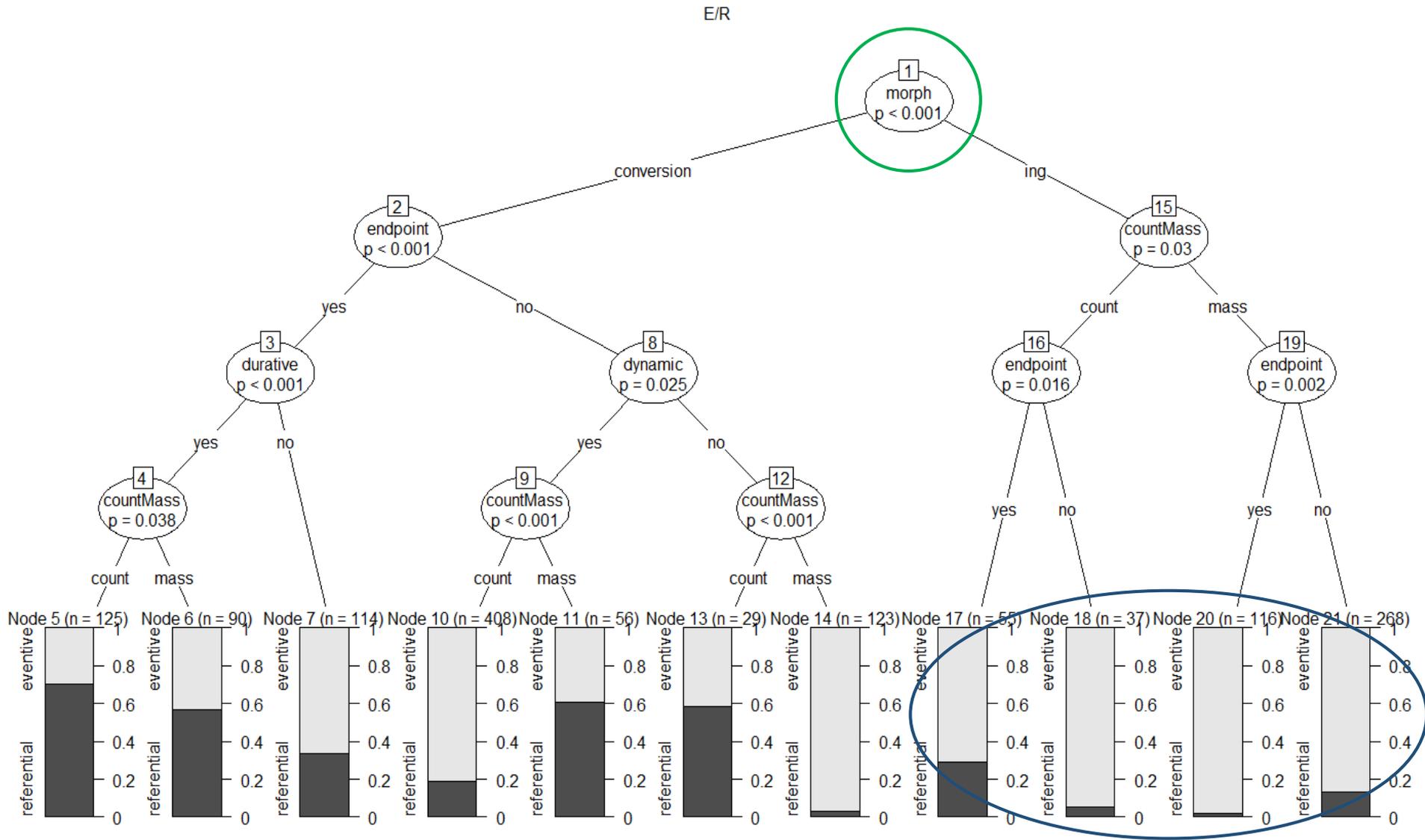
Figure 15: Conditional inference tree for the choice of eventive vs. referential reading

