

Grammar and Complexity

Constructions, complexity and word order variation

Peter W. Culicover
The Ohio State University

A simple idea about complexity

- I started thinking about complexity a few years ago when I had the idea that it might be possible to account for the occurrence of *do*-support in English in terms of a pressure to reduce complexity.
- But then it dawned on me (duh) that most languages don't have *do*-support, so how could complexity be an explanation?

Complexity is complex

- But (I asked myself) if reduction of complexity is not a factor in change, why do languages change, and why is there variation?
- The (short) answer, I think, is that there are **multiple dimensions** on which complexity is measured, and there are **competing pressures**, cognitive and social, that produce change and variation.

Outline

- Some ideas about complexity
- Some questions about word order variation
- Inspirations
- CWG verb clusters
- Variation in a constructional approach
- Variation and social dynamics - where complexity fits in
- Summary

Some ideas about complexity

- Formal complexity
 - A measure of the generality of grammatical ‘rules’
 - There is pressure to change in the direction of less complexity, other things being equal.
 - Complexity can be measured as a property of grammars formulated in terms of CONSTRUCTIONS
 - that is, form-meaning correspondences.

Formal complexity - an example

- if there is a filler in an infinitival relative, it must be a PP.

- (1) a. the man to whom to talk ____
 b. *the man who to talk to ____

- but in an infinitival question, the filler may be an NP.

- (2) a. I wonder to whom to talk ____
 b. I wonder who to talk to ____

- Why?

Infinitival relatives

- Old English **infinitival** relatives lacked relative pronouns entirely (Kemenade 1987:151); they used either zero or the equivalent of *that*.
- When what was relativized was in a PP, the relative was a zero-relative and the preposition was stranded, e.g.,
 (1) Drihten, þu þe gecure þæt fæt [on *t* to eardienne]
 Lord, you yourself chose that vessel in *t* to live
 [Blick 157/ Kemenade 1987:151]
- But finite relatives had relative pronouns.

OE relatives

- The OE tensed relatives did not permit P-stranding with a clause-initial *wh*-; only clause-initial pied-piped relative PP was possible (van Kemenade 1987:152-3).

OE relatives

OE	zero	PP	NP
Tensed			
Infinitival			

- It is therefore plausible to assume that infinitival wh-relatives are an innovation.
- The innovation is in the direction of increasing uniformity, by extending wh-relatives from the tensed to the infinitival cases.
- But note that it gives rise to an idiosyncrasy when interpreted with respect to tensed relatives (which allow all 3 types).

Infinitival relatives

- A likely possibility, given wh-NP in initial position infinitival questions (i.e. *wonder who to talk to*), is that the last cell will be filled in over time.

OE relatives

OE	zero	PP	NP
Tensed			
Infinitival			

- (i) a. where do i find the person who to talk to about the quest?

<http://answers.yahoo.com/question/index?qid=20100727153550AAPs5ed>

- b. In this case, I'll refer to the radio/club DJ as being the person who to target.

<http://independentmusicstartup.com/440/how-to-get-a-tastemaker-to-take-your-music-to-the-next-level/>

- c. All requests for aid should include: ... 3. The name and identity of the requesting person or the *person who to* contact upon arrival.

www.jdcap.org/SiteCollectionDocuments/EmergencyPlanExternal.pdf

- d. if you ae [sic]from out of Auckland and interested I can give you the number and person who to contact.

www.electricalforum.co.nz/index.php?action=more_details...

- e. Please provide a phone number and name of a person who to contact to resolve this.

blog.klm.com/“the-ocean-of-the-air-connects-us-all”/1189/

Some ideas about complexity

- Computational or processing complexity
 - measured indirectly by eye-tracking, reaction times, etc.
 - greater complexity corresponds to lower acceptability, other things being equal.
 - lower acceptability due to processing complexity is distinct from ungrammaticality.
 - greater complexity results in lower frequency, which may produce “surprisal” (\cong unacceptability) and in the limit lead to complete avoidance (easily confused with ungrammaticality).

‘Groundedness’ or ‘embodiment’

- Island constraints (Hofmeister et al; see also Phillips)
- “Freezing” (e.g. extraction from extraposed PP), topic islands, etc.
- and many other configurations where it can be plausibly argued that processing complexity \Rightarrow low frequency \Rightarrow surprisal \Rightarrow unacceptability \sim ungrammaticality
- frequency may vary across social groups, leading to different acceptability of the same configuration

Applying complexity to a real case of word order variation

- The case of Continental West Germanic verb clusters
 - ... that ... the book to-read wants ~ wants to-read
- Goal: devise a theory in which
 - variation is permitted,
 - alternative orders are explained,
 - rare or non-existent orders are explained,

possibly in terms of pressures to reduce complexity.

Some questions about word order variation

- What factors are involved in determining word order, and how do the factors relate to one another?
- What explains the linear order possibilities in a given language?
- What's the right way to talk about word order variation in a given language, across languages, and in time?
- What's the relationship between word order, constituent structure, and conceptual structure?

Main points about word order

1. Word order is specified in and by constructions, which allow for generality as well as lexical idiosyncrasy.
2. A lexically restricted construction may generalize to a broader set of lexical items through contact and learning, reducing complexity.
3. Alternative orders are in principle possible between sisters of a phrase, giving rise to sets of constructions; this is the soil in which variation may sprout and flourish.
4. Social factors (network topology, geography, different frequency distributions over agents, etc.) may favor one constructional alternative over another, leading to dialect variation.
5. Alternative orders may emerge as a way of reducing complexity on different dimensions.
6. Contact leads to spread of properties, resulting in mixed variants.

