Do we need probabilistic grammars for language change?

Richard Zimmermann

Université de Genève
VariaForMea Winter School

December 4, 2013

Outline

Grammatical variables

- Definition 'grammatical variable'

Grammatical variables

- Definition 'grammatical variable'
  1. a point along which linguistic output may vary

Grammatical variables

- Definition 'grammatical variable'
  1. a point along which linguistic output may vary
  2. has variants, i.e. realizations of the grammatical variable

Grammatical variables

- Definition 'grammatical variable'
  1. a point along which linguistic output may vary
  2. has variants, i.e. realizations of the grammatical variable
  3. variants may be primitive or in direct contrast with some other variant (cf. parameters)

Grammatical variables

- Definition 'grammatical variable'
  1. a point along which linguistic output may vary
  2. has variants, i.e. realizations of the grammatical variable
  3. variants may be primitive or in direct contrast with some other variant (cf. parameters)
  4. cf. phonemes (variable) - allophones (variants);
     alleles (variables) - genes (variants)

Grammatical variables

- Definition 'grammatical variable'
  1. a point along which linguistic output may vary
  2. has variants, i.e. realizations of the grammatical variable
  3. variants may be primitive or in direct contrast with some other variant (cf. parameters)
  4. cf. phonemes (variable) - allophones (variants); alleles (variables) - genes (variants)
  5. purely descriptive device

Example of a grammatical variable

Verb clusters with three verbs in German: the variable 'verb position' has two variants: the finite verb can appear before or after the VP:

(1) a. ... dass er ein Buch lesen wollen wird
... that he a book read want will
'... that he will want to read a book'

b. ... dass er ein Buch wird lesen wollen

Contextual factors

- Definition 'contextual factor'

Contextual factors

- Definition 'contextual factor'
  1. a controlled, independent variable that influences the realization of a grammatical variable

Contextual factors

- Definition 'contextual factor'
  1. a controlled, independent variable that influences the realization of a grammatical variable
  2. cf. contexts that determine the complementary distribution of allophones

Contextual factors

- Definition 'contextual factor'
  1. a controlled, independent variable that influences the realization of a grammatical variable
  2. cf. contexts that determine the complementary distribution of allophones
  3. can be categorical / hard or probabilistic / soft

Contextual factors

- Definition 'contextual factor'
  1. a controlled, independent variable that influences the realization of a grammatical variable
  2. cf. contexts that determine the complementary distribution of allophones
  3. can be categorical / hard or probabilistic / soft
  4. can be virtually anything: syntactic, semantic, phonological features, sociological characteristics, genre ...

Contextual factors

- Definition 'contextual factor'
  1. a controlled, independent variable that influences the realization of a grammatical variable
  2. cf. contexts that determine the complementary distribution of allophones
  3. can be categorical / hard or probabilistic / soft
  4. can be virtually anything: syntactic, semantic, phonological features, sociological characteristics, genre ...
  5. purely descriptive device

Examples of hard and soft factors

The contextual factor 'phonological context' influences the grammatical variable 'occurence of consonants' in Dutch:

(2) *velar nasal before schwa and full vowel in Dutch*

a. ^{ok}[eŋəl] /before schwa

b. *[taŋo] /before full vowel

(3) *other consonants before schwa and full vowel in Dutch*

a. preferred: [kɛpə] /before schwa

b. dispreferred: [kɛpo] /before full vowel

(Oostdorp 2011)

Context dependency

Definition

By 'context dependency' I mean the observation that grammatical variation can be constrained by unrelated factors, irrespective of what such factors may be, and irrespective of whether these factors determine the grammatical variable categorically or influence them statistically.

Two general explanations for contextual factors

structural reasons:

↓

should be part of a model of linguistic knowledge

non-structural reasons:

- processing ease
- discourse optimization
- general cognitive principles
- third factors
- sociology

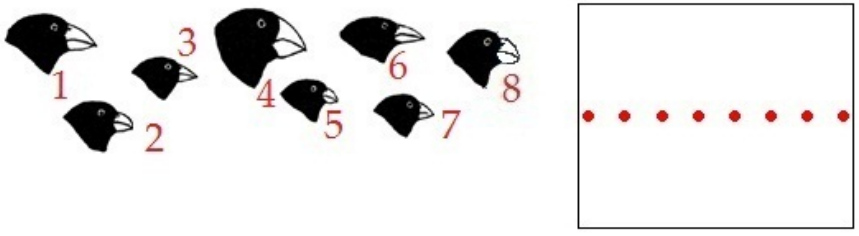
↓

should not be part of a model of linguistic knowledge

A biological metaphor...

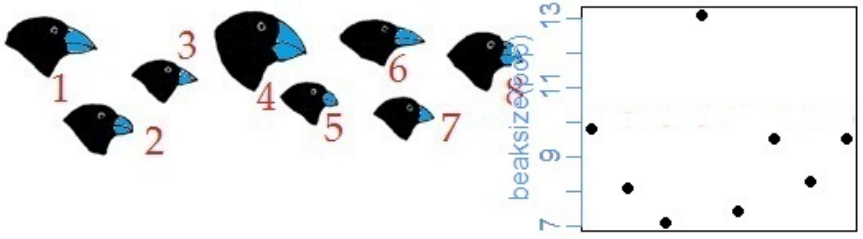
- a population of birds; no two birds are identical

A biological metaphor...



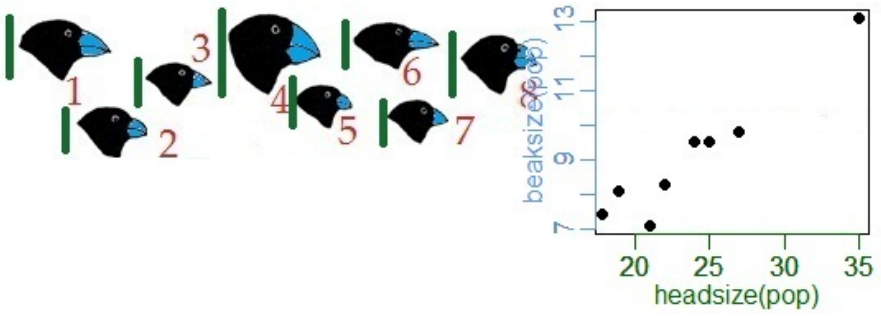
- a population of birds; no two birds are identical
- population = {1,2,... 8}

A biological metaphor...



- a population of birds; no two birds are identical
- population = {1,2,... 8}
- beaksize(pop) = {9.8mm, 8.1mm, ... 8.7mm}

A biological metaphor...



- a population of birds; no two birds are identical
- population = {1,2,... 8}
- beaksize(pop) = {9.8mm, 8.1mm, ... 8.7mm}
- headsize(pop) = {27mm, 19mm, ..29mm}

...for structural variation

- biology:

...for structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary

- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary

...for structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary
  2. head size correlates with (explains?) beak size

- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary
  2. a contextual factor constrains (explains?) the realization of a grammatical variable

...for structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary
  2. head size correlates with (explains?) beak size
  3. head size is an *inherent* property of bird physiology

- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary
  2. a contextual factor constrains (explains?) the realization of a grammatical variable
  3. contextual factor may be an *inherent* property of speakers' linguistic knowledge

...for structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary
  2. head size correlates with (explains?) beak size
  3. head size is an *inherent* property of bird physiology
  4. head size should be represented in our model of bird physiology
- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary
  2. a contextual factor constrains (explains?) the realization of a grammatical variable
  3. contextual factor may be an *inherent* property of speakers' linguistic knowledge
  4. such a contextual factor should be represented in our model of competence

Example 1: German verb position by 'clause type'

- (4) a. ... dass Peter Maria ein Buch gibt.
  ... that P. M. a book gives
  '...that Peter gives Mary a book'

  b. *... dass Peter gibt Maria ein Buch.

  c. Peter gibt Maria ein Buch
  P. gives M. a book
  'Peter gives Mary a book'

  d. *Peter Maria ein Buch gibt.

- Speakers of German probably *know* that verb position is constrained by the contextual factor "clause type", [+/- main clause].

Example 2: French indirect pronouns by 'verb semantics'

- (5) a. je pense à lui
  I think of him
  'I think of him'
  b. *je lui pense.

