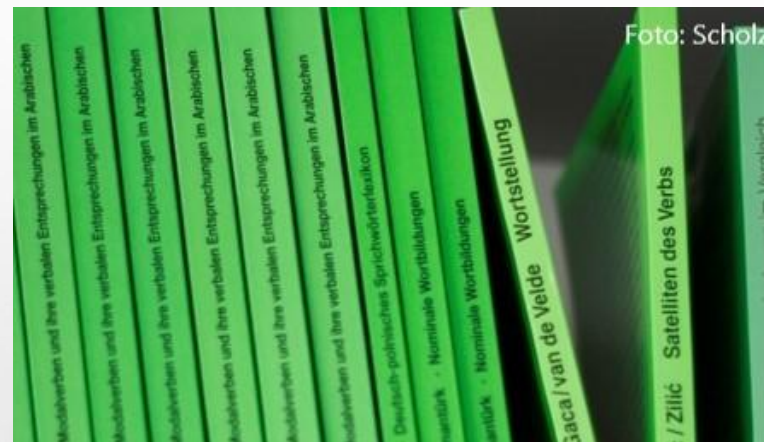




Thilo Weber

FREE VARIATION IN GRAMMAR? PERSPECTIVES FROM CORPUS LINGUISTICS AND GRAMMATICAL THEORY

LMU Munich, July 21 2026

“[T]here is a growing body of research highlighting the prevalence and diachronic stability of variation within and across individuals, which is not correlated with any grammatical or social meaning within language communities.”

(BAYLAT-FAPESP workshop, project description)

This talk

- A potential example of partially free variation in the grammar of German
- Free variation in grammatical theory
- My take on the idea of free variation

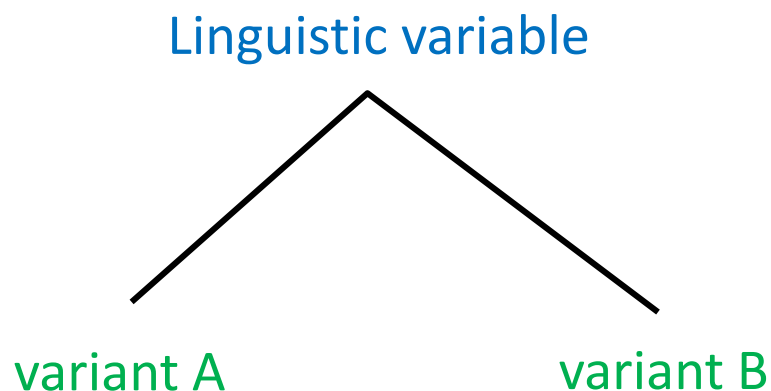
A POTENTIAL EXAMPLE

Background: IDS-project *Corpus Grammar – Grammatical variation in Standard and near-Standard German*

- Focus on morphological and syntactic variation in German
- Data sources (Bubenhof/Konopka/Schneider 2014):
 - German Reference Corpus (DeReKo), mainly newspaper texts (Kupietz et al. 2018)
 - webforum discussions from the DECOW web corpus (Schäfer/Bildhauer 2012)

A POTENTIAL EXAMPLE

“It is common for a language to have many alternate ways of saying ‘the same’ thing”. (Labov 1972: 188)



→ Can we – at least probabilistically – predict under what conditions speaker choose one variant rather than the other?

