How independent are the variants of a linguistic variable?

and how much does it matter?

Daniel Ezra Johnson (Lancaster University)

Bill Haddican (CUNY-Queens/Grad Center)

IRCS Common Ground, March 25, 2015

Thanks to Meredith Tamminga, Charles Yang, Nanna Haug Hilton, Anders Holmberg, and audiences at CUNY, Georgetown, NWAV, Cambridge, LAGB, Manchester, and NWAV

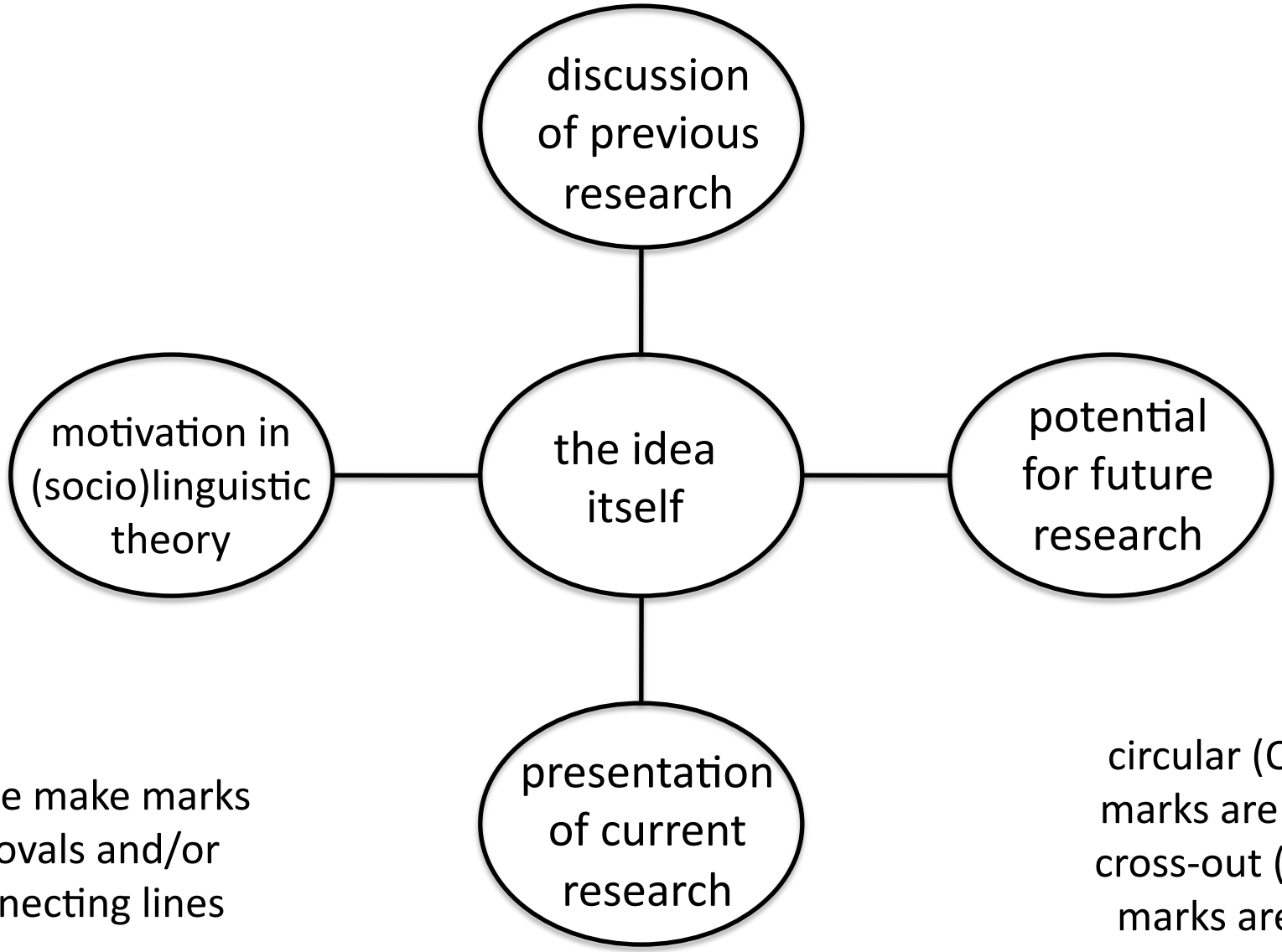
how are good ideas received?

- I had imagined a long and bitter struggle for my ideas, where I would push the social conditioning of language against hopeless odds, and finally win belated recognition as my hair was turning gray. But my romantic imagination was cut short. They ate it up! (William Labov, remembering his presentation of "The Social Motivation of a Sound Change" at the LSA Annual Meeting in 1962)
- All truth passes through three stages. First, it is made ridiculous or distorted. Then it is opposed. In the end, it is accepted as self-evident. (attributed to Schopenhauer)
- Truth is granted only a short victory celebration between two long periods of time: the first when it is condemned as paradoxical and the second when it is scorned as trivial. (Schopenhauer, 1818)
- First they ignore you, then they laugh at you, then they attack you, then you win. (attributed to Gandhi)
- First they ignore you. Then they ridicule you. And then they attack you and want to burn you. And then they build monuments to you. And that is what is going to happen to the Amalgamated Clothing Workers of America. (Nicholas Klein, 1918)

how are bad ideas received?

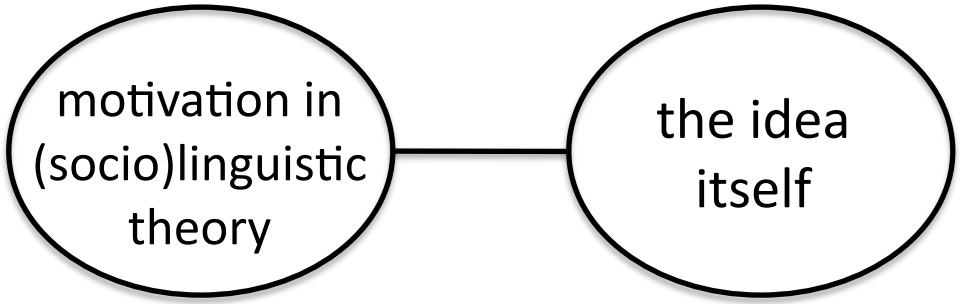
- ignored?
  - the converse doesn't hold
- not understood? (e.g. if incoherent)
  - this could lead to being ignored: save face, save time?
- laughed at?
  - the converse doesn't hold
  - politeness could translate to ignoring in public
- faint praise
- attempts to clarify, correct, or oppose
- we've gotten some of these (and better) but…
- feedback can be scanty, vague, biased, or forgotten

talk structure (feedback form)



please make marks on ovals and/or connecting lines

circular (O-like) marks are good, cross-out (X-like) marks are bad


motivation in (socio)linguistic theory
the idea itself

what is a linguistic variable?