Inspirations

- Complexity
 - Markedness & optimality (Chomsky, etc.)
 - Processing complexity & dependency (Gibson, Hawkins, Haider)
- Empirical work on verbal clusters and word order in Continental West Germanic [CWG] (Wurmbrand, van Riemsdijk & Haegeman, Zwart, T. Schmid, M. Bader, R. Vogel, Bies, Sapp, many others.)
- Factors bearing on word order (Lötscher, Wasow, Hawkins, Haider)
- Linearization & structure (Curry, Dowty, HPSG, many others)
- Social dynamics & epidemiology (A. Nowak, G. Seiler, N. Enfield)
- Constructional approaches to grammar (many)

CWG verb clusters

- There are ordering variants in CWG 2- and 3-verb clusters. [using Standard German orthography to illustrate variation]

2-verb

Maria glaubt, daß

Maria believes that

(1) a. sie die Arie singen kann. (2-1)

she the aria sing can

‘... she can sing the aria.’

b. sie die Arie kann singen. (1-2)

CWG verb clusters

- There are ordering variants in CWG 2- and 3-verb clusters, as well.

3-verb

- (2) a. sie Peter die Arie singen hören wird. (3-2-1)
 she Peter the aria sing hear will
 ‘...she will hear Peter sing the aria’
- b. sie Peter die Arie hören singen wird (2-3-1) [rare]
- c. sie Peter die Arie wird hören singen (1-2-3)
- d. sie Peter die Arie wird singen hören (1-3-2)
- e. sie Peter die Arie singen wird hören (3-1-2)
- f. sie Peter die Arie hören wird singen (2-1-3) [rare]

- How do we account for
 - the possible orderings in each variety?
 - why some orderings are more frequent than others? (Does complexity have anything to do with it?)
 - how clusters are integrated into grammatical descriptions in terms of syntax and semantics?

- The evidence suggests that the verb clusters are complex phrasal (but not headed) constituents.
 - They cannot be broken up by adverbs, etc.
 - They can be left-dislocated as units.
 - Their interpretation can be given by rule.
- Moreover, derivations of verb sequences as simply spellings out of branching structure are problematic (my assessment of Wurmbrand 2004; 2005).

They cannot be broken up by adverbs - Haider (2003)

Let me emphasize that the ‘no-intervener’ property is clear counterevidence to any analysis of cluster construction that operates with V-projections (or higher ones) as minimal building units of clusters. This applies to Koopman & Szabolcsi (2000) as well as to Wurmbrand (2001). Unfortunately, both publications pay no attention to this crucial property of clusters.

Clusters can't be broken up

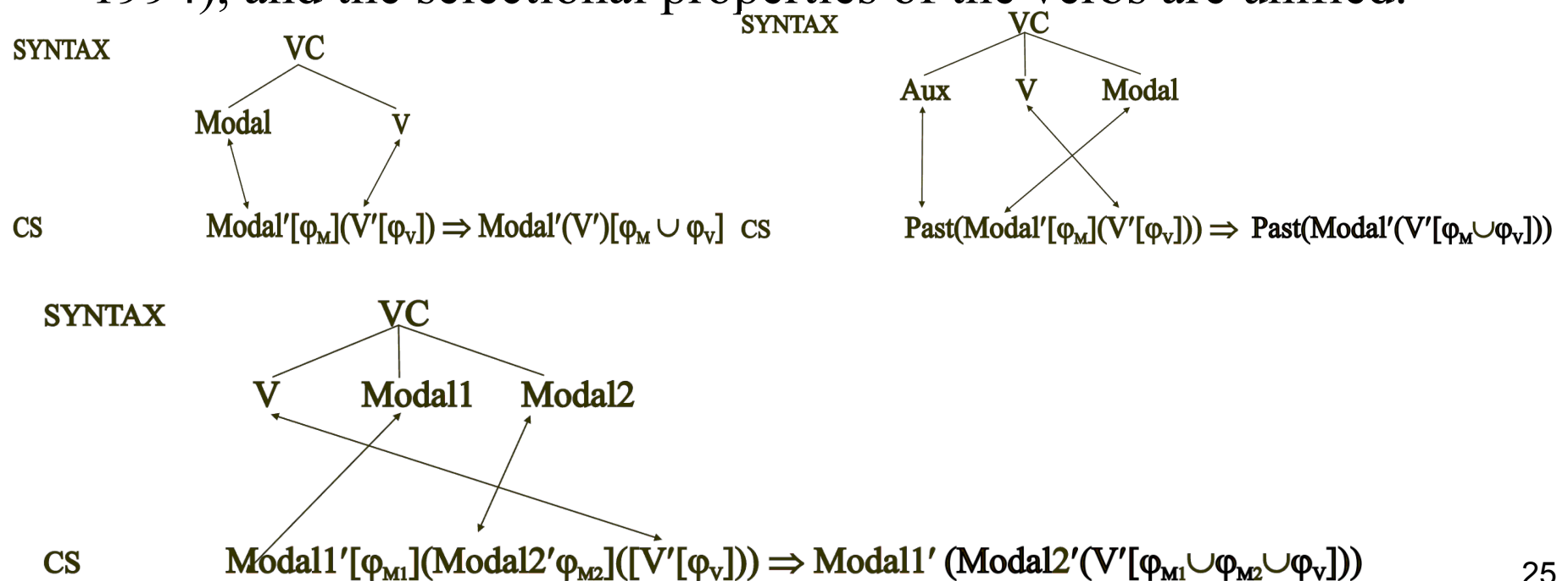
They can be left-dislocated as units - Haider (2003: 95)