Typical procedure

- Identification of variable phenomena
- Formation of hypotheses on influencing factors
 - internal factors
 - external factors
- Data extraction from the corpus
- Coding of the corpus data according to potentially influencing factors
- Quantitative analysis using inferential statistics

A POTENTIAL EXAMPLE

The phenomenon (Weber, Bildhauer & Münzberg 2024)

Linguistic variable: noun modification

variant A: *dass*-clause
'*that*-clause'

variant B: *zu*-infinitive
'*to*-infinitive'

A POTENTIAL EXAMPLE

The phenomenon (Weber, Bildhauer & Münzberg 2024)

- Interchangeability under some circumstances
- But not under all circumstances

*Solche Ausstellungen_i hätten den **Vorteil***
Such exhibitions have. 3RD.PAST.SUBJ the advantage
'Such exhibitions would have the advantage'

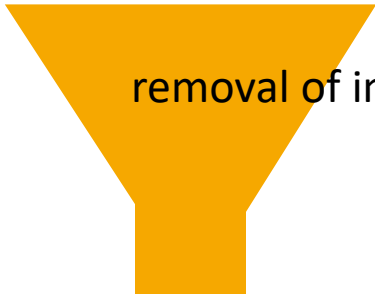
a. **dass** *sie_i Unternehmen zusammenbringen*
that they companies together.bring.3RD.PL.PRES
'that they bring companies together'

- No explicit subject
- Restriction to contexts where subject reference can still be reconstructed.

b. **PRO_i** *Unternehmen zusammenzubringen*
companies together.bring.INF
'of bringing companies together'

A POTENTIAL EXAMPLE

53.349 potential examples extracted from corpus



removal of invalid examples

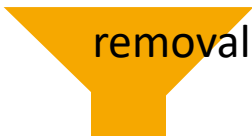
9.715 examples



removal of non-interchangeable examples

- following a set of criteria defining 'interchangeability'
- parallel and independent annotation by two annotators (kappa = 0.8: 'very substantial interrater-agreement')

6.101 examples



removal of examples dependent on a non-varying head-noun

1.928 examples = variable in our more narrow sense → subjected to statistical analysis: Felix Bildhauer