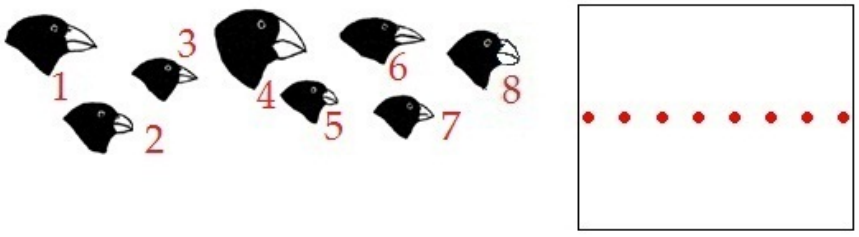
  c. je lui parle
  I him talk
  'I talk to him'
  d. *je parle à lui.

- Speakers of French probably *know* that cliticization is constrained by the contextual factor "verb semantics", [+/- verb of thinking].

Another biological metaphor...

- a population of birds; no two birds are identical

Another biological metaphor...

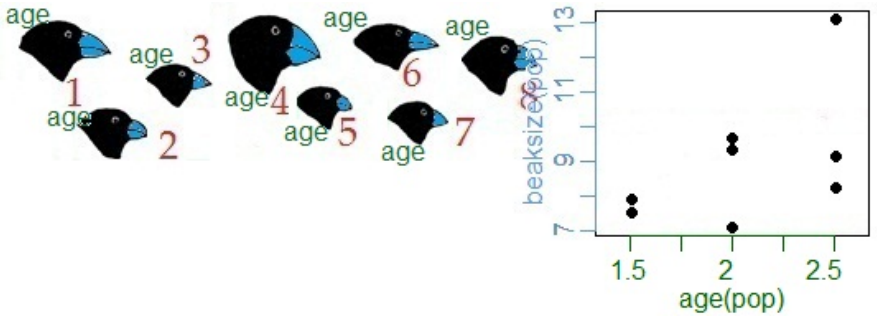


- a population of birds; no two birds are identical
- population = {1,2,... 8}

Another biological metaphor...

- a population of birds; no two birds are identical
- population = {1,2,... 8}
- beaksize(pop) = {9.8mm, 8.1mm, ... 8.7mm}

Another biological metaphor...



- a population of birds; no two birds are identical
- population = {1,2,... 8}
- beaksize(pop) = {9.8mm, 8.1mm, ... 8.7mm}
- independent variables like: nourishment, weather, habitat ...

  age(pop) = {2.5years, 1.5years, ..2years}

...for non-structural variation

- biology:

...for non-structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary

- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary

...for non-structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary
  2. age correlates with (explains?) beak size

- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary
  2. a contextual factor constrains (explains?) the realization of a grammatical variable

...for non-structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary
  2. age correlates with (explains?) beak size
  3. age is *not* an inherent property of bird physiology

- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary
  2. a contextual factor constrains (explains?) the realization of a grammatical variable
  3. contextual factor may *not* be an inherent property of speakers' linguistic knowledge

...for non-structural variation

- biology:
  1. beak size is a point of comparison along which birds may vary
  2. age correlates with (explains?) beak size
  3. age is *not* an inherent property of bird physiology
  4. age should not be represented in our model of bird physiology
- linguistics:
  1. a grammatical variable is a point of comparsion along which speakers' linguistic knowledge may vary
  2. a contextual factor constrains (explains?) the realization of a grammatical variable
  3. contextual factor may *not* be an inherent property of speakers' linguistic knowledge
  4. such a contextual factor should not be represented in our model of competence

Example 3: Topic-drop by 'register'

- (6) a. *_ Must try harder. (narrative, discourse)
  b. _ Must try harder. (school report)
  (Haegeman 1990)
- English speakers probably do not *know* that Topic-drop is constrained by the contextual factor "register", [+/- school report].

Example 4: Postposition by 'phonological weight'

- (7) a. The astronomer [gazed [into the dark sky] [through his effective but unwieldy telescope]]
  (27 examples in Hawkins' sample of 117 clauses with instrument and direction PP)

  b. ?The astronomer [gazed [through his effective but unwieldy telescope] [into the dark sky]]
  (2 examples in Hawkin's sample of 117 clauses with instrument and direction PP)

  (Hawkins 2000)
- English speakers probably do not *know* that postposition is constrained by the contextual factor "phonological weight", [+/- heavy].

Distinguishing structural and non-structural variation

- In actual practice, it is difficult to distinguish structural and non-structural reasons for contextual factors
- Generally, researchers proceed like this:

Distinguishing structural and non-structural variation

- In actual practice, it is difficult to distinguish structural and non-structural reasons for contextual factors
- Generally, researchers proceed like this:

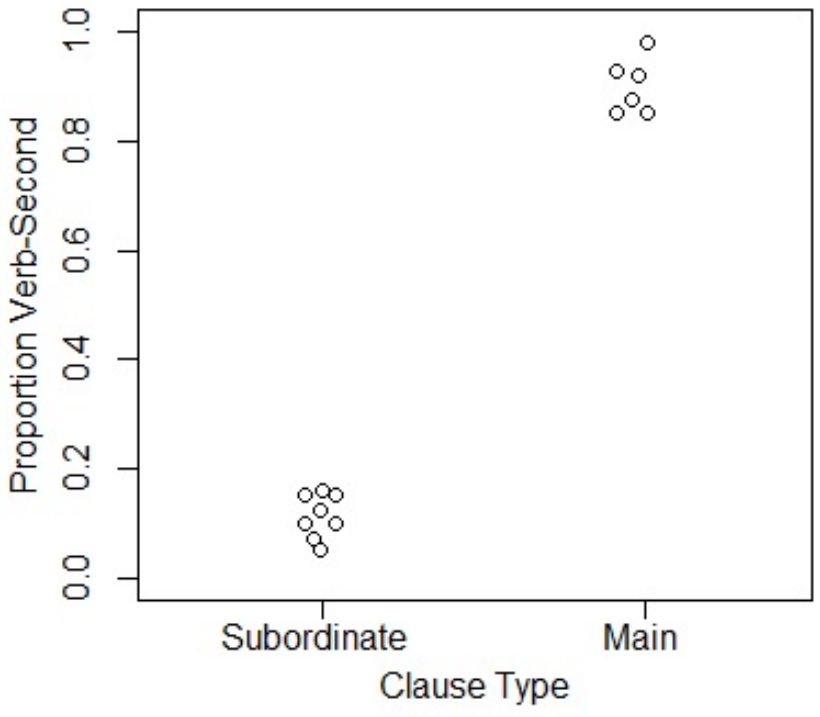
structural variation	non-structural variation
▸ a plausible formal algorithm can be devised that takes the contextual factor into consideration and generates different structures (e.g. V2 in German)	▸ for functional explanations, a general principle can be evoked that works outside of competence (e.g. John Hawkins' Minimize Domains Principle)

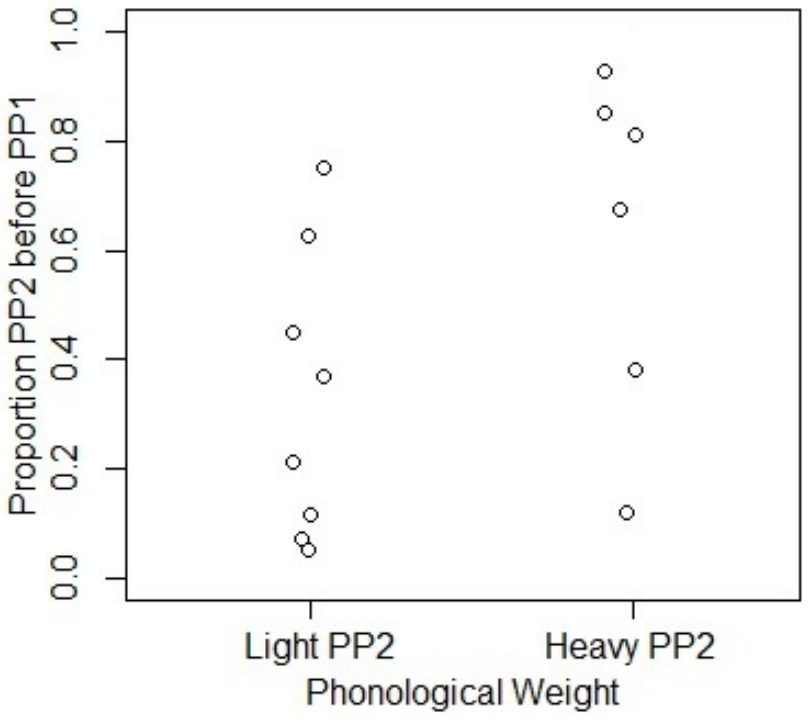
Distinguishing structural and non-structural variation