- (4)
- a. [Wiedererkennen können]ⁱ (dasⁱ) müßte er sie schon
 identify be-able (this) must he her well
 - b. [Vorsingen lassen müssen]ⁱ (dasⁱ) wird man ihn schon
 audit let must (this) shall one him well
 ‘One will have to let him audit’
 - c. [Übersehen haben]ⁱ (dasⁱ) wird man sie sicher nicht
 overlooked have (this) will one her surely not
 - d. [Übersehen worden]ⁱ (dasⁱ) ist sie noch nie
 overlooked been (this) has she never ever
 - e. [Zu reparieren versucht]ⁱ (dasⁱ) hat man ihn nicht
 to repair tried (this) has one it not

- (note that 2-3 are units, separated from 1)

Interpreting clusters - Bouma and van Noord, 1998

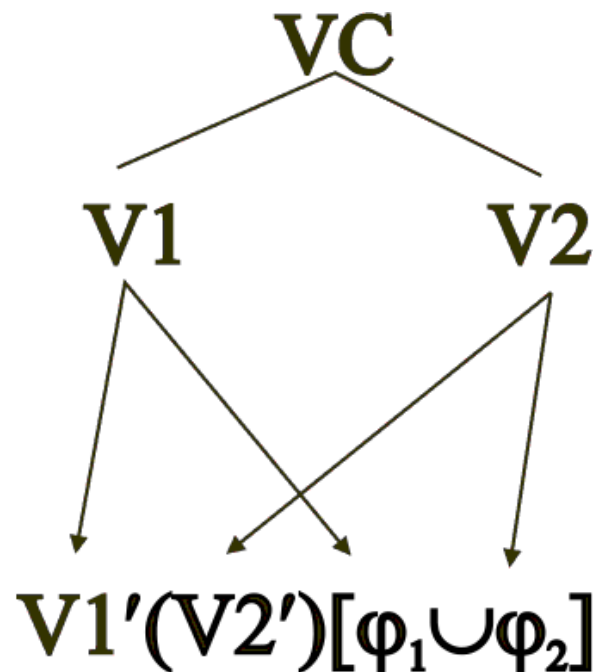
- The semantics do not require hierarchical branching structure.
- There is a single domain for the verb cluster (see also Reape, 1994), and the selectional properties of the verbs are unified.



2-verb clusters

- If the VC is a construction, then it has the following statement for the 2-verb case.

SYNTAX



2-verb clusters

- so, for example, *will*₁ *lesen*₂ acquires the Obj argument of *lesen*,
- and can select *das Buch* in a variety that permits
 (5) ... daß sie das Buch [*will*₁ *lesen*₂].
 that she the book wants to-read
- similarly for *lesen*₂ *will*₁.

Wurmbrand on derivation

- Wurmbrand (2004; 2005) showed that multiple derivational accounts of the possible sequences are possible – that is, we cannot demonstrate that there is a particular base order (e.g. 3-2-1) from which all the other possibilities are derived, because **the full set of surface orders can be derived from any underlying order** (given appropriate stipulations). (Cf. Culicover & Rochemont 1991 regarding HNPS and extraposition.)

A challenge

- There are gaps, asymmetries and correlations— why?
- E.g., in Swiss German dialects (Seiler, 2004), V2-V1 (using Standard German forms: *lesen₂ lassen₁* ‘make read’) is not found unless V-MOD (*lesen können* ‘can read’) is found, and V-MOD is not found unless V-AUX (*gelesen₂ haben₁* ‘have read’) is found.
- At the same time, 3-1-2 (*lesen₃ kann₁ haben₂* ‘read can have’) occurs only if 2-1 is fully general, suggesting a generalization of 1-2 > 2-1 to the 3-verb case (“place main verb first in the sequence”).

Seiler (2004)

“I have shown that the ordering of elements in Western dialects is strictly ascending (1–2–3), but the more we move eastwards the more the tendency for ascending ordering weakens. Second, the ordering of elements is sensitive to the category of the head. Auxiliaries tend most to be set at the right edge of the cluster. **This tendency is much weaker with modal verbs and almost absent with lexical verbs as heads of a cluster.** (emphasis mine - PWC)”

	<div>West ←</div> <div></div> <div>→ East</div>				
	I	II	III	IV	V
AUX,V	1–2	2–1	2–1	2–1	2–1
MOD,V	1–2	1–2	2–1	2–1	2–1
V ₁ ,V ₂	1–2	1–2	1–2	2–1	2–1
AUX,MOD,V	1–2–3	1–2–3	1–2–3	1–2–3	3–1–2

One interpretation of the data

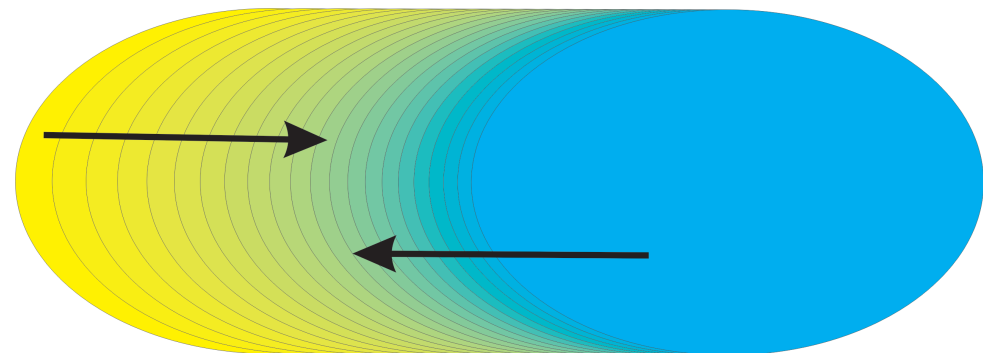
- The 2-1 order develops from the 1-2 order, starts with AUX preceding V, then extends to MOD preceding V, and then V1 (e.g. *lassen* ‘make’) preceding V2.
- Each construction spreads geographically, as well.
- (Such a pattern of change is reminiscent of the spread of *do*-support in EME documented by Ellegård and analyzed by Kroch.)