Illustration: Felix Bildhauer

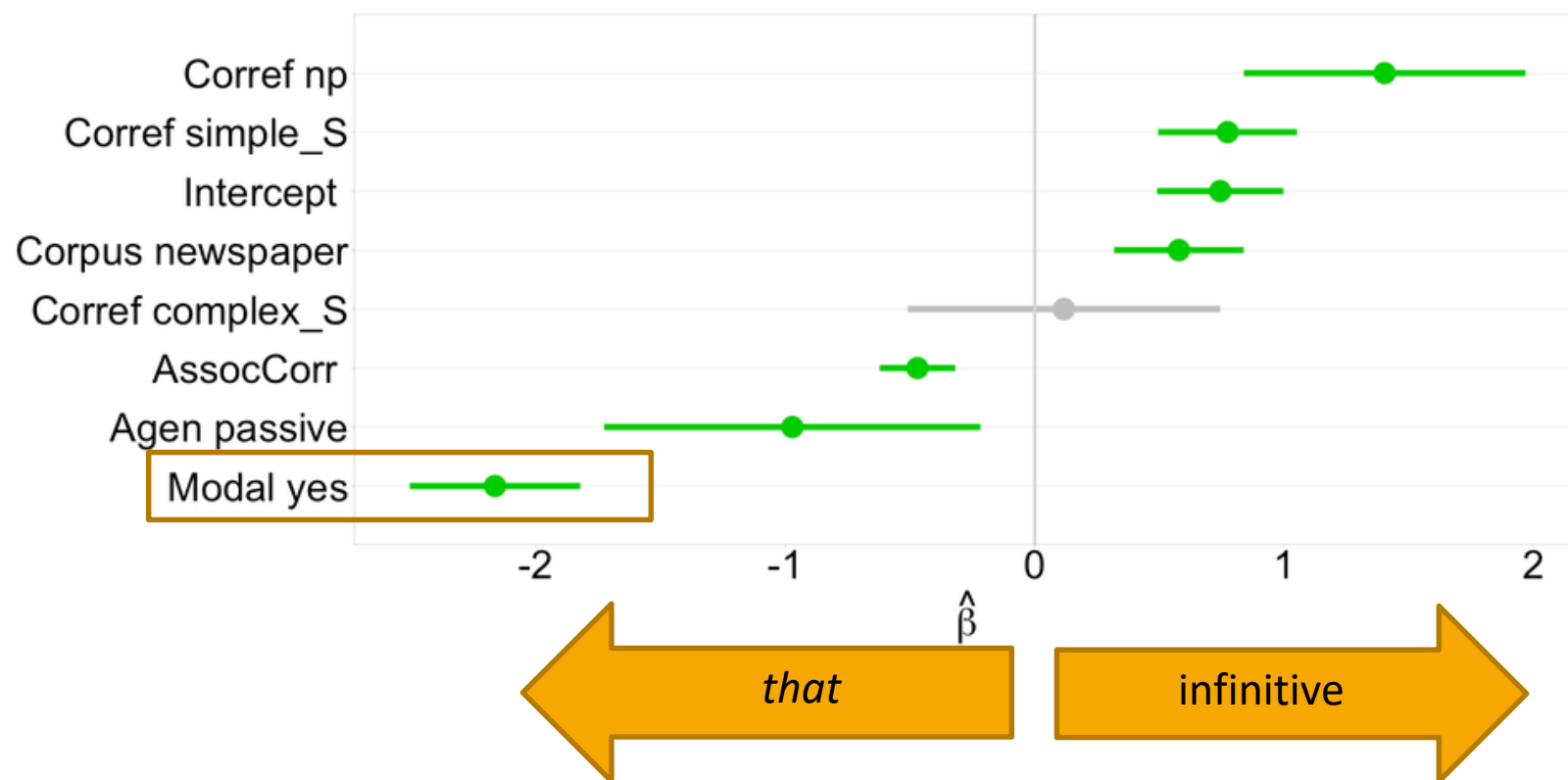
A POTENTIAL EXAMPLE

Statistical analysis: model specification

- logistic regression
- mixed model (GLMM)
- fixed effects: 5 independent variables
 - 4 internal/grammatical variables
 - 1 external variable: “register” (news paper vs. webforum)
- random effect: the head noun