- In actual practice, it is difficult to distinguish structural and non-structural reasons for contextual factors
- Generally, researchers proceed like this:

structural variation	non-structural variation
▸ a plausible formal algorithm can be devised that takes the contextual factor into consideration and generates different structures (e.g. V2 in German)	▸ for functional explanations, a general principle can be evoked that works outside of competence (e.g. John Hawkins' Minimize Domains Principle)
▸ categorical / hard factor	▸ probabilistic / soft factor

Hard vs. soft factors





Distinguishing structural and non-structural variation

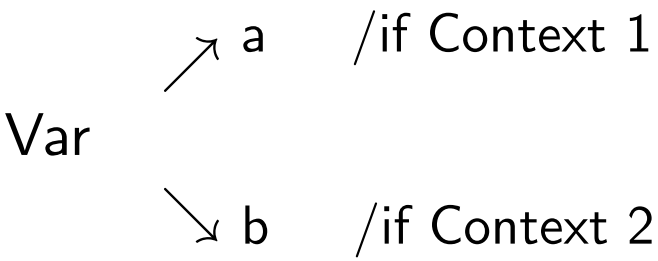
- In actual practice, it is difficult to distinguish structural and non-structural reasons for contextual factors
- Generally, researchers proceed like this:

structural variation	non-structural variation
▸ a plausible formal algorithm can be devised that takes the contextual factor into consideration and generates different structures (e.g. V2 in German)	▸ for functional explanations, a general principle can be evoked that works outside of competence (e.g. John Hawkins' Minimize Domains Principle)
▸ categorical / hard factor	▸ probabilistic / soft factor

A though nut...

- Hard contextual factors are often encoded in a formal model of competence, i.e. they are seen as structural.

Hard contextual factor - general description



Hard contextual factor - example

subject - auxiliary inversion

↗ SAI /if question

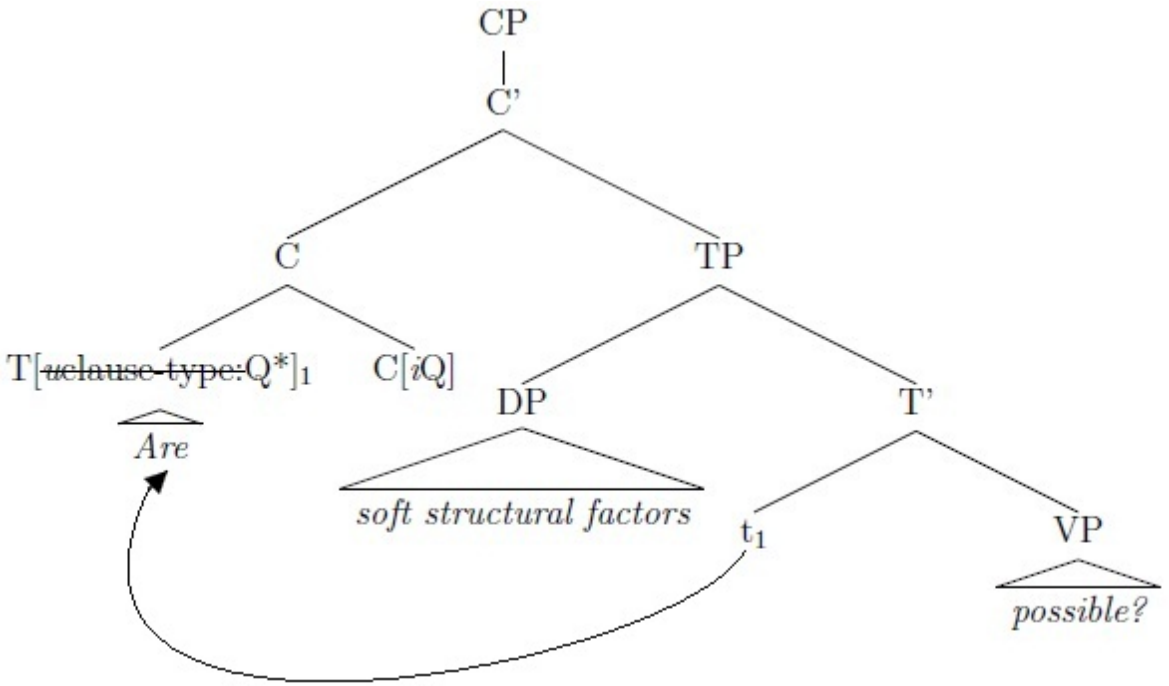
↘ no SAI /if not a question

Formal implementation

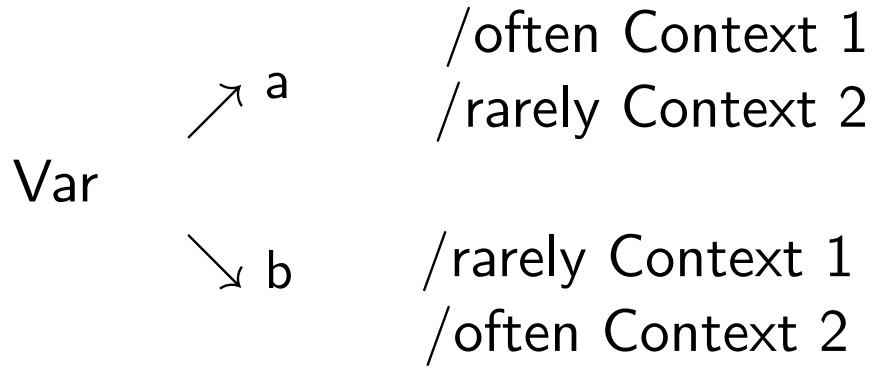
- researcher decides that the difference between question and non-question is structural / part of language knowledge / should be encoded in the formal model of competence

Formal implementation

- researcher decides that the difference between question and non-question is structural / part of language knowledge / should be encoded in the formal model of competence
- one possible implementation:



Soft contextual factor - general description



Dual representation

- Soft contextual factors require dual represenation:

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice
  2. variant = rule, parameter, constraint, feature, construction...

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice
  2. variant = rule, parameter, constraint, feature, construction...
  3. cf. free variation for allophones

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice
  2. variant = rule, parameter, constraint, feature, construction...
  3. cf. free variation for allophones
  4. real synonymy is extremely rare; so variants may have slightly different meanings

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice
  2. variant = rule, parameter, constraint, feature, construction...
  3. cf. free variation for allophones
  4. real synonymy is extremely rare; so variants may have slightly different meanings
  5. the formal algorithm sometimes picks up the one or the other representation and thus variably generates the observed structures

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice
  2. variant = rule, parameter, constraint, feature, construction...
  3. cf. free variation for allophones
  4. real synonymy is extremely rare; so variants may have slightly different meanings
  5. the formal algorithm sometimes picks up the one or the other representation and thus variably generates the observed structures
- Lots of different formulations of this idea (e.g., Roeper 1999)

Dual representation

- Soft contextual factors require dual represenation:
  1. a variant can be represented twice
  2. variant = rule, parameter, constraint, feature, construction...
  3. cf. free variation for allophones
  4. real synonymy is extremely rare; so variants may have slightly different meanings
  5. the formal algorithm sometimes picks up the one or the other representation and thus variably generates the observed structures
- Lots of different formulations of this idea (e.g., Roeper 1999)
- Alternative: leakage - one and the same representation can somehow belong to two different categories at the same time (Denison 2004)

Where do soft contextual factors come from?

Hypothesis 1:

- Statistical distribution of variants always arises for reasons outside of competence
- variation + non-structural (grammar-external) cause
- There are no soft structural factors.

Hypothesis 2:

- Statistical distribution of variants may arise for reasons within competence
- variation + structural (grammar-internal) cause
- Soft structural factors are possible.

Hypothesis 1 details

- Variant a and variant b can both be generated in the same context by the grammar.

Hypothesis 1 details

- Variant a and variant b can both be generated in the same context by the grammar.
- A functional constraint (e.g. Hawkin's Minimize Domain Principle) or other non-structural reasons cause the statistical tendencies observed in large scale corpora.

Hypothesis 1 details

- Variant a and variant b can both be generated in the same context by the grammar.
- A functional constraint (e.g. Hawkin's Minimize Domain Principle) or other non-structural reasons cause the statistical tendencies observed in large scale corpora.
- Hence, *all* soft constraints are "accidental". They "emerge" as grammatical knowledge is used.