Simulation

- Time for a computational simulation of the spread of two orders (or any two constructions), each one moving into the area dominated by the other.
- In the simulation with 2 constructions A and B, there are areas of pure A and areas of pure B, and mixed A and B in various degrees.

Variation and social dynamics

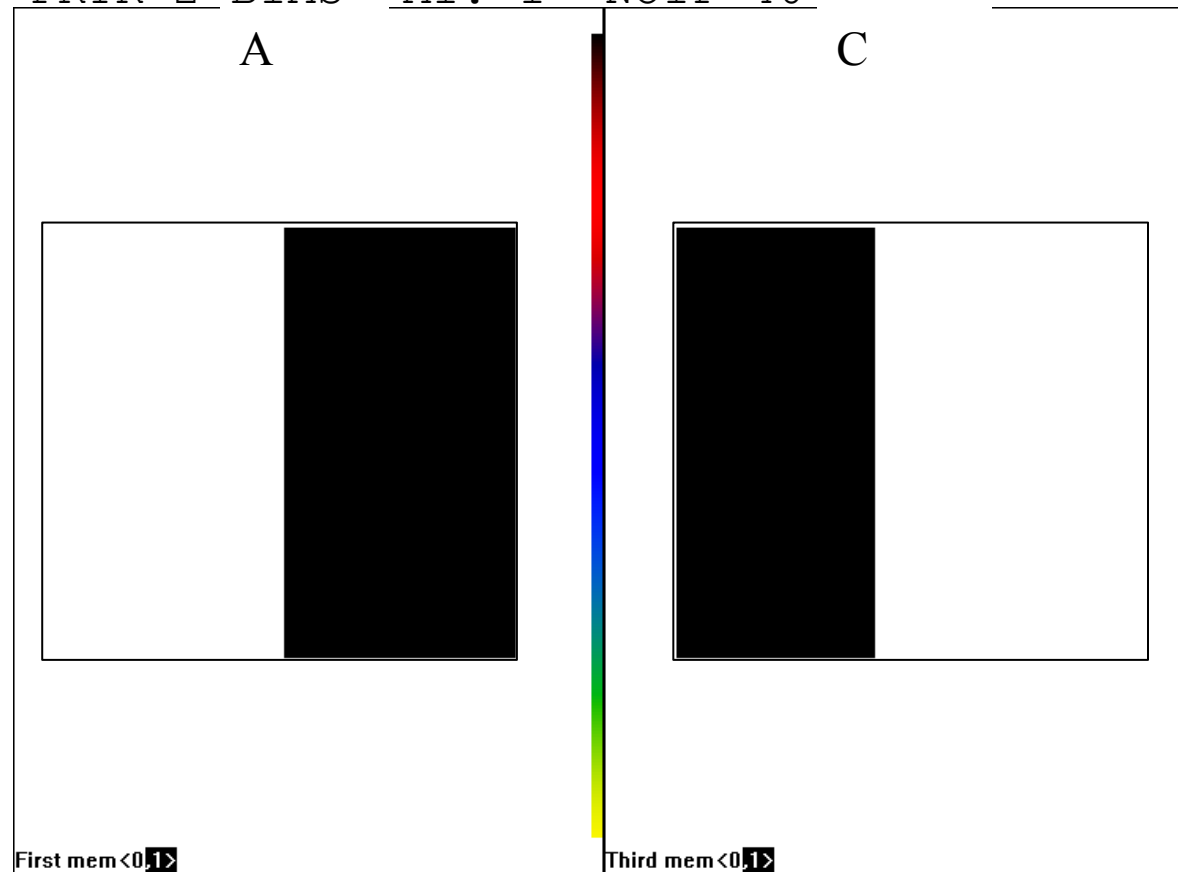
- An interpretation of the data.
 - The various constructions live in the ‘body linguistic’ like viruses.
 - They spread through contact. (cf. Enfield on ‘linguistic epidemiology’; also Seiler (2008) on ‘Plastizität’)



DEMO1

BLUE (A2, C1), GREEN (A1, C2), WHITE (A2, C2), RED (A1, C1)