A POTENTIAL EXAMPLE

Coefficient estimates



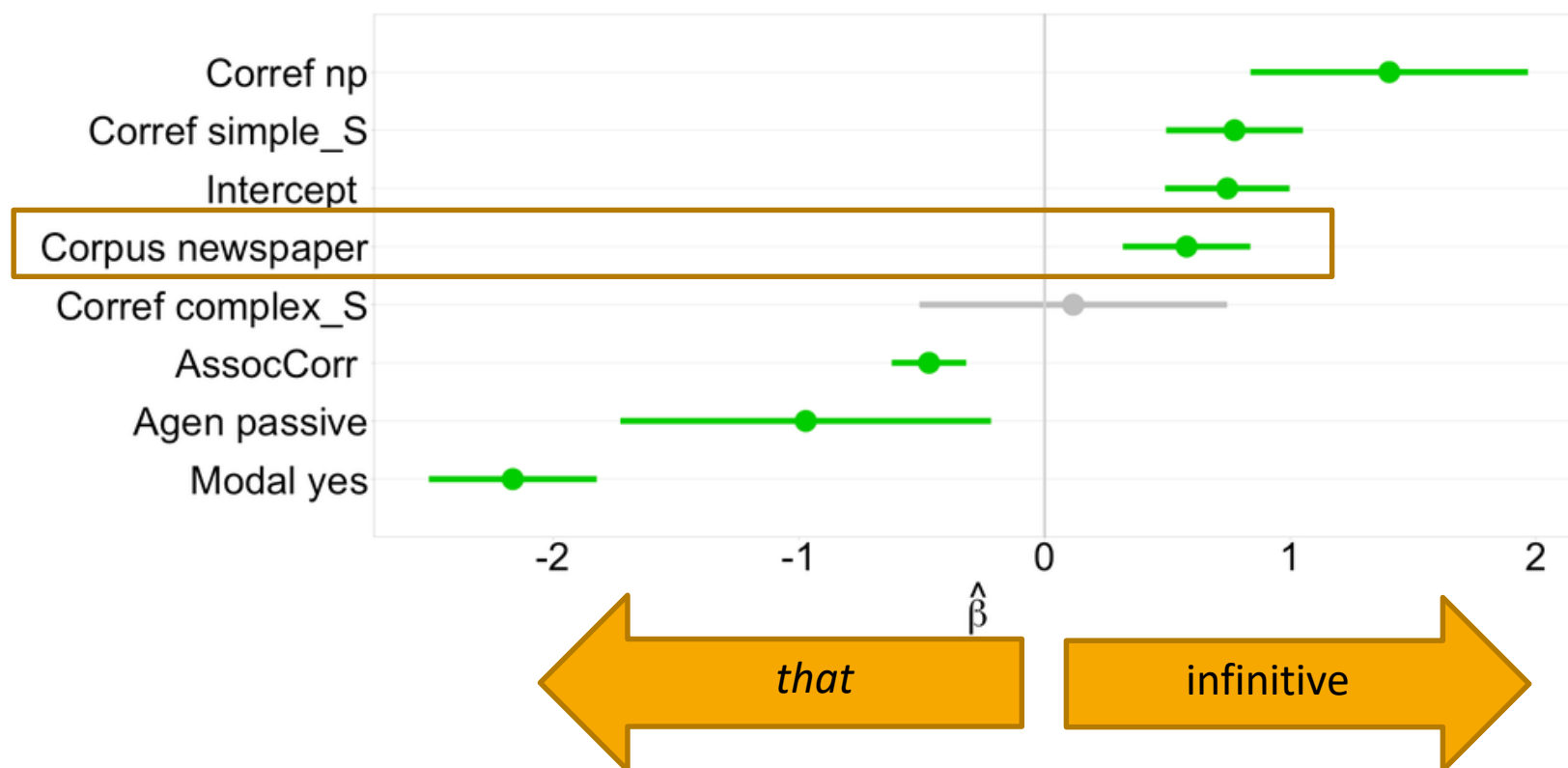
Internal factors, e.g. modal verb:

- Adnominal clause contains a modal verb → higher probability of *that*

R^2 marginal = 0,24, R^2 conditional = 0,34

A POTENTIAL EXAMPLE

Coefficient estimates



External factor:

- Newspaper text (vs. webforum) → higher probability of infinitive.

R^2 marginal = 0,24, R^2 conditional = 0,34

A POTENTIAL EXAMPLE

Coefficient estimates



R^2 marginal = 0,24, R^2 conditional = 0,34

- The model only explains a modest proportion of the variance.
- What about the rest?
 - Entirely due to factors not considered?
 - Or at least partly indicative of **free variation**?

- How does grammatical theory deal with the idea of free variation?

The epistemological problem

- But: The hypothesis of complete determinism is not falsifiable, either

“[A]re not such researchers [i.e. opponents to free variation] operating from a theoretical position that does not permit falsification? They believe that systematicity exists and that if we look hard enough we will always find it. Such a position is, I think, untenable both theoretically and methodologically.”

(Ellis 1999: 476)

- Required: investigating every variable phenomenon with respect to every potential influencing factor and showing that we are still left with unexplained variation.

The epistemological problem

- Existence/Absence of free variation **as a general phenomenon** appears to be treated as an **axiom**, e.g. Goldberg's (1995: 67) “Principle of No Synonymy”

“[I]f two constructions are syntactically distinct, they must be semantically or pragmatically distinct [...]. [...]

Corollary A: If two constructions are syntactically distinct and S(emantically)-synonymous, then they must not be P(ragmatically)-synonymous.

Corollary B: If two constructions are syntactically distinct and P-synonymous, then they must not be S-synonymous.”