Hypothesis 1 example one

"Let me give an example. Abney 1996 suggests that grammars have weighted probabilities for different parses of the same string. So take the phrase (39):

(39) the shooting of the hunters

[...] the NP is [overwhelmingly] interpreted as the object of the gerund, not as the subject. That is a raw fact about frequency of interpretation. [...] It seems to me that the most natural treatment is to say that the phrase is grammatically ambiguous and that extragrammatical factors determine which reading is both statistically preferred and likely to be preferred in a given instance." (Newmeyer 2003)

Hypothesis 1 example two

"Whether a speaker says *You will stay here*, *or You won't go anywhere*, or *You won't go nowhere*, will depend, apart from the meaning he intends to express, on his choice of interactional and social register (assuming he has sufficient command of the registers concerned), whereby frequency considerations may play a limited role. Other than that, frequency plays no role in linguistic competence. In particular, once the speaker has set his sociolinguistic register, the rule systems he employs do not contain a frequency parameter."
(Pieter Seuren's Blog, April 6, 2013)

Hypothesis 2 details

- Variant a and variant b can both be generated in the same context by the grammar.

Hypothesis 2 details

- Variant a and variant b can both be generated in the same context by the grammar.
- But each variant is represented with an inherent "strength"; they can be more or less firmly "entrenched"; they are more or less easily accessible.

Hypothesis 2 details

- Variant a and variant b can both be generated in the same context by the grammar.
- But each variant is represented with an inherent "strength"; they can be more or less firmly "entrenched"; they are more or less easily accessible.
- Hence, soft constraints *can be* structural. They can be a direct consequence of the "design" of variant representation.

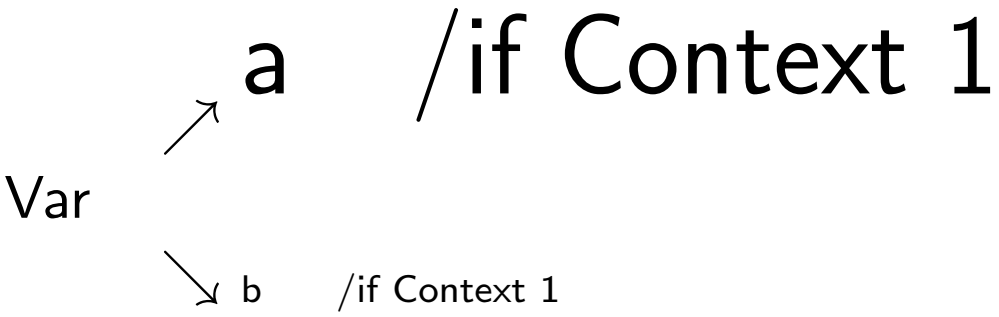
Hypothesis 2 details

- Variant a and variant b can both be generated in the same context by the grammar.
- But each variant is represented with an inherent "strength"; they can be more or less firmly "entrenched"; they are more or less easily accessible.
- Hence, soft constraints *can be* structural. They can be a direct consequence of the "design" of variant representation.
- The most common way to represent this idea is with *probabilistic grammars*; they include a frequency parameter for each dually represented variant.

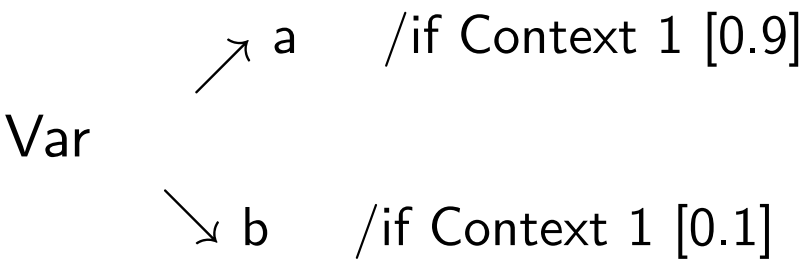
Hypothesis 2 details

- Variant a and variant b can both be generated in the same context by the grammar.
- But each variant is represented with an inherent "strength"; they can be more or less firmly "entrenched"; they are more or less easily accessible.
- Hence, soft constraints *can be* structural. They can be a direct consequence of the "design" of variant representation.
- The most common way to represent this idea is with *probabilistic grammars*; they include a frequency parameter for each dually represented variant.
- Note: This only concerns structural factors. Contextual factors like sociology, genres, registers, phonological weight, information structure etc. need not be encoded in our model of competence but can still be due to functional constraints.

Hypothesis 2 illustration

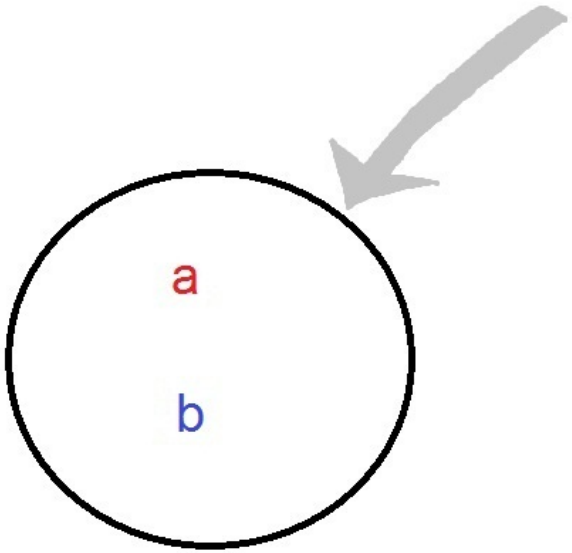


Hypothesis 2 illustration

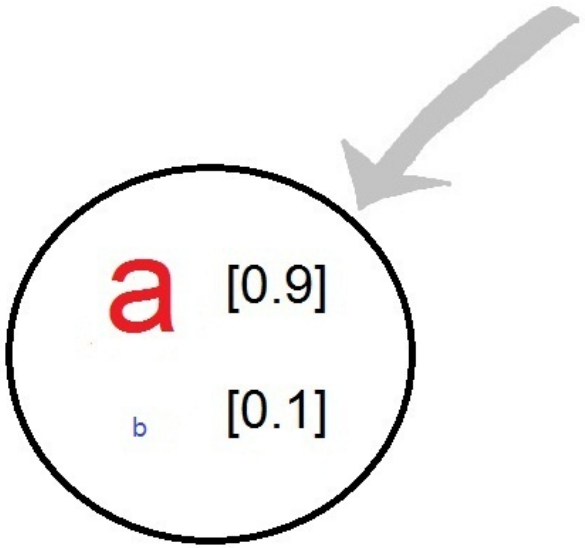


Graphical summary

Hypothesis 1:



Hypothesis 2:



Why bother?

- The question of whether soft structural constraints exist is of paramount importance for diachronic linguistics.

Why bother?

- The question of whether soft structural constraints exist is of paramount importance for diachronic linguistics.
  1. the question is frequently discussed but often only implicitly and vaguely

Why bother?