- (syntax) Lavandera 1978, Labov 1978, Romaine 1984, Cheshire 1987, Fasold 1991, Wolfram 1991…
- heuristic: practical, sufficient for the immediate goals
- Wolfram: until 1969, it was a way for sociolinguists to answering our own questions (language and society)
- w/ variable rules, offered as part of linguistic theory
  - as soon as you do this, you're tied to a linguistic theory!
- Wolfram recommends the original approach
- does this mean that if we have different questions, we would define the linguistic variable differently?
- also, especially lately, sociolinguists' questions are (again) questions of linguistic theory anyway

what causes a linguistic variable?

- conditioned/variable rule or process, A -> B (Type 1)
  - sometimes a distinction: inside or outside the grammar
- conditioned choice or competition, A vs. B (Type 2)
  - unclear if this replaces or supplements Type 1
  - locus not always clear, nor (relatedly) what is competing
    - phonemes, morphemes, anything? functional heads, grammars?
  - to compete/be chosen among, A and B must both exist
  - competition can occur at selection/insertion or later (?)
  - possibility of competition not just at one point
    - can a "surface variable" result from several choices/a prolonged competition? vs. assumption that all effects are simultaneous…
- Type 1 is theoretically clear (why mere heuristic?)
- Type 2 less clear (still aims at theory, why heuristic?)

method in variationist sociolinguistics

- a new (and better) type of data
  - the sociolinguistic interview preceded the variable rule
- earliest quantitative methods (Labov 1963, 1966) simple
- later, codification of data analysis principles
  - principle of accountability (on which the linguistic variable rests)
    - any variable form (a member of a set of ways of "saying the same thing") should be reported with the proportion of cases in which the form did occur in the relevant environment, compared to the total number of cases in which it might have occurred (Labov, 1972)
- a standard method of data analysis:
  - additive model, multiplicative models (c. 1969 - 1974)
  - logistic model (matches other fields, 1975 - present)
- method (multiple logistic regression) constrains data
  - possible data: cell proportions or individual binary tokens

what the standard method can tell us

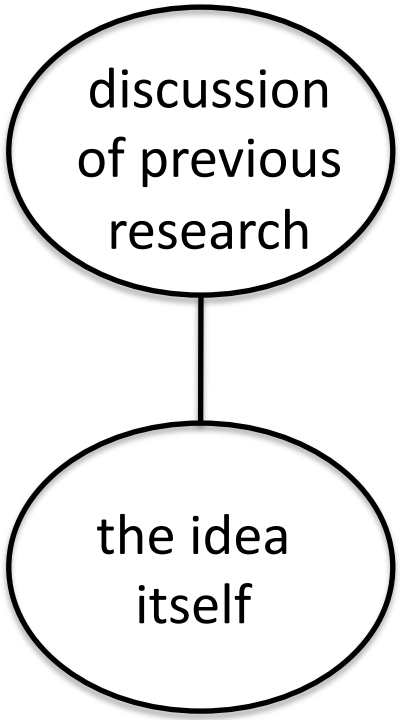
- assuming a variable of type 1 (A -> B) or type 2 (A vs. B)
  - standard approach quantifies constraints on the variation
  - size and significance of predictors of various types
- on the surface, an inverse pattern (or mirror pattern)
  - what favors A disfavors B, what favors B disfavors A
- for type 2 (A vs. B), inverse pattern is inherent/underlying
- for type 1 (A -> B): it depends
- imagine XAY -> XBY
  - if both input XAY and output XBY affect rule application, probable inverse pattern (indistinguishable from type 2?)
  - if the output XBY does not affect rule application, e.g. what favors A disfavors A -> B, so it disfavors B, but… what favors B doesn't favor A -> B, so doesn't disfavor A!

what the standard method can't tell us

- can't distinguish type 1 (rule) from type 2 (choice)
  - sometimes we must be dealing with a choice, most self-evidently when the variants cannot be related by a rule
  - harder to say that we ever must be dealing with a rule
    not about to solve this! cf. rules vs. constraints in phon.
- can't distinguish a possible type 3 (variant independence)
  - what favors/disfavors A may not affect B (and vice versa)
  - this is inherently a non-inverse pattern
  - the standard method of regression on proportions always "returns" an inverse pattern
  - the same goes for the experimental method of Bresnan 2007, in which subjects distribute 100 points between variant sentences
- "the idea": variants do sometimes behave independently, and this can be seen (for syntax) w/ acceptability ratings

other things it can't tell us (not in talk)

- if anything can be learned from variant frequencies
  - what if both variants increase/decrease (syn- or diachronically)?
  - what if one increases/decreases and the other stays the same?
  - what if one increases and the other decreases, but not equally?
  - for example, in a new study of filled pauses by Wieling et al.:
    - steady increase in proportion UM / UM + UH similar in English and Dutch
    - but for English, decrease in UH is at least twice as large as increase in UM
    - Dutch change in UH is non-monotonic, overall smaller than change in UM
- anything about single items (not variants)
  - these can be uninteresting/"non-linguistic" but not always
  - "but the principle of accountability" – but logistic regression!
- pushing the envelope: e.g. particle verb alternation
  - Bob threw his keys in (VOP) vs. Bob threw in his keys (VPO)
  - Bob threw his keys in the drawer (VOP, no VPO option)


discussion of previous research
the idea itself

a partial discussion of some partially-related research

- this is a very incomplete discussion
- please let me know of other research about on these issues!
- speaker-rating approaches
- experiments where the choice of variant is among IVs and the DV is some kind of rating of the speaker
- psycholinguistic approaches
- asymmetrical persistence patterns (not RT priming) not sure if results have been (or could be) interpreted as evidence for variant independence
- acceptability approaches
- evidence for variant independence
  not sure if results have been intepreted this way
- research on (syntactic) acceptability (but not variants)

speaker-rating approaches

- Labov et al. 2011 “Properties of the sociolinguistic monitor”
  - judging professional suitability based on LV (ING)
  - given the formal context, -in’ is marked, -ing is expected
  - logarithmic pattern: each additional -in’ has less effect
  - makes sense, but turn it around (or imagine production)… social factor not related to LV in (symmetric) logistic way
- Campbell-Kibler 2011
  “The sociolinguistic variant as a carrier of social meaning”
  - rating on a range of social attributes based on (ING)
  - (unlike dissertation?) three guises: [IN], [In], and [obscured by noise]
  - we can observe independence: “difficult or impossible to detect through methods that contrast one variant against the other” (435)
  - -ing: intelligent, educated, articulate, not a student
  - -in’: casual, not gay
  - independent social meanings because it’s two morphemes?
  - affects (third-wave) concept of LV

psycholinguistic approaches

- Tamminga 2014 “Persistence in the production of linguistic variation”
  - use of one variant favors use of the same variant again
  - depends on many things (time, grammatical similarity. . .)
- persistence asymmetries (not all classic variables)
  - BP: non-marking inflection stronger effect than marking it
  - Standard English: passive stronger than active, -in’ stronger than –ing
  - AAVE: 3sg zero stronger than 3sg -s (inanimate prime)
  - generalization: “inverse frequency effect” (expectation)
- does it suggest we don’t just store/attend to proportions?
  - given 10% A, 90% B, “surprise” at A requires proportions
  - yet over time, knowing proportion, how could A “surprise”?
- Loudermilk 2013 measured “N400-like” potentials (EEG)
  - those for -in affected by (ING) use in preceding context*
  - those for -ing not affected