Attitudes towards free variation cut across functional and formal linguistics

- Rejection of free variation in functional linguistics, e.g.
 - Bolinger's (1977) book *Meaning and Form* explicitly aims to “reaffirm the old principle that the natural condition of a language is to preserve one form for one meaning and one meaning for one form” (1977: x).
 - Partly in reaction to then-current notions of generative grammar: rejection of the idea that “an abstract structure could be converted into x number of surface structures – in either case without gain or loss of meaning” (Bolinger 1977: ix).
 - But: “Did Bolinger's generativist ever exist?” (Newmeyer 1983: 116); disagreement perhaps at least in part merely due to different definitions of 'meaning'.

Attitudes towards free variation cut across functional and formal linguistics

- Rejection of free variation in formal linguistics, e.g.
 - Free variation explained away as diglossia/bilingualism ('competing grammars'), e.g.

“In general, individual grammars do not manifest optional alternations of this type. [...] Where a language has such an alternation, we say that this manifests diglossia, and that speakers have access to two grammars”.

(Lightfoot 1999:94)
 - Free variation explained away as pseudo-optionality, “the favoured solution to optionality in optimality-theoretic syntax” (Anttila 2002: 218)
 - Basic idea: Interchangeability between two forms is only apparent, whereas on closer inspection there is a difference in meaning after all.
 - Applied to OT: The two forms are not alternative outputs of the same input and thus not in competition with one another.

Attitudes towards free variation cut across functional and formal linguistics

- Openness towards free variation in functional linguistics, e.g.

“The rejection of the generativist position in which one variant was regarded as primary and the other as derived from the primary variant, is of course justified and necessary in a cognitive linguistic approach [...]. ”

(Uhrig 2015: 323, 323)

“However, if it is used as a hard and fast rule [as the corollaries in Goldberg (1995) suggest] or even as dogma, it causes more problems than it solves and – if taken at face value by researchers – will stop them from observing exceptions to this general rule.”

(Uhrig 2015: 335)

Attitudes towards free variation cut across functional and formal linguistics

- Openness towards free variation in formal linguistics, e.g.
 - Adger's (2006) investigation of variation in subject-copula-agreement in Buckie English (e.g. *we was/we were*) within the framework of Minimalism:

“What I want to show in this paper is that the combinatorial mechanisms that create structure in language from discrete elements also give rise to variability in the sense we are interested in here: the **non-deterministic** choice of form.”

(Adger 2006: 505, my emphasis)