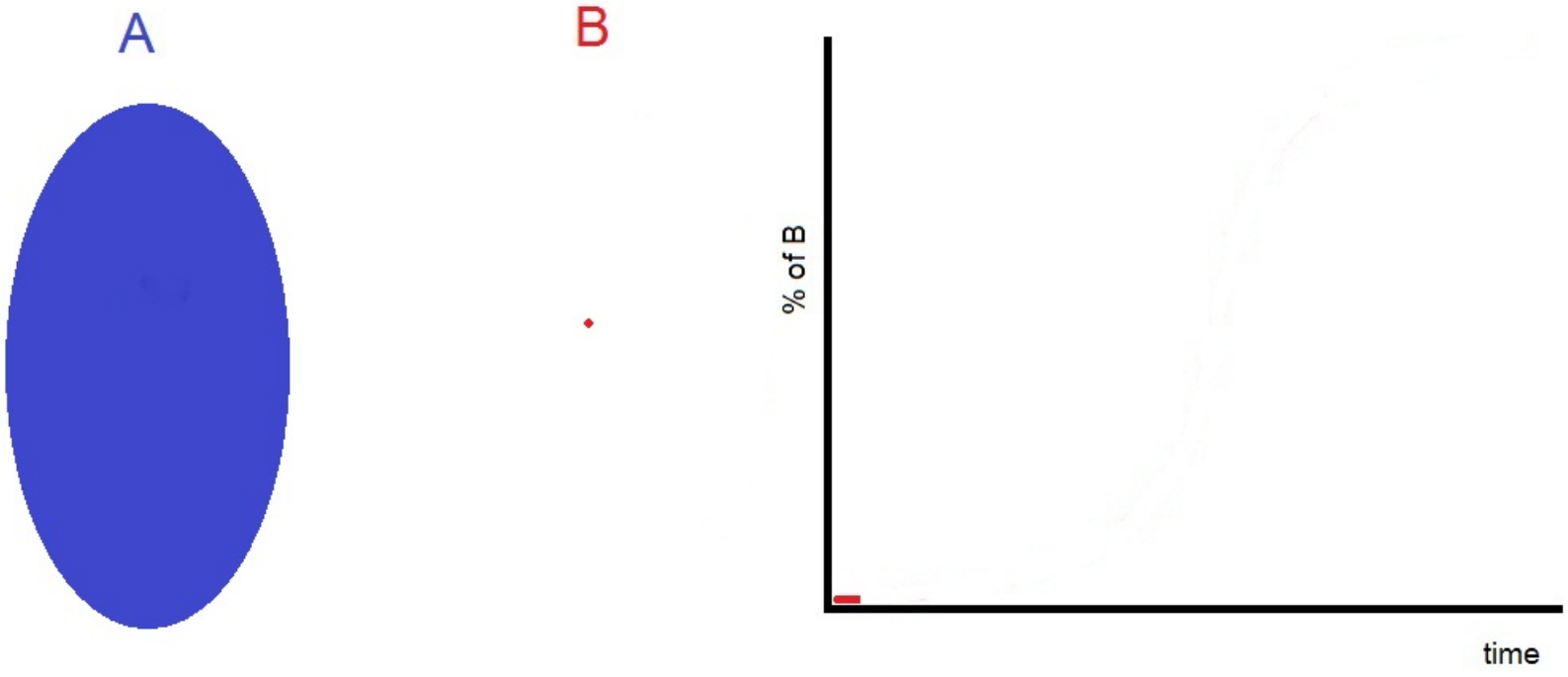
- The question of whether soft structural constraints exist is of paramount importance for diachronic linguistics.
  1. the question is frequently discussed but often only implicitly and vaguely
  2. the answer will affect models of change. e.g. variational learning (e.g., Yang 2002)

Why bother?

- The question of whether soft structural constraints exist is of paramount importance for diachronic linguistics.
  1. the question is frequently discussed but often only implicitly and vaguely
  2. the answer will affect models of change. e.g. variational learning (e.g., Yang 2002)
  3. methodology, how to use historical corpora

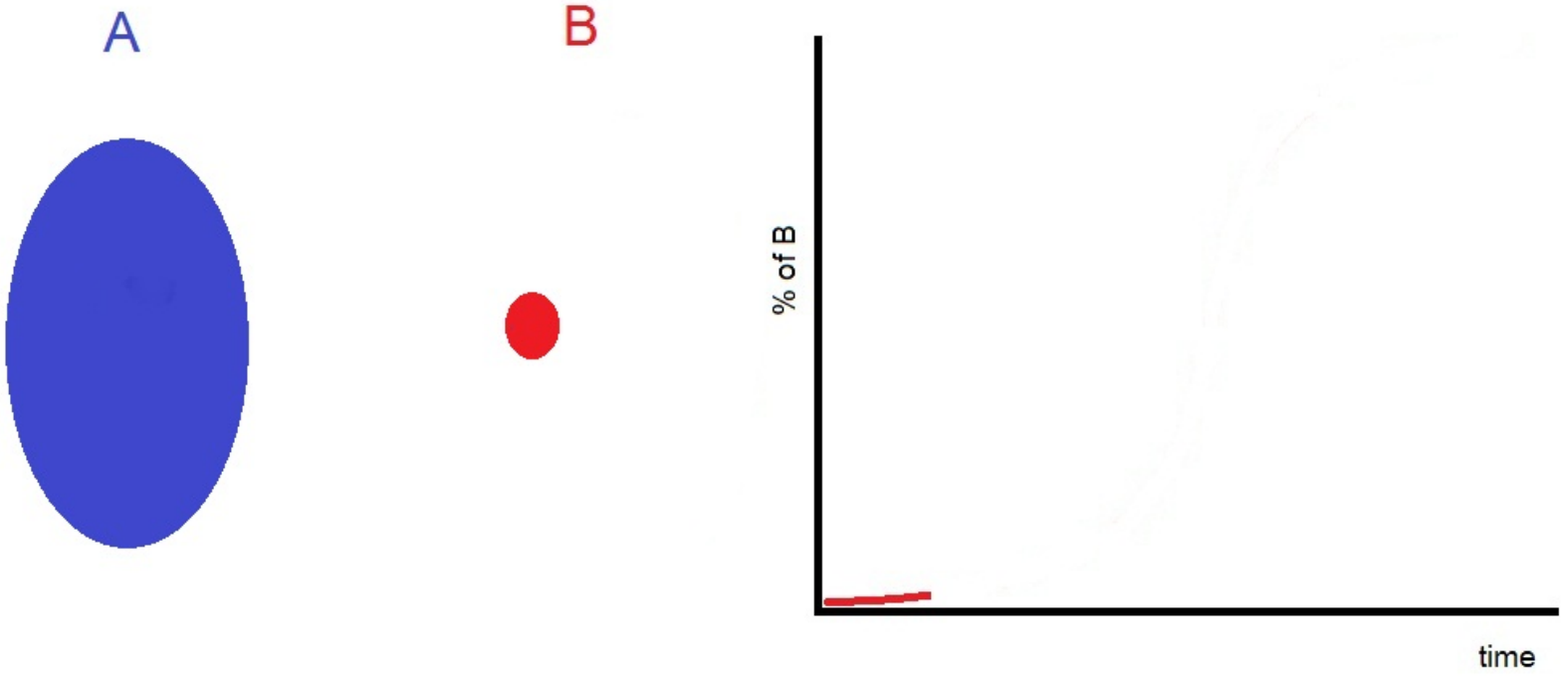
Descriptive facts of language change

- New variant B arises. Actuation.



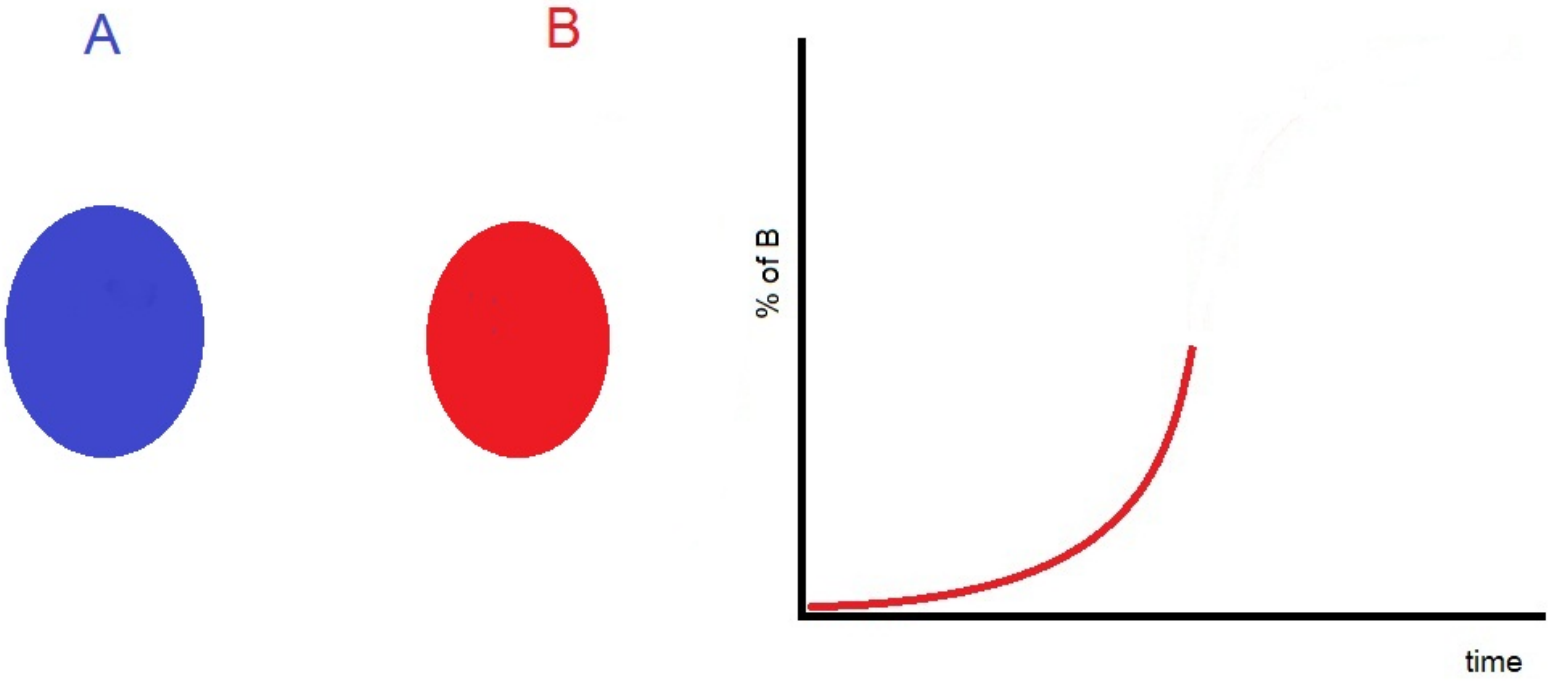
Descriptive facts of language change

- The new form can be found only rarely. Initially slow spread.