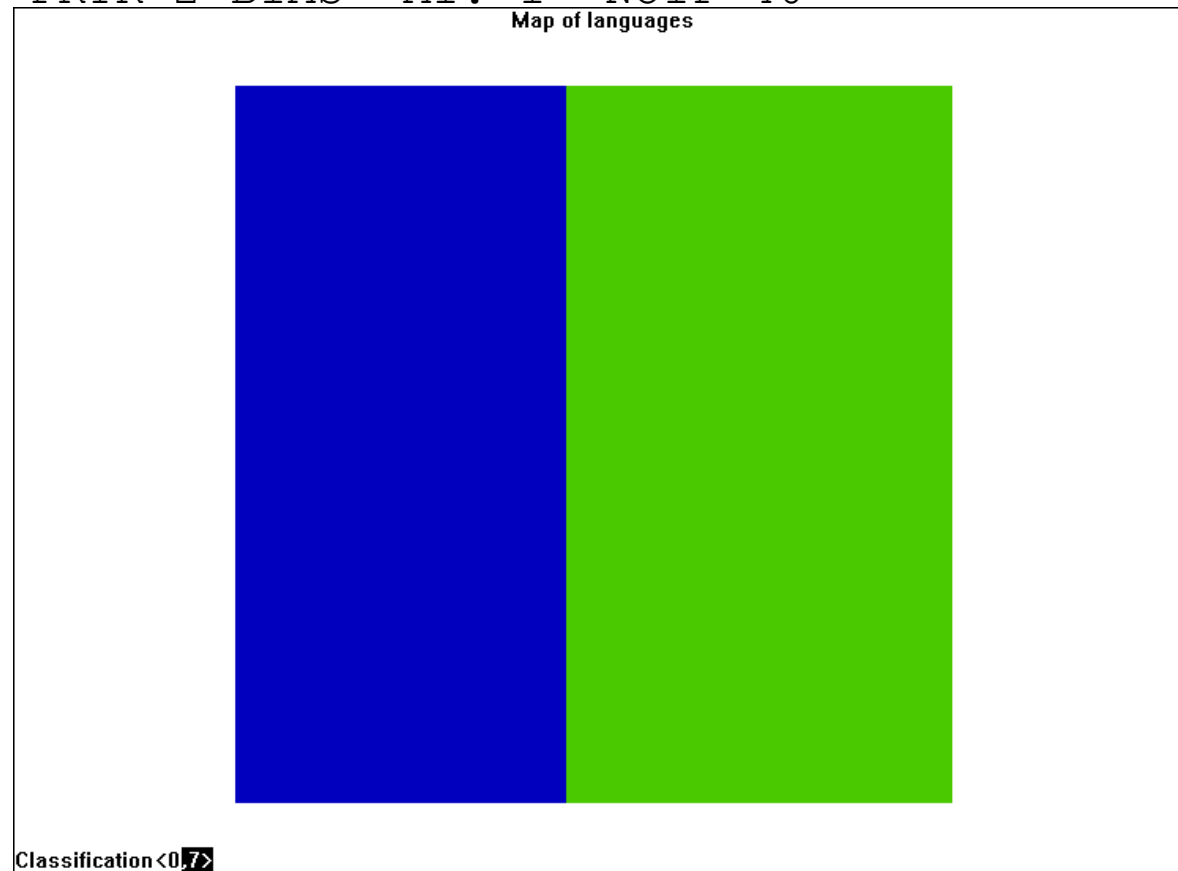
INDI=3 PRTR=2 BIAS="A1:-1" NOIP=40



DEMO1

BLUE (A2, C1), GREEN (A1, C2), WHITE (A2, C2), RED (A1, C1)

INDI=3 PRTR=2 BIAS="A1:-1" NOIP=40

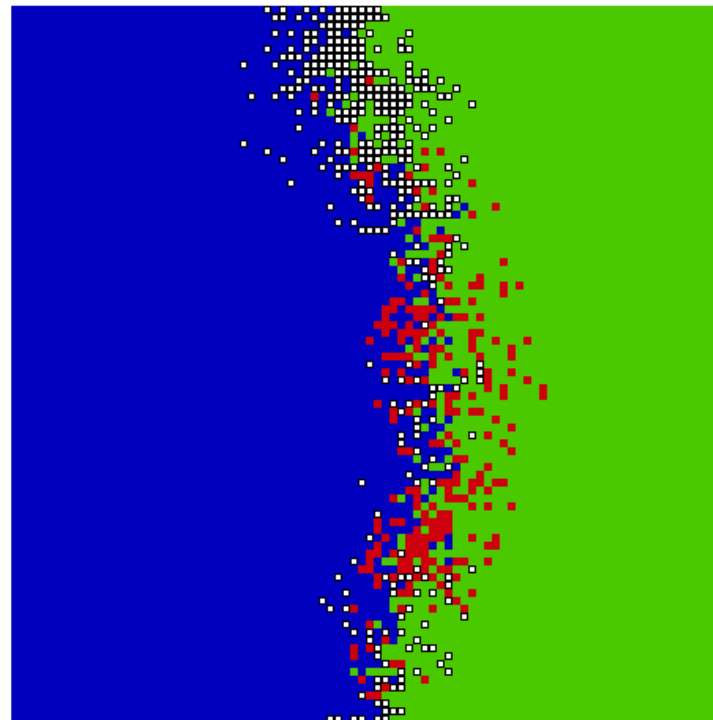


DEMO1

BLUE (A2, C1), GREEN (A1, C2), WHITE (A2, C2), RED (A1, C1)