Attitudes towards free variation cut across functional and formal linguistics

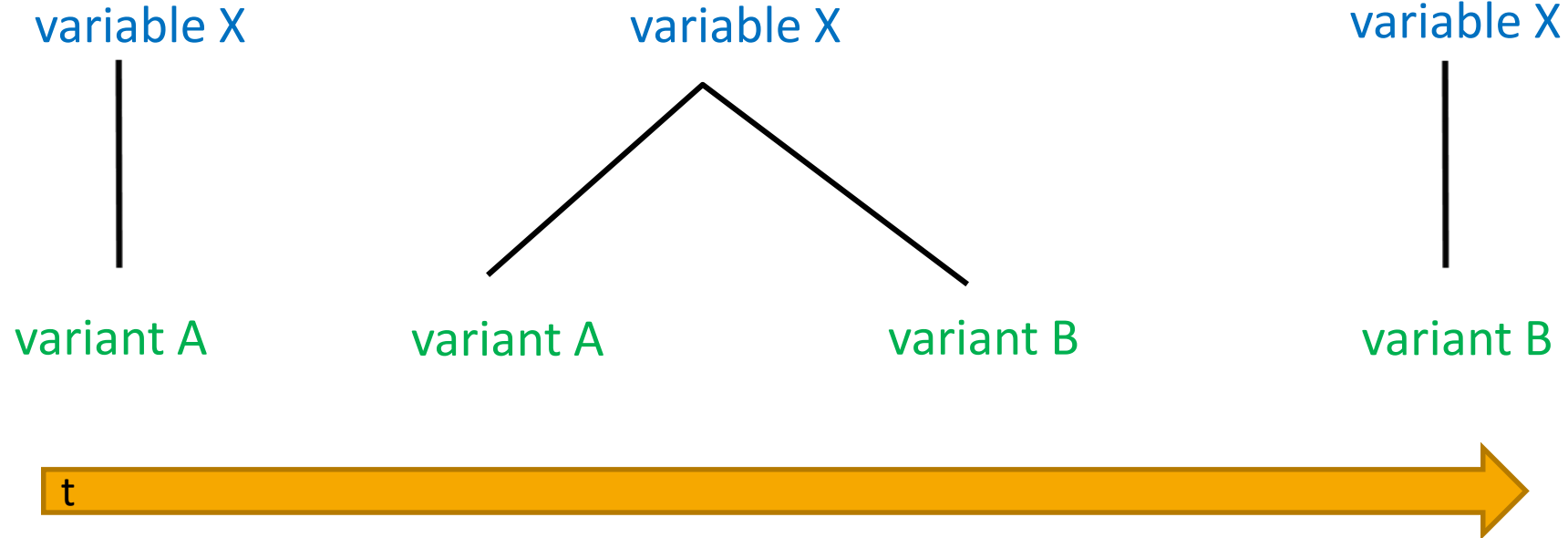
- Openness towards free variation in formal linguistics, e.g.
 - Seiler's (2004) investigation of word order variation in verb-clusters in Swiss German within Optimality Theory:

“I am convinced that completely free variation exists, though it is perhaps not very common.”

(Seiler 2004: 382)