The role of the base verb:
Linear mixed effects models

- Inclusion of VERB as predictor ('random effect')
- No interactions
- Strength of idividual predictors

model	predicted outcome	significant predictors	explained variance *fixed effects*	explained variance *RE of VERB*
1	MORPHOLOGY	QUANTIFICATION EVENTIVITY [DYNAMIC]		
2	QUANTIFICATION	MORPHOLOGY [DYNAMIC]		
3	EVENTIVTIY	MORPHOLOGY [IMPLIED ENDPOINT]		

The role of the base verb:
Linear mixed effects models

- Inclusion of VERB as predictor ('random effect')
- No interactions
- Strength of idividual predictors

model	predicted outcome	significant predictors	explained variance *fixed effects*	explained variance *RE of VERB*
1	MORPHOLOGY	QUANTIFICATION EVENTIVITY [DYNAMIC]	0.47	0.26
2	QUANTIFICATION	MORPHOLOGY [DYNAMIC]		
3	EVENTIVTIY	MORPHOLOGY [IMPLIED ENDPOINT]		

The role of the base verb:
Linear mixed effects models

- Inclusion of VERB as predictor ('random effect')
- No interactions
- Strength of idividual predictors

model	predicted outcome	significant predictors	explained variance *fixed effects*	explained variance *RE of VERB*
1	MORPHOLOGY	QUANTIFICATION EVENTIVITY [DYNAMIC]	0.47	0.26
2	QUANTIFICATION	MORPHOLOGY [DYNAMIC]	0.33	0.34
3	EVENTIVTIY	MORPHOLOGY [IMPLIED ENDPOINT]		

The role of the base verb:
Linear mixed effects models

- Inclusion of VERB as predictor ('random effect')
- No interactions
- Strength of idividual predictors

model	predicted outcome	significant predictors	explained variance *fixed effects*	explained variance *RE of VERB*
1	MORPHOLOGY	QUANTIFICATION EVENTIVITY [DYNAMIC]	0.47	0.26
2	QUANTIFICATION	MORPHOLOGY [DYNAMIC]	0.33	0.34
3	EVENTIVTIY	MORPHOLOGY [IMPLIED ENDPOINT]	0.18	0.51

Summary

- Some tendencies in the expected directions for quantification and eventivity

- Some tendencies in the opposite direction (e.g. 56% eventive readings for conversion)

- A lot of variation, in both directions (form to meaning, meaning to form)

- Weak effect of aspectual categories (apart from [dynamic])

- Complex constellations of particular properties may go together with certain preferences in interpretation, or with preferences of morphological form.

- The individual base verb plays a very important role (sometimes the most important role) in the interpretation of a given nominalization form, or in the choice of the morphological form.

Bottom line

- Predictions based on a single feature are not satisfactory.

- Categorical claims, as found in the literature, are clearly wrong.

What does this mean for morphological theory?

Syntactic approaches

- Syntactic approaches like Distributed Morphology (DM) predict that ideally the association between morphological form, type of quantification, and eventivity ought to be categorical.

- Semantic nuances like count versus mass or eventive versus referential are encoded by adding functional projections or positioning nominalizers higher or lower in the syntactic structure

- Each combination of semantic nuances would require a different structural configuration with different functional categories and placements of nominalizers.

- The fewer the possible combinations, the neater the analysis.

Lexical Semantic approaches

- A purely lexical semantic approach like Lieber's Lexical Semantic Framework (LSF) predicts that ideally the association between morphological form, type of quantification and eventivity ought to be completely free.

- Different semantic interpretations are the result of underlying semantic underspecification and resolution of underspecification in context.

- If lexical semantic representations are highly underspecified and underspecification is resolved by context, then there should not be statistical preferences towards one combination of features or another.

Other possibilities

- Analogical modeling (e.g. A::M): well-suited for the task
  No study available with semantic features

- Distributional semantics (Lapesa et al. 2018, Wauquier 2020): flexible, well-suited to deal with polysemy
  No study available for English *-ing* and conversion nouns

- Watch this space ☺

Thanks for listening!