INDI=3 PRTR=2 BIAS="A1:-1" NOIP=40

Map of languages



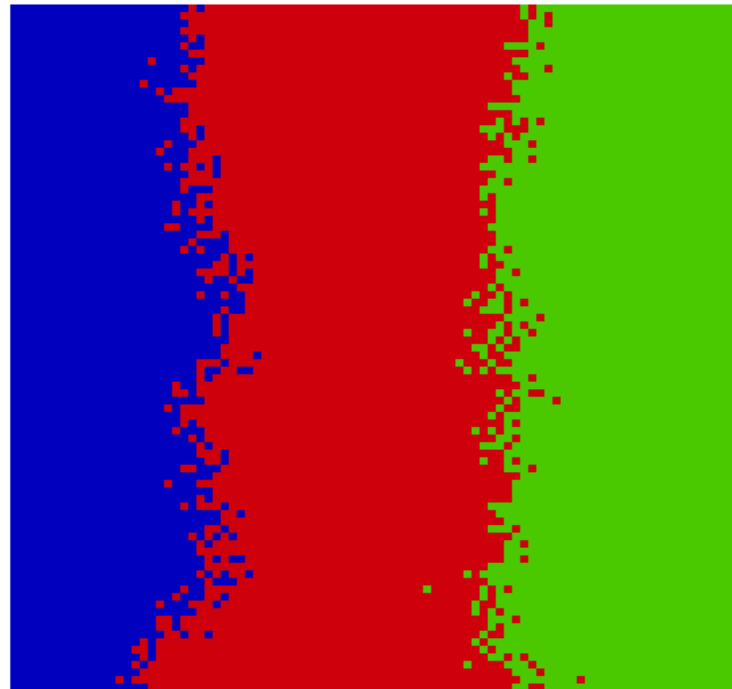
Classification <0.7>

DEMO2

BLUE (A2,C1), GREEN (A1,C2), WHITE (A2,C2), RED (A1,C1)

INDI=3 PRTR=2 BIAS="A1:10 A2:-10 C1:10 C2:-10

Map of languages



Classification<0,7>

Why do both different orders exist in CWG?

- and not just 1-2(-3) (or (3-)2-1)?
- That is, why isn't one order stronger than the others, and in the long run wipes them all out?
- ... as in the following DEMO3.

DEMO3

BLUE (A2, C1), GREEN (A1, C2), WHITE (A2, C2), RED (A1, C1)

INDI=3 PRTR=2 BIAS="A1:25 C1:25"

Map of languages



Classification <0.7>

Constructional change

- Two general reasons (at least) for dominance of a construction B over alternative construction A.
 - *Bias*: Alternative B is less complex (in some sense) than alternative A, and so there is pressure for learners to abandon A in favor of B.
 - *Frequency*: Alternative B occurs more often in the corpus (more speakers?, more uses?) than alternative A, so learners begin to prefer B in favor of A.
 - In such a case, alternative B may actually be more complex than A in some computational sense, but still wins the competition.

Types of Bias

- ‘Bias’ leads learners to prefer B over A, given equal frequency etc.
- B is more general than A (e.g. covers a broader set of lexical items, has fewer exceptions).
- B is easier to process than A, e.g.,
 - puts dependent elements closer together (Hawkins);
 - referents are more accessible;
 - fewer intervening computations (Gibson);
 - more congruent with information structure.
- B is more prestigious/cooler than A
- ...

Why 1-2 (over 2-1)?

- *Scope bias*: Alignment of scope and linear order may facilitate one aspect of the computation of CS representation.

Scope-Order principle

“The preferred scope ordering of operators corresponds to the left-to-right ordering of the phrases in the surface structure of the sentence.” (Familiar intuition, many sources)

Why 1-2 (over 2-1)

- Hence modals, tense and aspect markers, etc. (=1) are preferred **before** their arguments, i.e. VPs (=2).

Why 2-1 (over 1-2)

- *Dependency bias*: 2-1 gets the main V adjacent closer to its arguments & adjuncts, reducing dependency length (cf. e.g. Hawkins)
- Hence arguments and adjuncts (=2) are preferred before V (=1).
- This only works in V-final languages.
 - we would expect 1-2 & 2-1 in such languages,
 - but only 1-2 in V-initial constructions.

Effects of Bias in final clusters

$X\ 1-2-3 \implies X\ \mathbf{3}-1-2$ [dependency bias]

$X\ 3-2-1 \implies X\ \mathbf{1}-3-2$ [scope bias]

$X\ 1-2-3 \implies X\ \mathbf{2}-3-1$ [weak dependency bias,
assumes unitary 2-3; cf.
Haider 2003]

- 2-1-3 should be rare, expresses contrast/
givenness bias on 2 (appears to exist in ZüGe
just for particular lexical V2)

Swiss German 2-1-3

(7) s aagfange₂ hat₁ rägne₃, ... [ZüGe.]

it begin.INF has rain.INF

(Lötscher 1978)

(8) dass i en gsee₂ ha₁ schaffe₃

that I him seen have.1S work.INF

‘that I saw him work’

(M. Salzmann, p.c.)

2-1-3!

- Notice that in the case of (8) ('him seen₂ have₁ work₃'), at least, 2-1-3 puts the verb closer to its overt argument, satisfying the dependency bias.
- A similar case can be made for (7) ('it start₂ have₁ rain₃'), if 'it' is an argument.
- 2-1-3 apparently occurs only with such verbs.

2-1-3!

- Notice that in the case of (8) ('him **seen**₂ have₁ work₃'), at least, 2-1-3 puts the verb closer to its overt argument, satisfying the dependency bias.
- A similar case can be made for (7) ('it **start**₂ have₁ rain₃'), if 'it' is an argument.
- 2-1-3 apparently occurs only with such verbs.

2-1-3!

- Notice that in the case of (8) (**him** **seen**₂ have₁ work₃'), at least, 2-1-3 puts the verb closer to its overt argument, satisfying the dependency bias.
- A similar case can be made for (7) (**it** **start**₂ have₁ rain₃'), if 'it' is an argument.
- 2-1-3 apparently occurs only with such verbs.

Correlations

- If a variety permits 1-2 or 2-1 for a particular pair of verbs, then it should permit 1-(2-3) and (3-2)-1.
- Crucially, we do not expect to find 3-2-1 in the absence of 2-1, or 1-2-3 in the absence of 1-2 (and we don't, as far as I can tell).