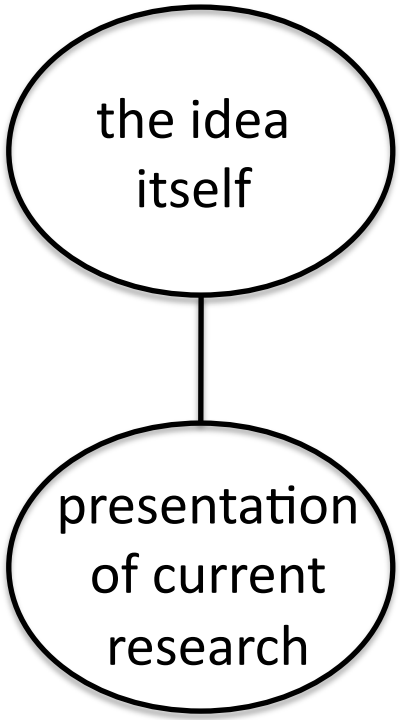
*anti-priming effect?

acceptability approaches

- Grondalaers and Speelman, 2008
  "Constructional near-synonymy, individual variation, and grammaticality judgments: Can careful design and participant ignorance overcome the ill reputation of questionnaires?"
    - (Warning: Theory "cancels the traditional distinctions between sociolinguistics and theoretical syntax, and between sociolinguistics and psycholinguistics")
- since appropriate corpus not available, "we have no choice" but to elicit grammaticality [sic] judgments (7-point scale)
  - In de asbak lag (er) een hagelkorrel.
    In the ashtray lay (there) a hailstone.
- sentences without er get worse with certain factors:
  - vague locative: -0.41 temporal: -0.98 verb *zijn*: -1.06
- sentences with er do not get better with same factors:
  - vague locative: -0.01 temporal: -0.03 verb *zijn*: -0.24
- makes sense if er is for repair, but there's no inverse pattern
  - variants er and zero are independent - if they are even variants!

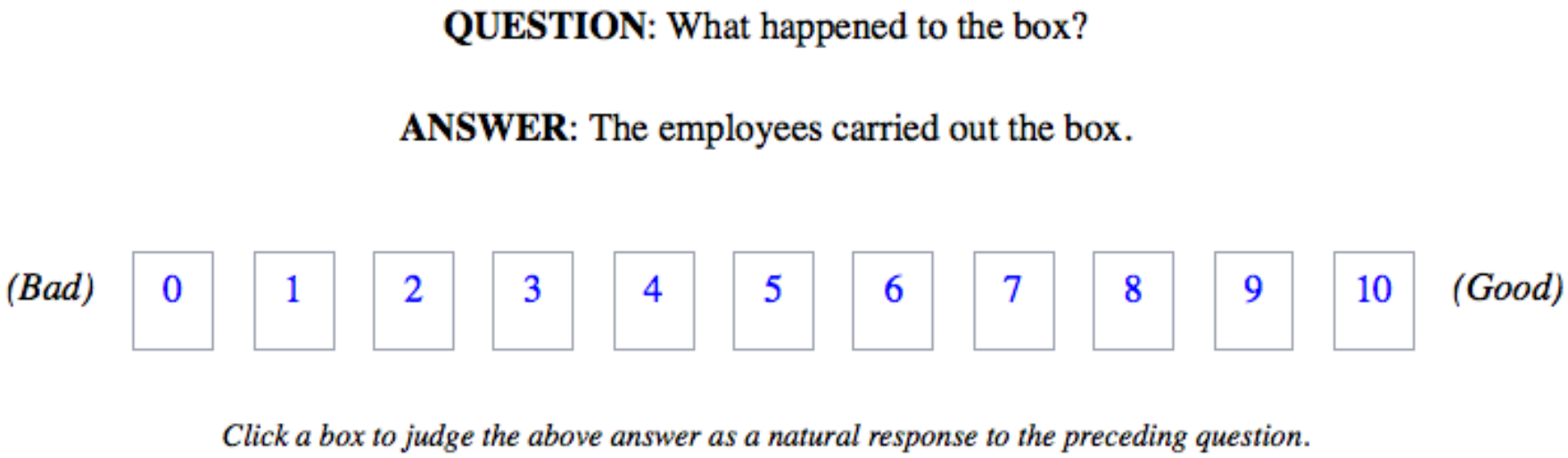
why might it be one way vs. the other?

- most linguistic variables affected by both social factors (have social meaning) and linguistic factors (more obscure)
- social meanings are complex and (following C-K) it may be unsurprising for them to attach to individual forms
- Labov (p.c.?): overt comment attaches to individual forms, rarely to alternations, distinctions, mergers, proportions
- speculation: linguistic factors could be of two types?
  - 1) where the association is arbitrary/historical
    - example: (ING): -ing more nominal, -in' more verbal, due to historical accident
    - this is more like social factors (so expect variant independence)
  - 2) where one variant is better than the other b/c articulation, processing, persistence, etc.
    - example: /t,d-/-deletion, following vowel vs. consonant
    - more like rule/process/choice (so expect inverse pattern)
- note: a LV can behave both ways, depending on the factor


the idea
itself
presentation
of current
research

Experimental methods

- All three experiments were essentially the same, except:
- Norwegian stimuli presented with no preceding context.
  English stimuli had a preceding sentence, and said:
  "Judge this sentence as a response to the preceding context."
- 11-point rating scale (0 = 'bad', 10 = 'good')
  Various lexicalizations (sentence types) created, 50% fillers,
  experiments done online with Ibex Farm (Drummond, 2013)



Data analysis methods

- 1) normalize responses
  - for each subject, get mean / s.d. of fillers (ranging OK to *)
  - for each response (0-10), create z-score using mean / s.d.
  - this step is actually barely even necessary