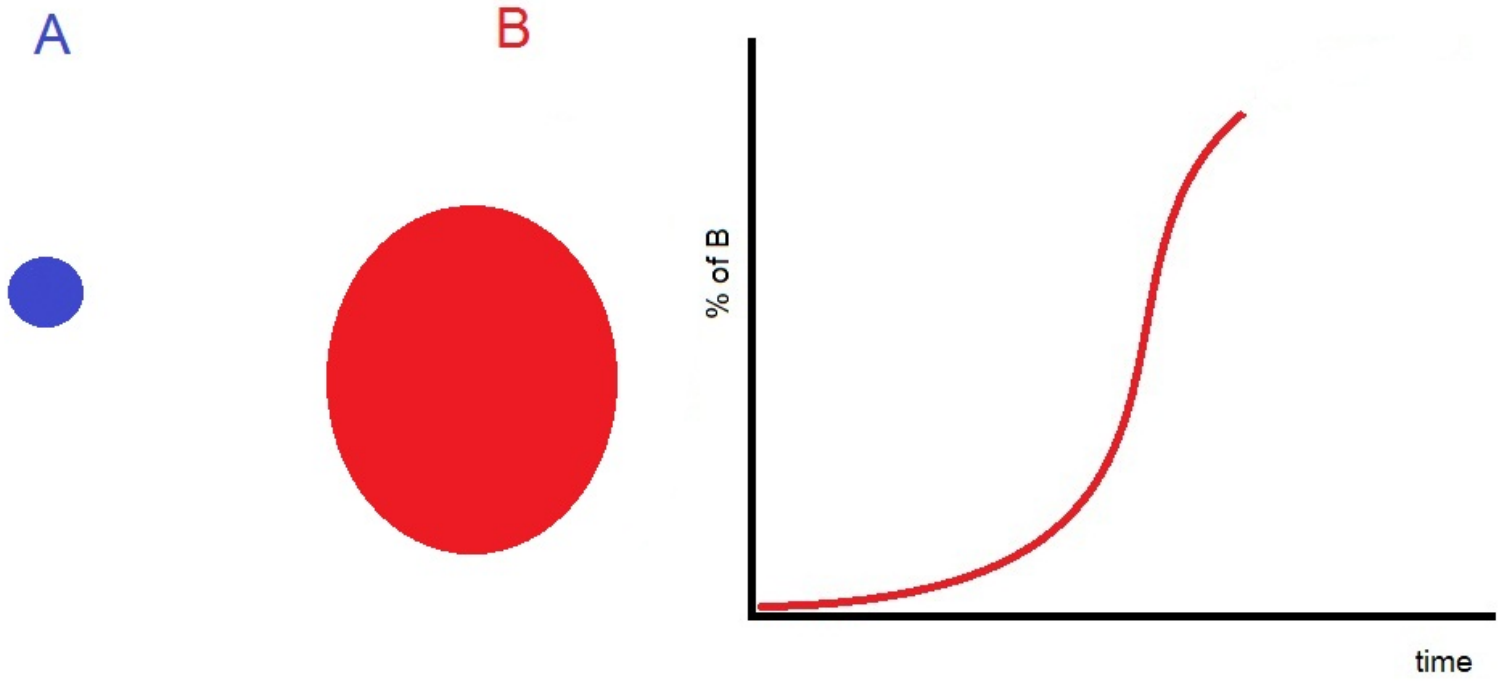
Descriptive facts of language change

- The implementation then accelerates. B overtakes A.



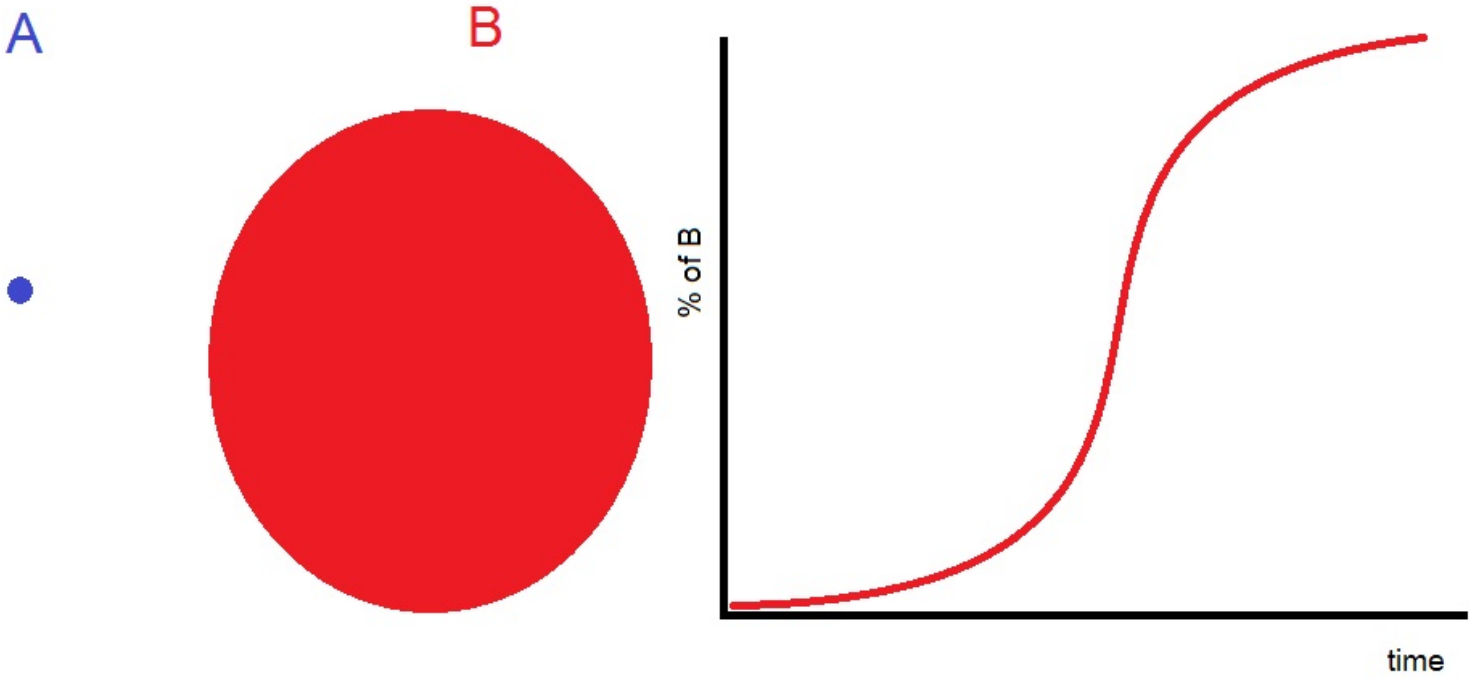
Descriptive facts of language change

- The new variant drives out the old one. A change occurred.



Descriptive facts of language change

- S-shaped curves (e.g., Weinreich et al. 1968; Kroch 1989)



Descriptive facts of language change

- This is not just an idealization of change!

Descriptive facts of language change

- This is not just an idealization of change!
- Syntactic changes really do proceed in this manner, descriptviely.

Descriptive facts of language change

- This is not just an idealization of change!
- Syntactic changes really do proceed in this manner, descriptviely.
- Example: *do*-support with possessive have in American English

Do-support in the history of English

- From early Modern English on, main verbs started requiring *do* for question formation and negation:

  (8) a. Why puttyn men her sonys rair to laue ciuile?
  R. Wimbledon, Serm. (c. 1450)

  b. Why doe men put their sons to the Civill Law?
  R. Wimbledon, Serm. (c. 1635)

  (from Denison 1993)

Do-support with possessive *have*

- For AmE possessive *have*, the development is recent, and is in fact only now coming to completion.