We gratefully acknowledge financial support by Deutsche Forschungsgemeinschaft (Grants SFB 991/2-C08 and PL 151/11-1 'Semantics of derivational morphology' to Ingo Plag)
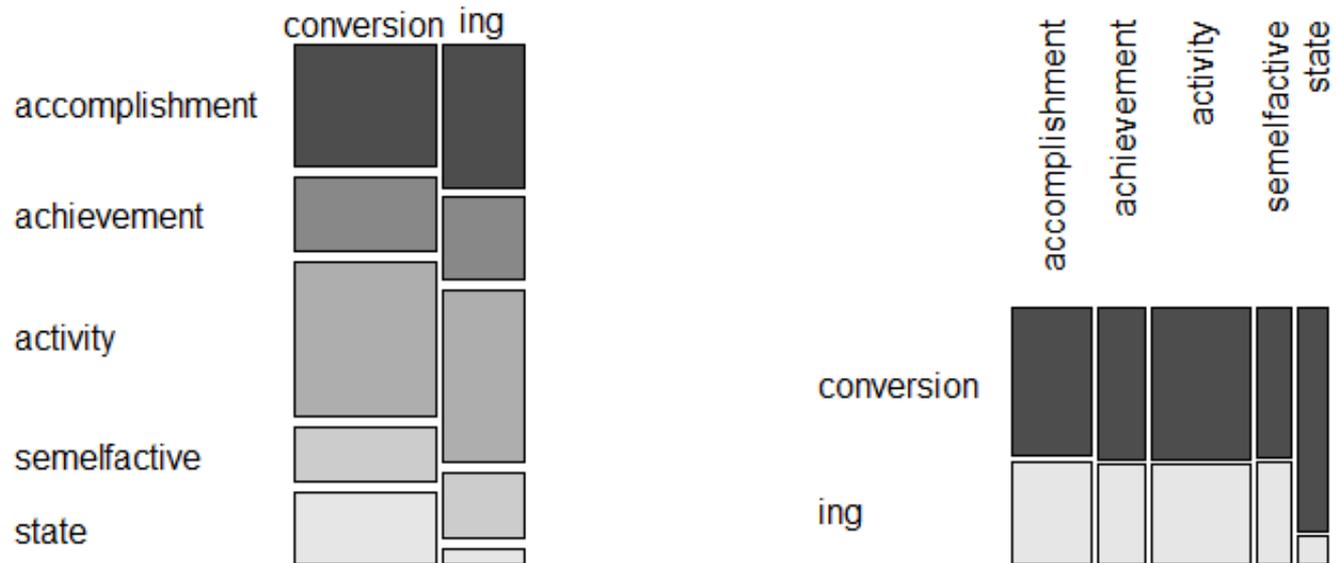
Aspectual classes

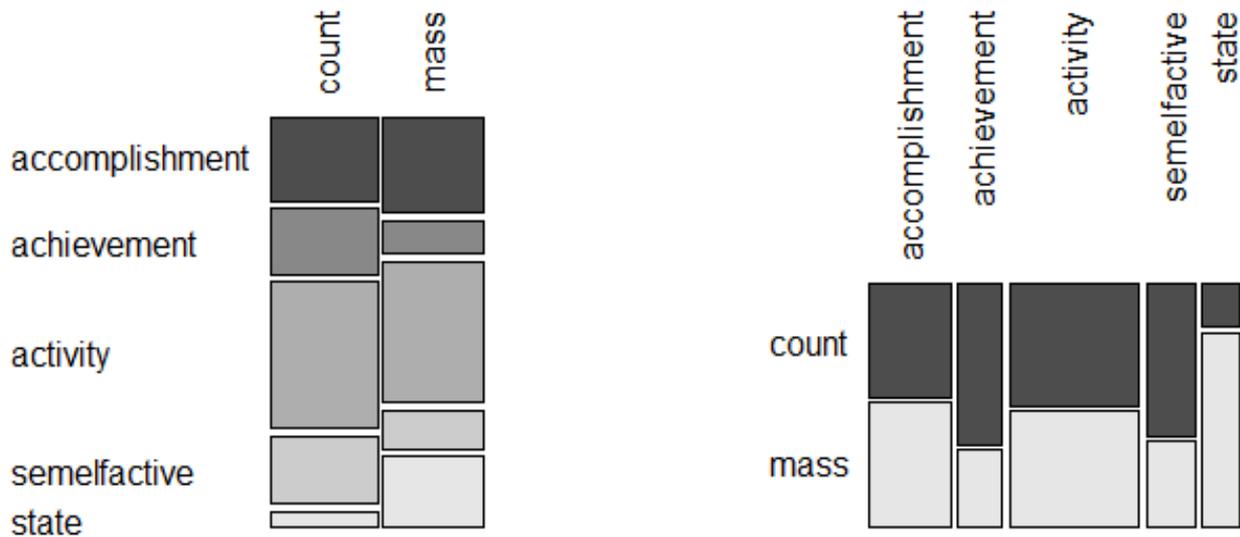Figure 3: Morphology by aspectual class (N=3405)