- 2) adjust responses, removing lexicalization-specific effects
  - fit maximal mixed-effects model (lme4 package)
    - within-speaker fixed effects:
      - e.g. order, object weight
    - random effects:
      - intercept / all slopes by subject
      - intercept / all slopes by lexicalization
  - subtract lexicalization effects, keep fixed and subject effects
  - this method was made up by me
    - if you want to know why I resorted to doing this, I could tell you

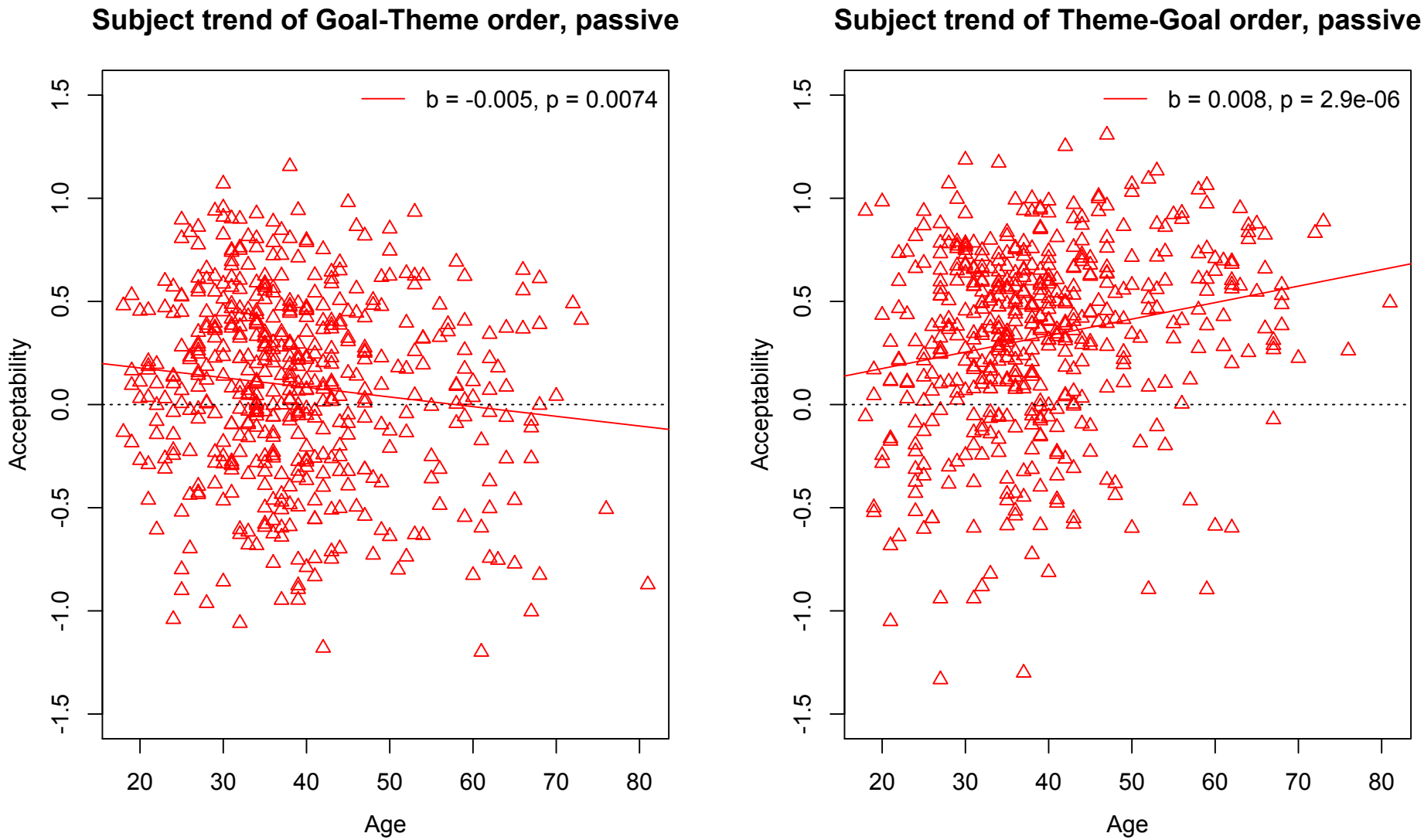
Experiment 1: Object ordering in Norwegian

- 500 subjects (.01% of population of Norway)
- age 18-81, mean 39
- 2 x 3 design, 4 items/condition

Context	theme-goal	goal-theme
Passives	Den ble gitt ham. 'It was given him.'	Han ble gitt den. 'He was given it.'
Active OS	Elsa ga den ham ikke. 'Elsa didn't give it him.'	Elsa ga ham den ikke. 'Elsa didn't give him it.'
Active-non-OS	Elsa har ikke gitt den ham. 'Elsa hasn't given it him.'	Elsa har ikke gitt ham den. 'Elsa hasn't given him it.'

- Table 1: Example sentences for six conditions

Diachronic effect: Norwegian passives



- Figure 1: Norwegian passives: acceptability of goal-theme (Han ble gitt den) and theme-goal (Den ble gitt han), by speaker, in apparent time (N = 500)

Diachronic effect: Norwegian passives

- Mirroring slopes for the effect of age is exactly the pattern we expect if grammatical change reflects incremental change in the probability of choosing one abstract representation vs. a competing one—"grammar competition" in the terms of Kroch (1989, 1994).

- It's not that this particular example – "He was given it" vs. "It was given him" – contains the kind of functional doublet Kroch had in mind. (Not sure.) The point is that the inverse patterning is not just unsurprising, it's hard to imagine any other pattern arising given the basic assumptions we make.