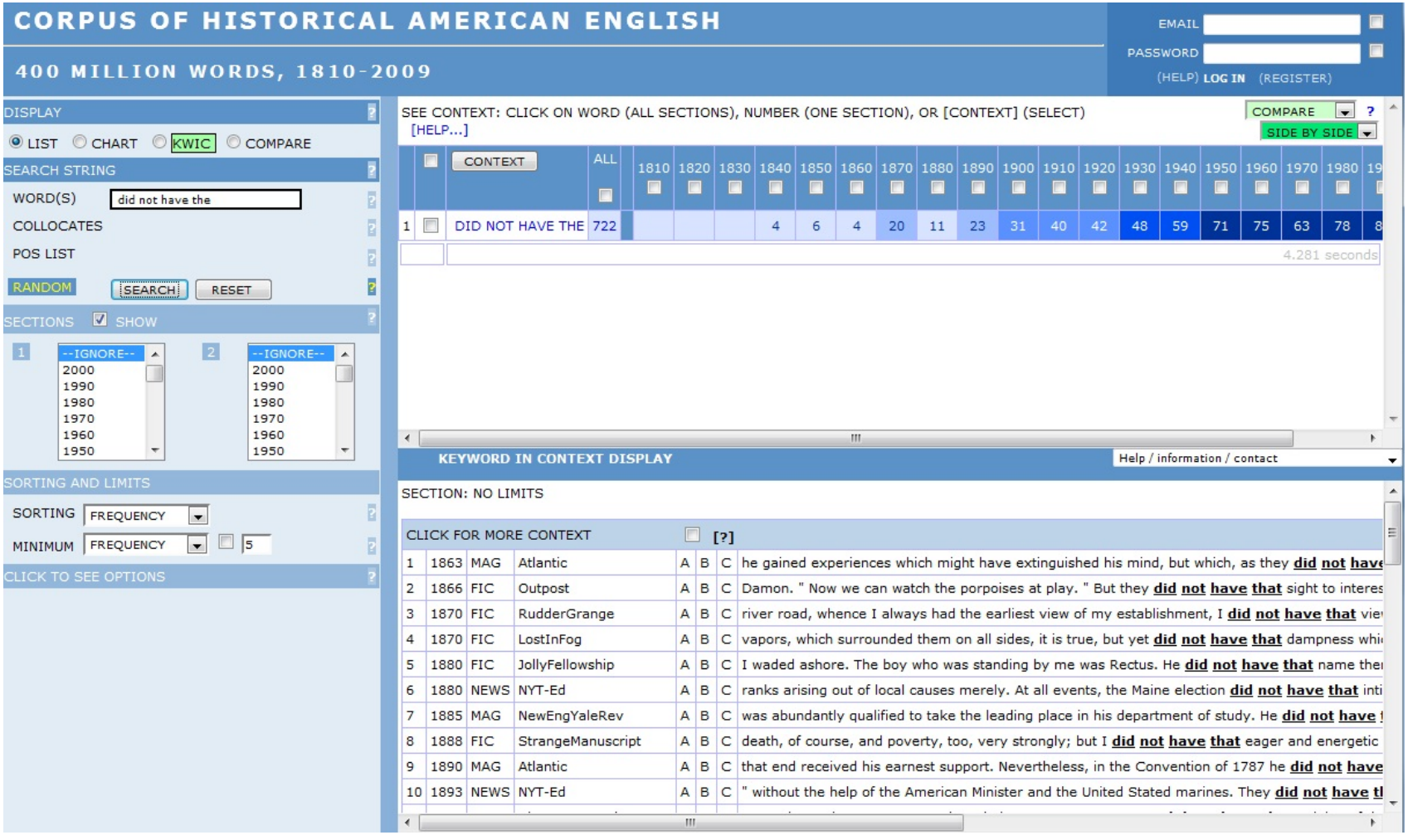
(9) a. I may add that it is a matter in which I have not the slightest interest.
ViolaGwyn (1922)

b. I do not have the slightest intention of even mentioning the word " Bermuda"
Play:CatsPaw (1998)

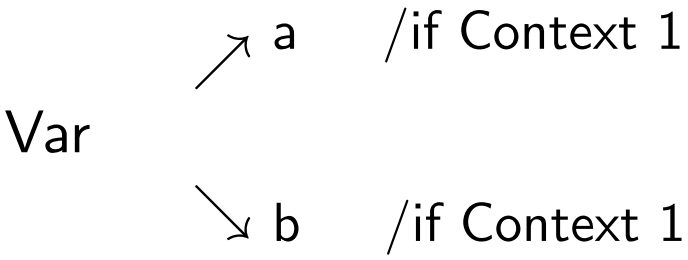
Data collection

- 400m word Corpus of Historical American English (Davies 2010)
- collection of all sentences with negation and possessive *have*
  1. 120 queries like "have not the", "hasn't any", "don't have a", "did not have his" etc.
  2. manual clean-up (still in progress)

Data collection - COHA



Variants and contextual factor



Variants and contextual factor

Negation ↗ *do*-support /if possessive *have*

Negation ↘ no *do*-support /if possessive *have*

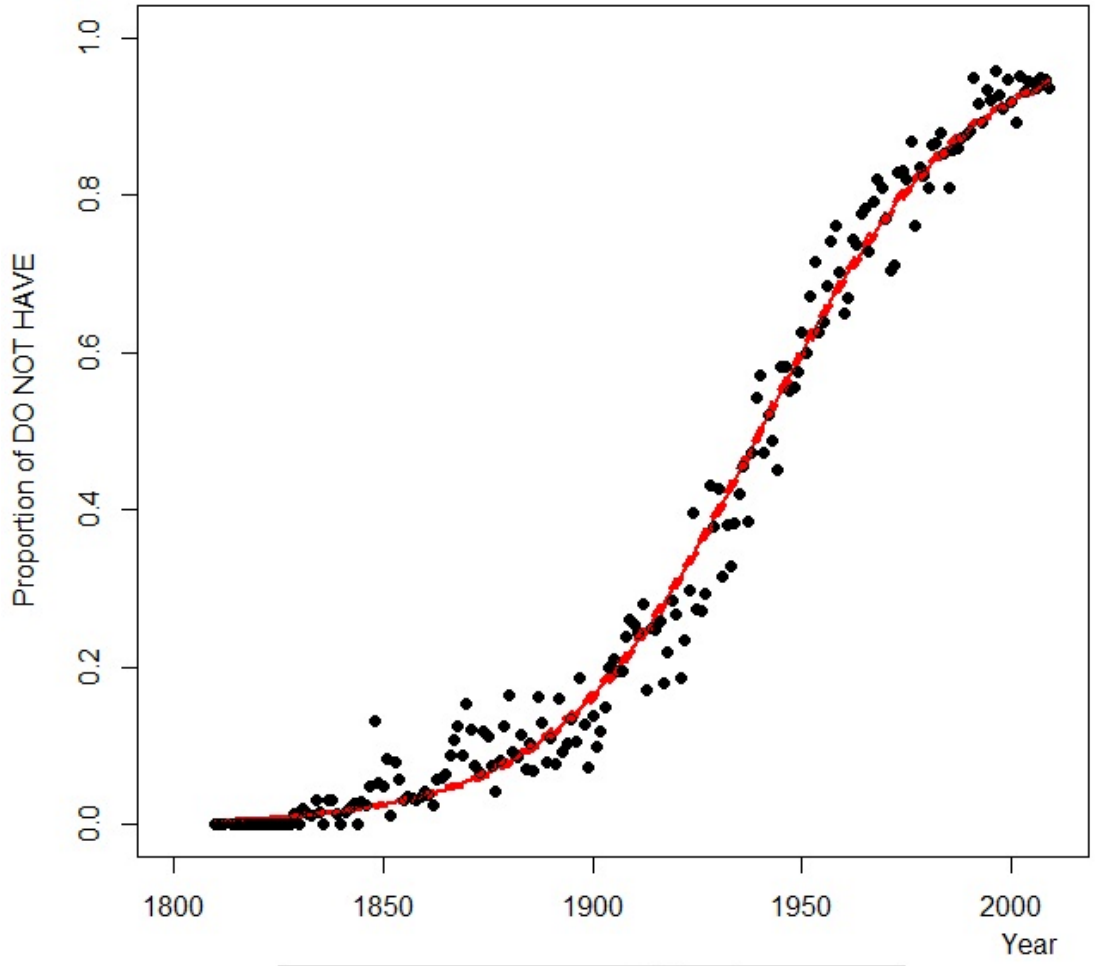
Variants and contextual factor

Negation ↗ *do*-support /if possessive *have*

Negation ↘ no *do*-support /if possessive *have*

- The contextual factor (possessive *have* vs. other kinds of verbs) is plausibly structural.