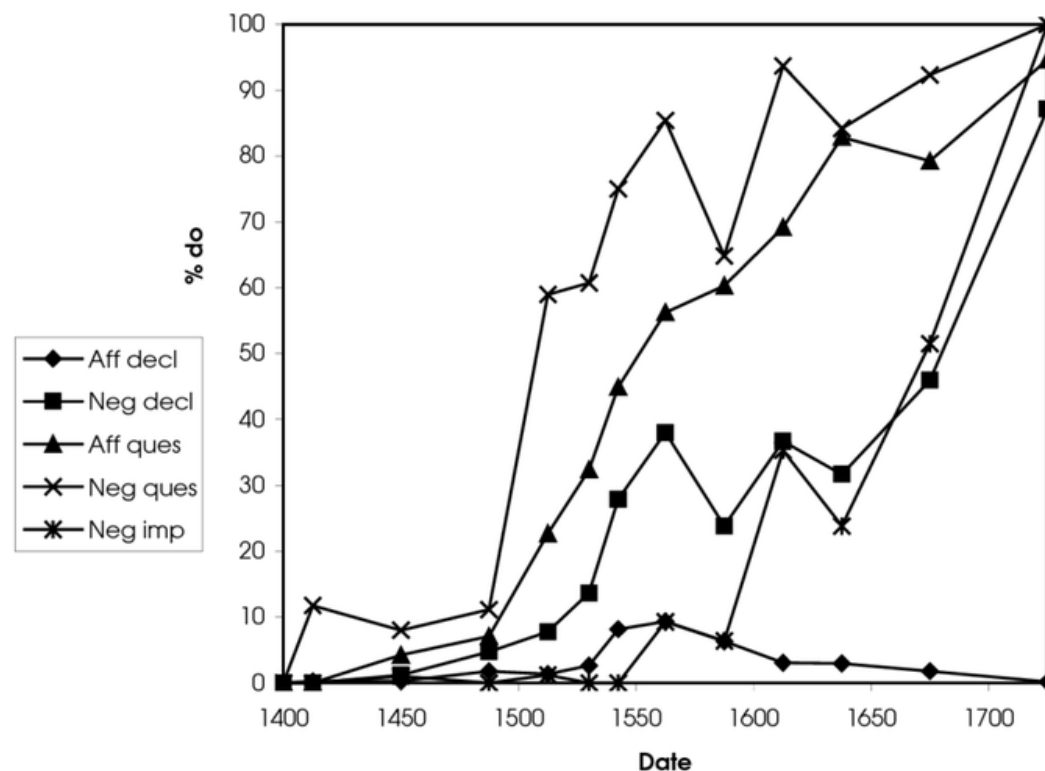
MY TAKE ON FREE VARIATION

A strong argument in favour: Language change



A strong argument in favour: Language change

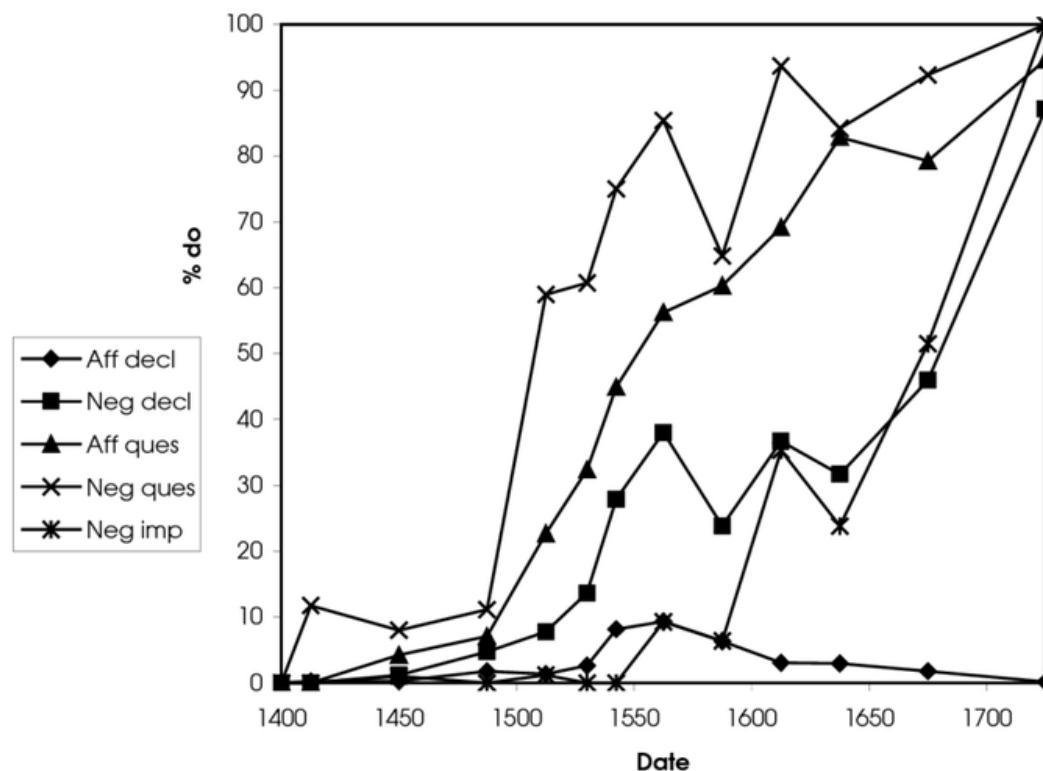
- e.g. Ellegård (1953) on the rise of *do*-support in English
 - Do you know?*
 - I don't know*



Ellegård (1953: 162)

A strong argument in favour: Language change

- At any given point in time, the (non-)use of *do* can be predicted to some extent
- Still, for each particular combination of influencing factors eventually there comes the point where speakers switch
- How would this switch be possible if language were completely determined?