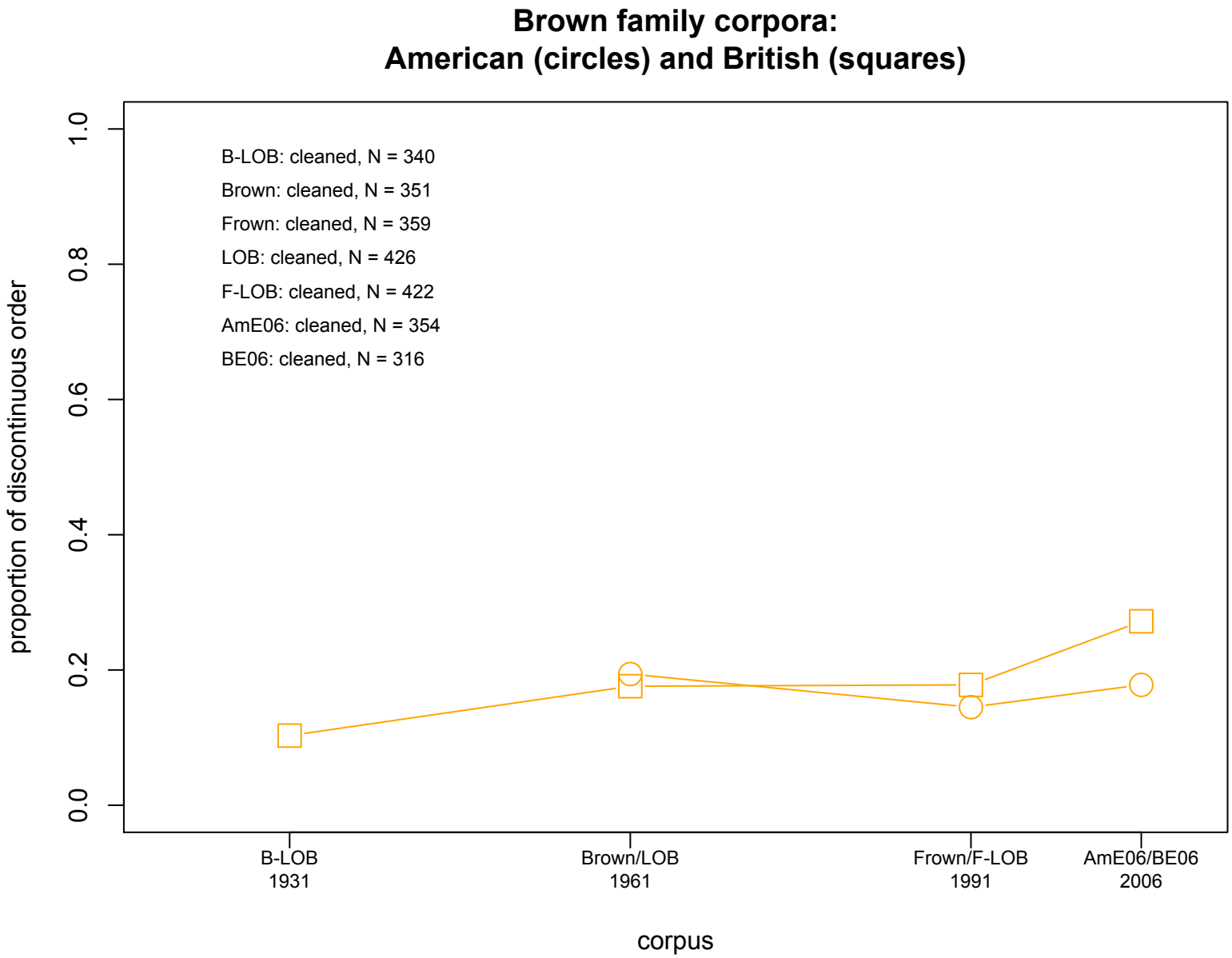
- But such isn't always the case . . . .

Introduction to the particle verb alternation

- Kim cut the melon open. [VOP]
- Kim cut open the melon. [VPO]

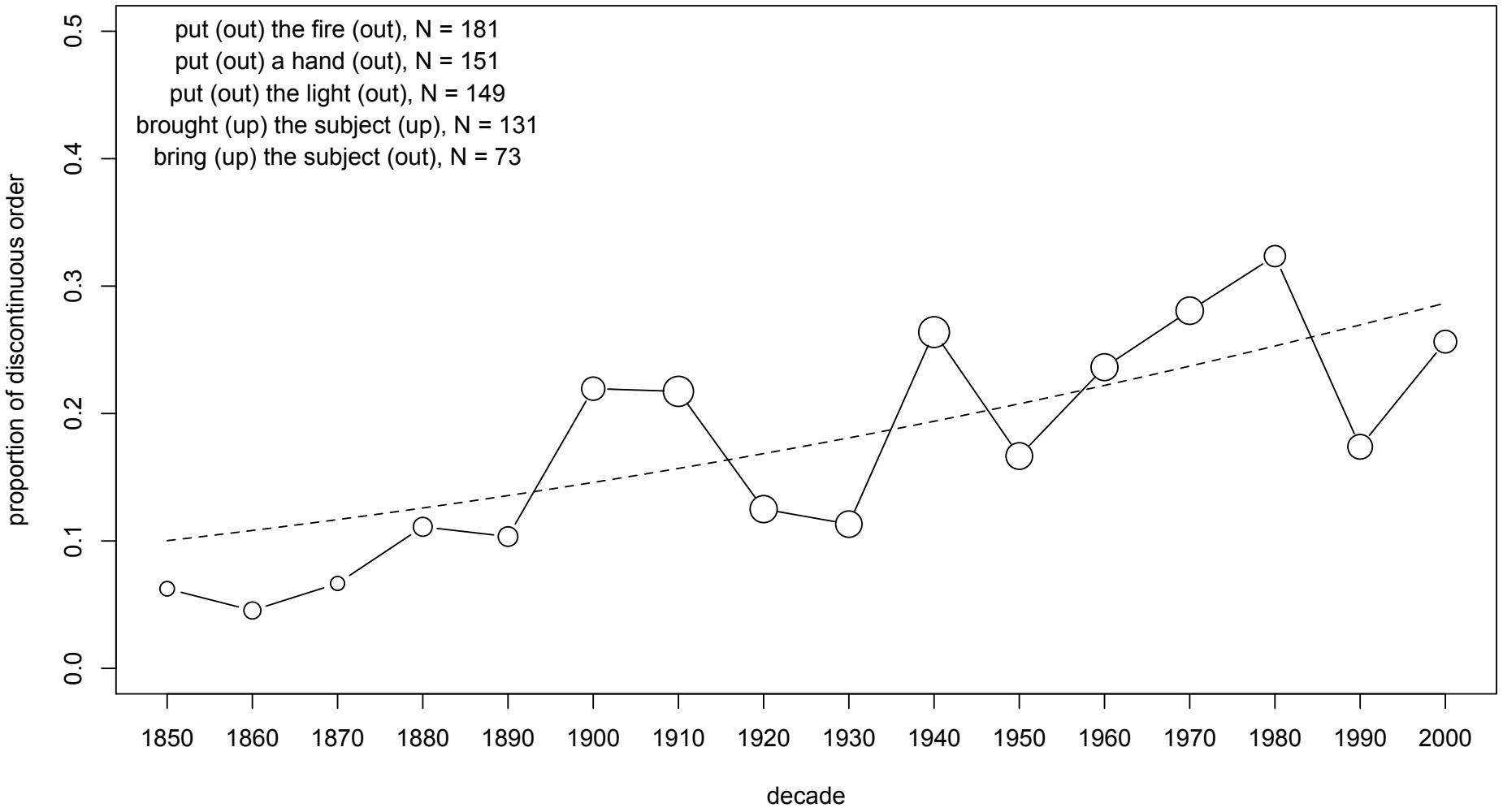
- two effects identified long ago - in the choice paradigm:
  - 'light' object favors VOP, 'heavy' object favors VPO
  - 'old' object favors VOP, 'new' object favors VPO
- two effects identified in our work - in the choice paradigm:
  - UK subjects favor VOP, US subjects favor VPO
  - younger subjects favor VOP, older subjects favor VPO
- recent literature:
  - Gries (2001, 2003): broad multi-factorial approach to alternation
  - Cappelle (2009): 'contextual' factors affecting alternation
  - Larsen (2014): syntax of alternants (445 pp., little on the alternation)