Aspectual classes

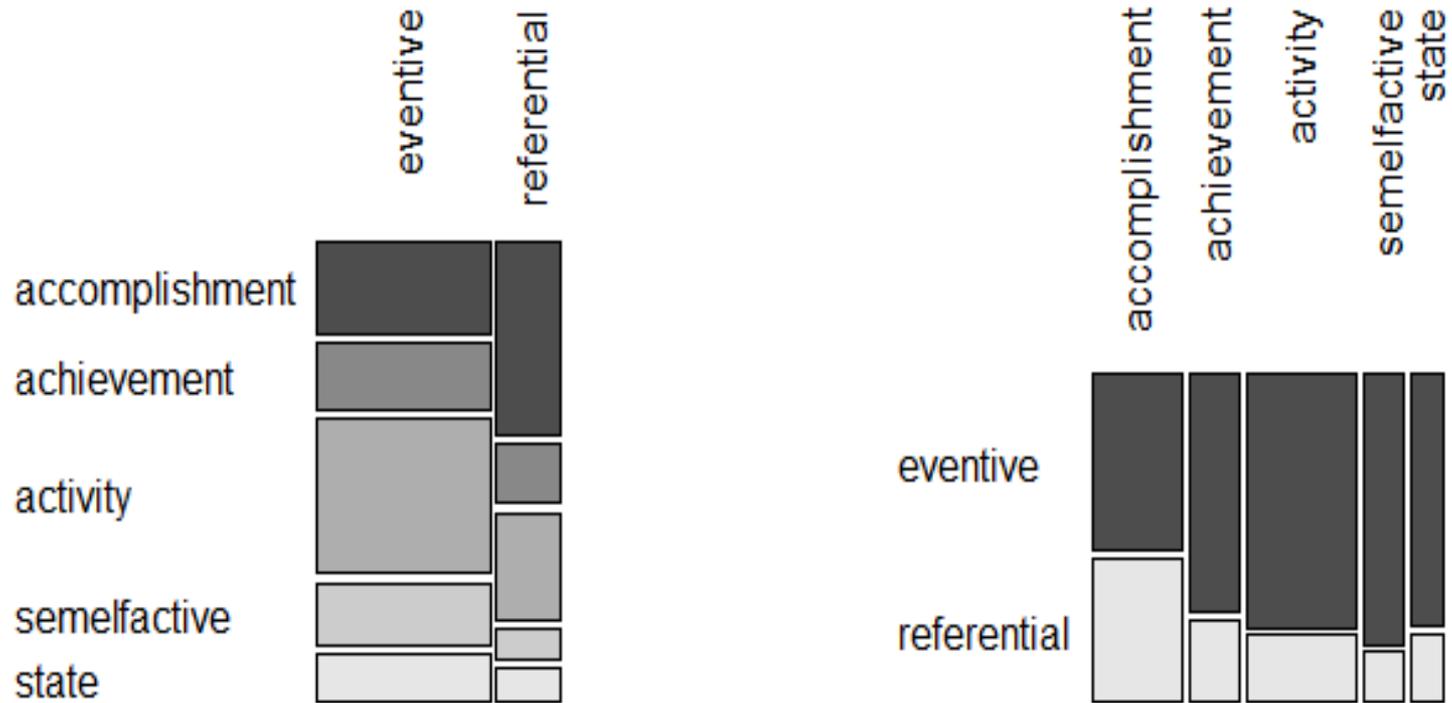Figure 4: Quantification by aspectual class (N=1934)

Aspectual classes

Figure 5: Eventivity by aspectual class (N=2568)