Correlations

- In West Flemish, 2-1 is required for $V1=AUX$ and 1-2 is required for $V1=MOD$ (Haegeman, 1994):
V-AUX & MOD-V
- So we expect 2-3-1 where $V1=AUX$, $V2=MOD$, since V follows MOD and AUX is maximally final.

(3) ...da Valère nie nor us will-en₂ kom-en₃ eet₁
that Valery not to house want-INF come-INF has
'..that Valery did not want to come home.'

[Haegeman 1994]

Correlations

- 2-3-1 clusters with a range of verbs for V2 are well-documented in Afrikaans (Biberauer n.d.) and quite general; e.g.,

(4) ... dat dit ophou₂ reën₃ het₁ [Afrikaans]
 that it stop.INF rain.INF have
 ‘... that it has stopped raining’
 [Biberauer n.d.]

Correlations

- Moreover, 2-3-1 is the preferred alternative to 1-2-3, which suggests that it occurs when 2-3 is a construction licensed by domain union – that is, 2-3 precedes 1 due to the Dependency Bias.
- However, the extent of 2-3-1 varies: in Afrikaans it appears to be quite general (occurs for all V2 but restricted to V1='have'), while in Dutch and West Flemish it is more restricted.
- On the other hand, 2-1-3 does not satisfy any linear order bias, and appears to be maximally disharmonic as well (Culicover *et al.*, 2003).

Summing up: Ideas about determinants of word order

- Word order is expressed in terms of constructions, and is restricted by
 - complexity (of the syntax-CS correspondences)
 - information structure, and
 - style/noise (= all other factors; cf. Wasow, Lötscher)

*‘Dabei muß die zusätzliche Komplikation berücksichtigt werden, daß für die Erklärung der Wortstellung wohl kaum ein einziger Regeltyp vorausgesetzt werden kann. Vielmehr lassen sich mindestens drei interagierende, aber primär voneinander unabhängige Arten von Regeln ansetzen: Erstens **grammatisch bedingte Regeln** [...] die mehr oder minder willkürlich eine Abfolge bestimmen [...]; **performanzbedingte Regeln** [...]; endlich **funktional bedingte Regeln** [...] deren Zweck die Ermöglichung von bestimmten funktionalen Satzverhältnissen im Sinne der Thema-Rhema Unterscheidung ist.’* ((Lötscher, 1978, 11); boldfacing by TS)

‘The additional complication, that one single rule type is hardly sufficient to account for word order, must be taken into account as well. Rather, there are at least three interacting but primarily independent kind of rules: First, **grammatical rules** [...] that determine an order more or less arbitrarily [...]; **performance rules** [...]; at last, **functional rules** [...] that allow for certain functional relations in a sentence in the sense of the topic-comment distinction.’ (translation by TS)

Wasow 2002

- Wasow (2002) shows that word order variation in the English VP is sensitive to a number of factors, including grammatical weight, discourse newness, and collocation frequency.

Summary

- Constructions are the right level of representation for the purpose of defining certain types of complexity.
- Constructions spread through the individual grammar and through the 'body linguistic', generalizing as they spread.
- Different constructions may be responses to pressures to reduce complexity on different dimensions.
- Constructions compete with one another in the social network.
- The structure of the network (and other factors, such as frequency) may result in the persistence of complexity even when there are cognitive pressures to reduce complexity.

Acknowledgements

- The Alexander von Humboldt Foundation
- Erhard Hinrichs & Marga Reis
- Ray Jackendoff
- Susanne Winkler
- Guido Seiler
- Martin Salzmann
- audiences at OSU, Humboldt University, University of Tübingen

- Thank you

References

- Bader, M., Schmid, T. 2009. Minimality in verb-cluster formation. *Lingua* 119, 1458-81.
- Bader, M., Schmid, T., Häussler, J. 2009. Optionality in verb-cluster formation. *The Fruits of Empirical Linguistics: Product 37*.
- Bader, M., Schmid, T. 2009. Verb clusters in colloquial German. *The Journal of Comparative Germanic Linguistics* 12, 175-228.
- Bader, M., Schmid, T. 2009. CAT meets GO: 2-verb clusters in German. In Van Craenenbroeck, J. (Ed), *Mouton de Gruyter*, Berlin. 203-44.
- Biberauer, T. Disharmonic word order, quirky morphology and the Afrikaans verb cluster. Unpublished ms. <http://www.hum.au.dk/engelsk/engsv/nyvad-abstracts/biberauer-ho.pdf>.
- Bies, A. E. 1996. *Syntax and Discourse Factors in Early New High German: Evidence for Verb-Final Word Order*. Unpublished doctoral dissertation. University of Pennsylvania,
- Bouma, G., Van Noord, G. 1998. Word order constraints on Germanic verb clusters. In Hinrichs, E., Nakazawa, T., Kathol, A. (Eds), *Complex Predicates in Nonderivational Syntax*. Academic Press, 43-71.
- Chomsky, N. 1964. *Current Issues in Linguistic Theory*. Mouton, The Hague.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.
- Chomsky, N. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.
- Culicover, P. W., Nowak, A., Borkowski, W. 2003. Linguistic theory, explanation and the dynamics of language. In Moore, J., Polinsky, M. (Eds), *The Nature of Explanation in Linguistic Theory*. CSLI Publications, Stanford, CA.
- Curry, H. B. 1961. Some logical aspects of grammatical structure. In Jakobson, R. (Ed), *Structure of Language and its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics*. American Mathematical Society, 56-68.
- Dowty, D. R. 1996. Towards a minimalist theory of syntactic structure. In Bunt, H., Van Horeck, A. (Eds), *Discontinuous Constituency*. Mouton de Gruyter, Berlin.