Particle verb alternation: diachronic effects

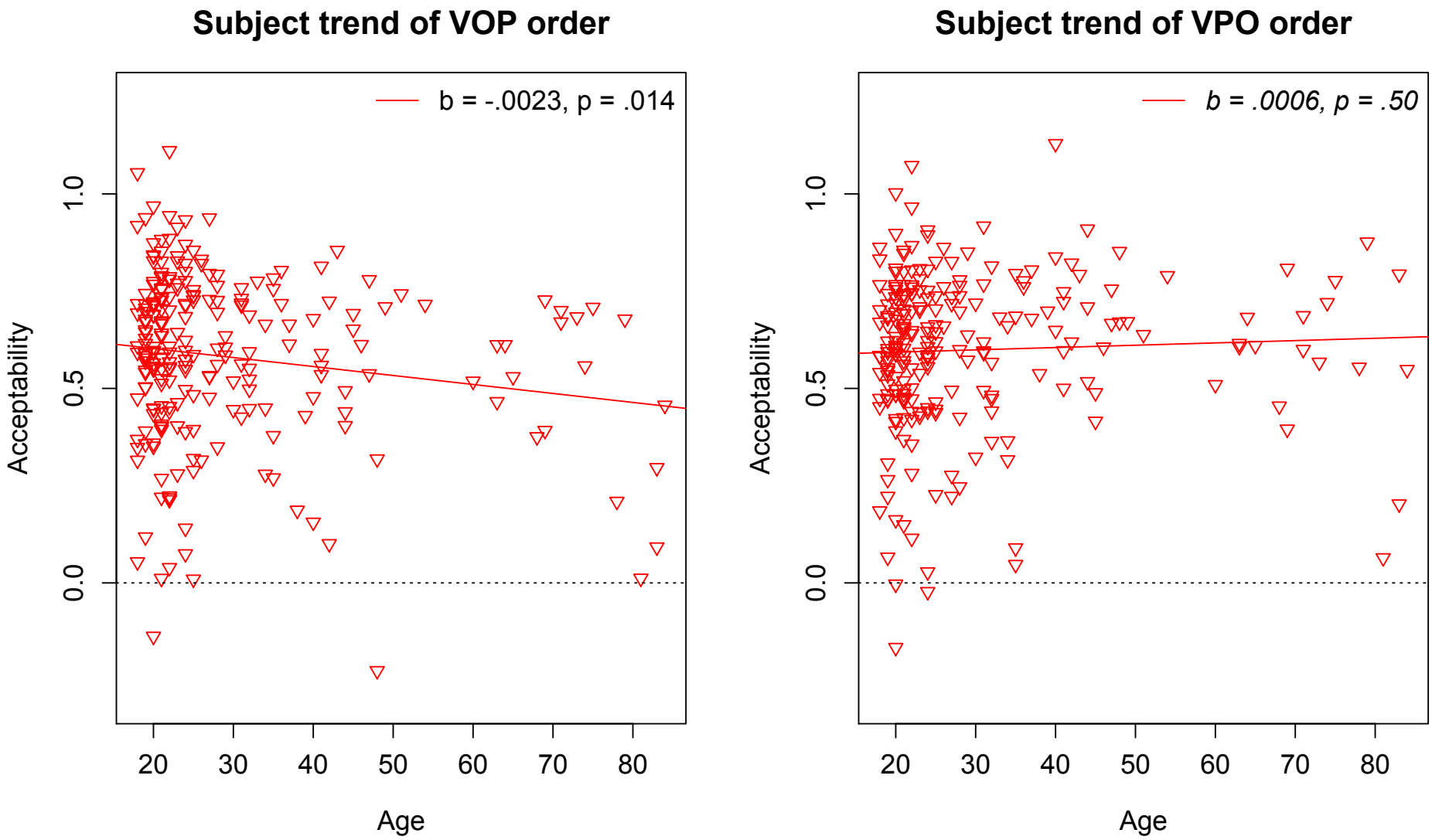


Particle verb alternation: diachronic effects

COHA (Corpus of Historical American English)



Particle verb alternation: diachronic effects



- Figure 2: English particle verb alternation: acceptability of VOP (Kim cut the melon open) and VPO (Kim cut open the melon), by speaker, in apparent time (N = 239)

Particle verb alternation: diachronic effects

- Figure 2 shows that this change has co-occurred with an apparent time increase in the acceptability of the VOP order, but no significant change in the acceptability of the VPO order.
- This is not expected on an approach to change in acceptability where both variants are always affected inversely and equally.
- Why some processes of change should show this inverse pattern, but not others, we can't answer.
- But from the perspective of the competition model, we feel no one can answer this question (if they could even ask it).

Particle verb alternation: constraints

- the 'social effects'
  - little "social meaning" (this variable seems to have very low salience)
  - any associations likely arbitrary - variant independence?
- the effect of object weight
  - head-initial Ls: "end-weight" (Behaghel, 1909; Quirk et al., 1972)
  - heavy middle item hard to process (Hawkins 1995; Lohse et al. 2004)
    - implies effect on VOP only?
- the effect of object newness
  - in English: old/topic before new/focus (19th c., Prague School)
  - high-level effect, predicts inverse pattern (?):
    - VPO order - new object > old object
    - VOP order - new object < old object
  - unless only real effect is *late-old
    in that case, predict variant independence:
    only one order would be affected by old/new manipulation