Results



Factor	Estimate	Wald (z^2)	df	p
(Intercept)	-79.04	9846.6	1	<0.001***
Year	0.04	9868.4	1	<0.001***

Evaluation

- one of the most extensive studies of a syntactic change, N=33,456 examples
- sufficient material to have a data point for every single year!
- clear support of the hypothesis that syntactic changes proceed in s-shaped curves

Hypothesis 1

- If speakers cannot represent near-synonymous variants with different strengths, non-structural explanations for a change must be sought, i.e. something outside of competence must change as the cause of competence.
  1. Functional constraints (e.g. Hawkins Minimize Domains principle) should be a relatively universal aspect of the human processor. How can that change?
  2. Sociological factors are not really plausible. *Do*-support with possessive *have* is not an identity marker. The American speech community did not experience any major disruptions (e.g. no invasion, no substratum effects). Americans learned English continuously from one generation to the next.
- The s-shaped curve is unexpected. Change should be more chaotic, like a drift, or catastrophic.

Hypothesis 2

- If speakers can represent near-synonymous variants with different strengths, those variants can emerge, grow and die out.
- With any kind of replicator-dynamics, the s-shaped curve follows mathematically.

  E.g. Variational learning:

$$p_{t+1} = \frac{\alpha p_t}{\alpha p_t + \beta q_t}$$

Conclusion

Conclusion

$\rightarrow$ the facts of language change seem to me to support the hypothesis that variants can be represented more or less strongly, i.e. of probabilistic grammars.

A Counterargument

- The development found in corpus studies are irrelevant. They can never reveal anything about the competence of an individual.

Some quotes

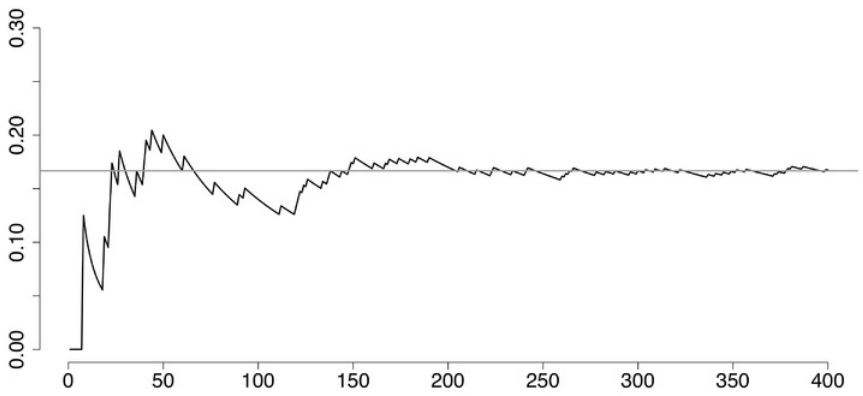
"corpus-derived statistical information is [hardly] relevant to the nature of the grammar of an individual speaker."
"There is no way that one can draw conclusions about the grammar of an individual from usage facts about communities"
"we do not have group minds"
"the corpora upon which stochastic grammars of American English are based do not [...] explicitly exclude spoken language samples from AAVE. And why stop at the American border? If corpora contain utterances from diverse dialects of American English, then why not from British, Australian, and Indian English as well?"
(Newmeyer 2003, 2005)

Shared linguistic knowledge

- It is not quite correct to say that we do not have "group minds" in this context. A relatively large proportion of competence must be communal , i.e. shared by speakers of a speech community.
  1. Communication would be impossible; all speech would be mutually unintelligble.
  2. Abstractions like "English", "French" would not make sense.

Law of large numbers

- If a trial is repeated many times, the average of the results will approach the expected value.



As a trial is repeated many times (x-axis), the expected value (y-axis) is approached

- Similarly, if examples in a corpus are numerous enough, the "true average" in the population will be approached.
- If absolutely nothing was known about an individual, the expected value would be predicted for that individual. Most people are close to the average.

Controlled variables

- Corpora control for variables that would introduce a lot of noise, i.e. corpora of particular modes (spoken or written), genres, dialects. This is standard practice in corpus construction.

Predictions

- If corpus information controls for individual variables and there are enough examples, predictions will be very accurate.
- Example: if theme / recipient pronominality, theme:recipient length ratio, animacy, spoken vs. written mode, age and gender of speaker are known, 94% of individual utterances of dative alternation are correctly predicted based on training data from corpora (Bresnan et al. 2007)

S-shaped curve again

- If the corpus data were completely divorced from individual competences, we could not expect a regular pattern, but chaos.
- The noise in the data may easily come from individual differences, functional constraints and measuring errors but the regularity in the data cannot be explained with such factors.
- Put differently, the large number of observation smoothes out individual preferences to the population average. It is therefore likely that the *regular* patterns cannot be explained away with *heterogenity* in the data.

Conclusion

- It is much too simplistic to assert that corpus information cannot in principle reveal anything about he competence of individuals.

Conclusion

- It is much too simplistic to assert that corpus information cannot in principle reveal anything about he competence of individuals.
  1. There is shared linguistic knowledge.

Conclusion

- It is much too simplistic to assert that corpus information cannot in principle reveal anything about he competence of individuals.
  1. There is shared linguistic knowledge.
  2. Large numbers of examples approach population average.

Conclusion

- It is much too simplistic to assert that corpus information cannot in principle reveal anything about he competence of individuals.
  1. There is shared linguistic knowledge.
  2. Large numbers of examples approach population average.
  3. Important variables can be controlled for.

Conclusion

- It is much too simplistic to assert that corpus information cannot in principle reveal anything about he competence of individuals.
  1. There is shared linguistic knowledge.
  2. Large numbers of examples approach population average.
  3. Important variables can be controlled for.
  4. Consequence: individual predictions based on corpus data are accurate.

Conclusion

- It is much too simplistic to assert that corpus information cannot in principle reveal anything about he competence of individuals.
  1. There is shared linguistic knowledge.
  2. Large numbers of examples approach population average.
  3. Important variables can be controlled for.
  4. Consequence: individual predictions based on corpus data are accurate.
  5. S-shaped curves are explainable only if there is some correlation between group averages and individual preferences.

Thank you for your attention!

Bresnan, J., T., N. and Baayen, H. R. (2007), Predicting the dative alternation, *in* G. Bouma, I. Kraemer and J. Zwarts, eds, 'Cognitive foundations of interpretationg', KNAW, Amsterdam, pp. 69–94.

Davies, M. (2010), *The Corpus of Historical American English: 400 million words, 1810-2009*, http://corpus.byu.edu/coha/ (Accessed 10 September 2013).

Denison, D. (2004), Do grammars change when they leak?, *in* C. Kay, S. Horobin and J. Smith, eds, 'New perspectives on English historical linguistics: Selected papers from 12 ICEHL, Glasgow, 21-26 August 2002, vol. 1, Syntax and morphology', Amsterdam, Benjamins, pp. 15–29.

Denison, D., ed. (1993), *English historical syntax: verbal constructions*, Longman Linguistics Library, London.

Haegeman, L. (1990), Non-overt subjects in diary contexts, *in* J. Mascaro and M. Nespor, eds, 'Grammar in Progress, GLOW essays for Henk van Rijmsdijk', Forisl, Dordrecht, pp. 167–174.

Hawkins, J. (2000), 'The relative order of prepositional phrases in english: Going beyond manner-place-time', *Language Variation and Change* **11**, 231–266.

Kroch, A. (1989), 'Reflexes of grammar in patterns of language change', *Journal of Language Variation and Change* **1.3**, 199–244.

Newmeyer, F. J. (2003), 'Grammar is grammar and usage is usage', *Language* **79**, 682–707.

Newmeyer, F. J. (2005), 'A reply to the critiques of 'grammar is grammar and usage is usage"', *Language* **81.1**, 229–236.

Oostdorp, M. v. (2011), 'On statistical effects', *Paper delivered at Workshop on West Germanic Phonology, Mannheim, September 2011* .

Roeper, T. (1999), 'Universal bilingualism', *Bilingualism: Language and Cognition* **2.3**, 169–186.

Weinreich, U., Labov, W. and Herzog, M. (1968), Empirical foundations for a theory of language change, *in* W. Lehmann and Y. Malkiel, eds, 'Directions for historical linguistics', Directions for historical linguistics, University of Texas Pressl, Austin, pp. 95–198.

Yang, C. (2002), *Knowledge and Learning in Natural Language*, Oxford University Press, Oxford.