Ellegård (1953: 162)

- From a theoretical perspective, it is reasonable to assume that language is not determined completely, i.e. that **there is free variation** to some degree
- Still, of course, we want to explain as much as possible of the variation that is not free.
 - Methodologically, it may be reasonable to assume (or: pretend) that **there is no such thing as free variation**.

“A certain phenomenon might never be accounted for in your lifetime or mine, but the label ‘free’ does not excuse us from trying. The descriptive linguist is committed to a deterministic philosophy; without determinism, he could never have gotten started, and having put his hand to the plow he can never turn back.”

(Joos 1968[1959]: 185–186)

MY TAKE ON FREE VARIATION

Thank you!

- Adger, David. 2006. Combinatorial Variability. *Journal of Linguistics* 42(3): 505–530.
- Anttila, Arto. 2002. Variation and phonological theory. In *The Handbook of Language Variation and Change*, Jack K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds), 206–243. Malden (MA), Oxford: Blackwell.
- Bolinger, Dwight L. 1977. *Meaning and Form*. New York: Longman.
- Bubenhofer, Noah, Marek Konopka & Roman Schneider. 2014. *Präliminarien einer Korpusgrammatik*. Tübingen: Narr.
- Ellegård, Alvar. 1953. *The auxiliary do: The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell. (Gothenburg Studies in English 2.)
- Ellis, Rod. 1999. Item versus system learning: explaining free variation. *Applied Linguistics* 20(4): 460–480.
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure* [Cognitive Theory of Language and Culture]. Chicago: University of Chicago Press.
- Joos, Martin. 1968[1959]. The isolation of styles. In *Readings in the Sociology of Language*, Joshua A. Fishman (ed), 185–191. Berlin, Boston: De Gruyter Mouton.
- Labov, William. 1972. *Sociolinguistic patterns* [Conduct and Communications Series]. Philadelphia: University of Philadelphia Press.
- Lightfoot, David. 1999. *The Development of Language. Acquisition, Change, and Evolution* [Blackwell/Maryland Lectures in Language and Cognition]. Malden, MA, Oxford: Blackwell.

Newmeyer, Frederick J. 1983. *Grammatical Theory. Its Limits and Its Possibilities*. Chicago, London: The University of Chicago Press.

Preston, Dennis R. 1996. Variationist Perspectives on Second Language Acquisition. In *Second language acquisition and linguistic variation* [Studies in Bilingualism (SiBil)], Dennis R. Preston & Robert Bayley (eds), 1–45. Amsterdam, Philadelphia: John Benjamins.

Schäfer, Roland & Felix Bildhauer. 2012. Building large corpora from the web using a new efficient tool chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asunción Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*, 486–493. Istanbul: European Language Resources Association (ELRA).

Seiler, Guido. 2004. On three types of dialect variation and their implications for linguistic theory. Evidence from verb clusters in Swiss German dialects. In *Dialectology meets Typology: Dialect Grammar from a Cross-Linguistic Perspective* [Trends in Linguistics. Studies and Monographs [TiLSM] 153], Bernd Kortmann (ed), 367–399. Berlin, New York: De Gruyter.

Uhrig, Peter. 2015. Why the Principle of No Synonymy is Overrated. *Zeitschrift für Anglistik und Amerikanistik* 63(3): 323–337.

Weber, Thilo, Felix Bildhauer & Franziska Münzberg. 2024. Finite vs. infinite Attributsätze: zu-dass-Alternation bei Substantiven. In Marek Konopka, Angelika Wöllstein & Ekkehard Felder (eds.), *Substantivflexion, Attributsätze, Präfix- und Partikelverben* (Bausteine einer Korpusgrammatik des Deutschen 3), 59–113. Heidelberg: Heidelberg University Publishing.

Weber, Thilo & Kristin Kopf. 2023. Free variation, unexplained variation? In Kristin Kopf & Thilo Weber (eds.), *Free Variation in Grammar: Empirical and Theoretical Approaches*, 1–20. Amsterdam: John Benjamins.