- Ellegaard, A. 1953. *The Auxiliary Do: The Establishment and Regulation of Its Use in English*. Almqvist and Wiksell., Stockholm.
- Enfield, N. J. 2008. Transmission biases in linguistic epidemiology. *Journal of Language Contact* 2, 299-310.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68, 1-76.
- Haegeman, L. 1994. Verb raising as verb projection raising: Some empirical problems. *Linguistic Inquiry* 25, 509-22.
- Haegeman, L., Van Riemsdijk, H. 1986. Verb projection raising, scope and the typology of rules affecting verbs. *Linguistic Inquiry* 17, 417-66.
- Haider, H. 2003. V-Clustering and clause union. In Seuren, P. A. M., Kempen, G. (Eds), *Verb constructions in German and Dutch*. John Benjamins Publishing Company, Amsterdam. 91-126.
- Hawkins, J. A. 2004. *Complexity and Efficiency in Grammars*. Oxford University Press, Oxford.
- Hawkins, J. A. 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press, Cambridge.
- Koopman, H., Szabolcsi, A. 2000. Verbal Complexes.
- Kroch, A. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1, 199-244.
- Lötscher, A. 1978. Zur verbstellung im zürichdeutschen und in anderen varianten des deutschen. *Zeitschrift für Dialektologie und Linguistik* 45, 1-29.
- Lötscher, A. 1993. Zur Genese der Verbverdopplungsverben bei gaa, choo, laa, aafaa), im Schweizerdeutschen. In Abraham, W., Bayer, J. (Eds), *Dialektsyntax*. Westdeutscher Verlag, Opladen. 180-200.
- Nowak, A., Szamrej, J., Latane', B. 1990. From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review* 97, 362-76.
- Pollard, C., Sag, I. A. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.
- Reape, M. 1994. Domain union and word order variation in German. In Nerbonne, J., Netter, K., Pollard, C. J. (Eds), *German in Head-Driven Phrase Structure Grammar*. CSLI Publications, Stanford. 151-98.

- Sapp, C. 2006. Verb Order in Subordinate Clauses From Early New High German to Modern German. Unpublished doctoral dissertation. faculty of the University Graduate School in partial fulfillment of the requirements for the degree Doctor of Philosophy in the Department of Germanic Studies, Indiana University,
- Sapp, C. 2007. Ov and vo in early new high german: Discourse and other factors. International Workshop on Information Structure and Word Order Variation in Older Germanic, HU Berlin,
- Sapp, C. 2011. Verb Order in Subordinate Clauses From Early New High German to Modern German. John Benjamins, Amsterdam.
- Sapp, C. D. 2005. Factors Favoring Aux-V Order in 15th-century German Subordinate Clauses. In Roehrs, D., Kim, O.-H., Kitagawa, Y. (Eds), Syntax and beyond: Indiana University Working Papers in Linguistics, vol. 5. Indiana University Linguistics Club Publications, Bloomington. 149-69.
- Sapp, C. D. 2007. Focus and verb order in Early New High German: Historical and contemporary evidence. In Featherston, S., Sternefeld, W. (Eds), Roots: Linguistics in search of its evidential base. Walter de Gruyter., Berlin.
- Schmid, T. 2005. Infinitival Syntax. Infinitivus Pro Participio as a Repair Strategy. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Schmid, T., Vogel, R. 2004. Dialectal variation in german 3-verb clusters: A surface-oriented optimality theoretic account. The Journal of Comparative Germanic Linguistics 7, 235-74.
- Seiler, G. 2004. On three types of dialect variation and their implications for linguistic theory. Evidence from verb clusters in Swiss German dialects. In Kortmann, B. (Ed), Dialectology meets typology: dialect grammar from a cross-linguistic perspective. Mouton de Gruyter, Berlin / New York. 367-99.
- Seiler, G. 2008. Syntaxgeographie und Plastizität der Grammatik. In Donhauser, K., Glaser, E., Vuillaume, M. (Eds), Empirische Grundlagen moderner Grammatikforschung: Akten des XI. Internationalen Germanistenkongresses Paris 2005, Vol 4. Peter Lang, Bern. 49-58.

- Vogel, R. 2003. Dialectal variation in german 3-verb clusters: Looking for the best analysis. *Head Movement and Syntactic Theory* 10, 199–235.
- Vogel, R. 2009. Skandal im verbkomplex. *Zeitschrift fuer Sprachwissenschaft* 28, 307-46.
- Wasow, T. 1997. End-weight from the speaker's perspective. *Journal of Psycholinguistic Research* 26, 347-61.
- Wasow, T. 1997. Remarks on grammatical weight. *Language variation and change* 9, 81-106.
- Wasow, T. 2002. *Postverbal Behavior*. CSLI, Stanford.
- Wurmbrand, S. 2004. West germanic verb clusters: The empirical domain. *Verb clusters: A study of Hungarian, German, and Dutch* 43-85.
- Wurmbrand, S. 2005. Verb clusters, verb raising, and restructuring. In Everaert, M., Van Riemsdijk, H. (Eds), *The syntax companion*. Blackwell, Oxford. 227-341.
- Zwart, C. J.-W. 1995. A note on verb clusters in the Stellingwerf dialect. In Den Dikken, M., Hengeveld, K. (Eds), *Linguistics in the Netherlands 1995*. John Benjamins, Amsterdam/Philadelphia. 215-26.
- Zwart, C. J.-W. 1996. Verb clusters in continental west germanic dialects. *Microparametric Syntax and Dialect Variation*, 229-58.