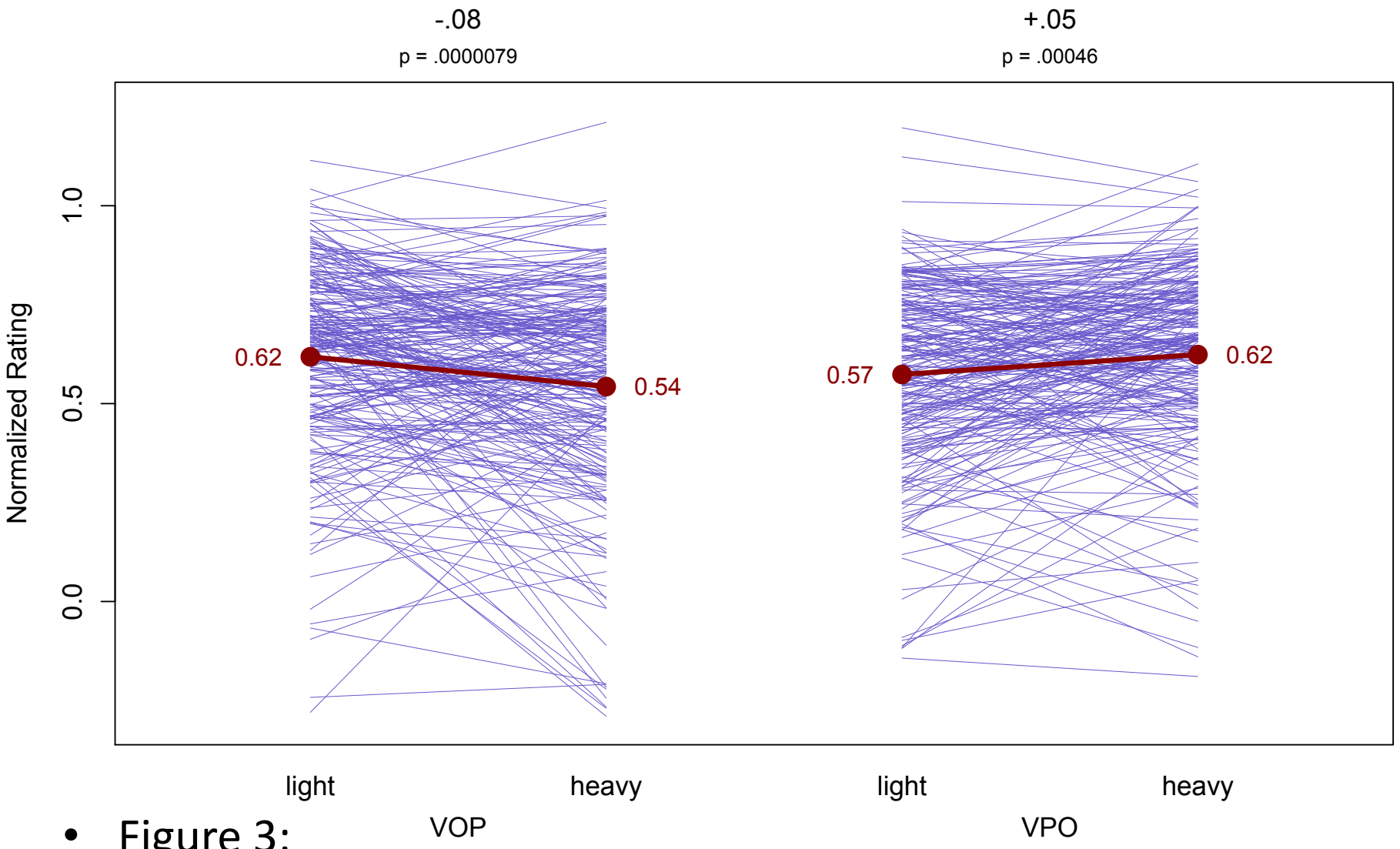
Particle verb alternation: object weight

- 113 US vs. 126 UK subjects, age 18-84 (mean 30)
- crossed: order, object weight (3 vs. 7 syllables), object focus
- object focus removed: 'old object' made all ratings go down

Object Weight	Verb-Object-Particle	Verb-Particle-Object
Light	*. . . cut the melon open*	*. . . cut open the melon*
Heavy	*. . . cut the heavy juicy melon open*	*. . . cut open the heavy juicy melon*

- Table 2: Four conditions

Particle verb alternation: object weight



- Figure 3:
- Weight effects on acceptability of VOP and VPO orders

Particle verb alternation: object weight

- If we call this an inverse or mirror pattern, it is only partial.
- Increasing the weight of the object from from three to seven syllables disfavors the Verb-Object-Particle order 50% more than doing so favors the Verb-Particle-Object order.

Object Weight	VOP	VPO
Light	0.618	0.574
Heavy	0.543	0.624
$\|\Delta\|$	-0.075	$+0.050$

- Table 3: Average acceptability for four conditions

Particle verb alternation: object weight

- The result for the VOP order can be explained by Lohse et al.'s (2004) processing-based account. A heavier object separating the verb and particle increases the size of the verb-particle processing domain.

- For object weight to affect the VPO order is unexpected from this perspective, since a heavier final object NP should have no effect at all on the size of the processing domain for the relevant dependency relation.

- Maybe when subjects evaluate the acceptability of a given syntactic structure, they implicitly compare it with a competing structure in the same environment.

- If so, which structures compete? Why doesn't this always happen?

Particle verb alternation: object focus

- Several authors have noted subtle focus effects on particle placement. New information objects prefer the VOP order.
- Sometimes explained in terms of focus-drive movement (Kayne, 1998; Haddican & Johnson, 2014), or in terms of syntax-prosody mapping (Svenonius, 1996; Dehé, 2002).

- Q: Who will you pick up? (Obj. focus)
- A: I'll pick (?the girls) up (the girls).

- Q: How are Turid and Ingrid going to get here? (Obj. given)
- A: I'll pick (the girls) up (?the girls).

(Svenonius, 1996)

Particle verb alternation: object focus

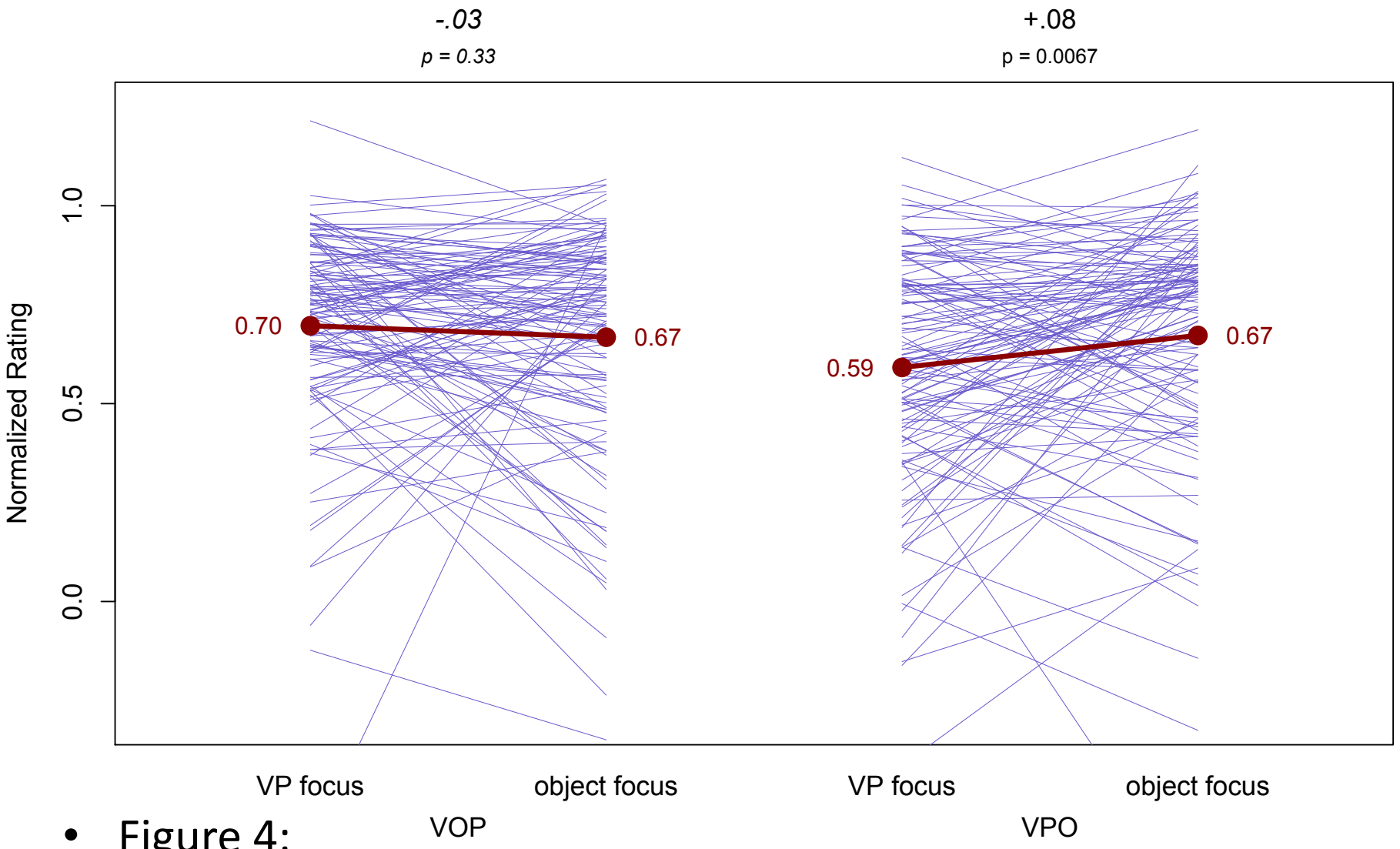
- 125 US subjects (age 18-52, mean 23)
- crossed: order, focus (sentence, VP, object, object topic)

- Q: What happened? (wide focus)
- A: Ann cut (the tree) down (the tree).

- Q: What did Ann do? (VP focus)
  A: Ann cut (the tree) down (the tree).
- Q: What did Ann cut down? (object focus)
  A: Ann cut (the tree) down (the tree).

- Q: What happened to the tree? (object topic)
  A: Ann cut (the tree) down (the tree).

Particle verb alternation: object focus



- Figure 4:
- Focus effects on acceptability of VOP and VPO orders

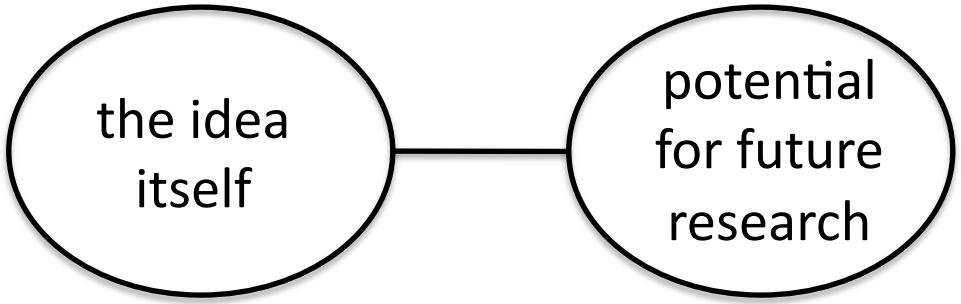
Particle verb alternation: object focus

Object Weight	VOP	VPO
VP focus	0.697	0.591
Obj. focus	0.668	0.672
$\|\Delta\|$	-0.029	$+0.081$

- Focus has a relatively strong effect on VPO orders
- At best it has a weaker effect on VOP (not significant)
- None of the available theories (our previous work included) predicts that VPO orders should be more sensitive to focus effects than VOP orders.
- Object is more likely to be included in a modern syntactic theory than object weight, but… does this really make sense?

Main points

- Much has been learned from the standard methodology that treats variants of a linguistic variable as choices (or the input and output of rules/processes).
- From this perspective, binary variants always appear to respond inversely to the factors affecting variation.
- However, in some respects variants can also behave independently with respect to these factors.
- Our results suggest that different factors (both internal ones and social/diachronic ones) can affect the variants of the same variable either inversely or independently – or perhaps along a continuum in between?
- Achieving a better understanding of these differences should surely be a goal of variationist sociolinguistics.


the idea itself
potential for future research

Potential for future research

- Again, distinguish between the idea (variant independence) and the method using acceptability judgments, which are:
  - hard to collect (in part because very noisy and sensitive to everything)
  - basically useless diachronically/historically
  - not easy to "calibrate" to production, a two-step process
    - what does the scale mean?
    - if variants are/were independent, (how) do they still feed into a production choice?
    - still, work of Bresnan is upbeat that speakers' judgments do reflect production
  - an inverse effect (+A, -B) plus a common effect (+A, +B) can look like anything: (+A, -B), (+A, B), (+A, +B)
- What are the alternatives?
  - What traditions, journals, could be relevant? Corpus linguistics (gasp)?
- What is the connection (if any) to syntactic structures?
  - or phonological structures, if we were looking at those?
  - syntacticians are looking at experimental acceptability judgments too
    - different sentence types, rarely alternations, not easy to make the connection

Potential for future research

- methodological shackles
  - using a "heuristic" should mean using other heuristics are OK
  - Labovian revolution looked at new data in new ways
    - it mainly became one single way, but never fully, and this is clearly changing
  - it (necessarily) sets some data to the side in order to look at other data
- "we" (or at least "I") still tell the story of 1960-1969
- today, "our" variationism can seem conservative or restrictive
  - if not worse, but I don't even want to go there, as it's not my opinion
- not trying to argue against logistic regression or its results!
- with this tool, can't measure something I think we'd like to
- expanding methods, we have nothing to lose but our chains
- and something to gain… our chain shifts?
- *Soziolinguisten aller Länder, vereinigt Euch!*