A unified theory of stable variation, syntactic optionality, and syntactic change

Joel C. Wallenberg

Newcastle University

`joel.wallenberg@ncl.ac.uk`

August 2, 2013

University of Ottawa

DiGS 15

Introduction

Variation in grammar is often described as falling into one of two categories.

1. Competing Grammars
   - Leads to language change via the replacement of one grammatical process by another.
   - Competition is parameterized in some fashion, as in competing flavors of the same functional head (Kroch, 1994).
2. Optionality (within a grammar?)
   - Diachronically stable variation between grammatical processes.

Introduction

Hypothesis: all variation, including grammatical optionality, is formally **competing grammars**, with the following consequences (Fruehwald and Wallenberg, 2013):

- We expect variation (apparent optionality) between two grammatical forms to be diachronically unstable, generally.

- True optionality = stable variation, and its difference from language change must be explained by some mechanism of language use, outside of the grammar itself.

- We argue that it depends on the mathematical character of some extragrammatical dimension with which the variation interacts.

  - Partial specialization of variants along a continuous dimension.

Outline

Blocking and Contrast

"Blocking Effect" (Aronoff, 1976)

- General cognitive pressure against two forms existing for one function ("doublet"), unstable (Kroch, 1994).

  {*lough, laughed*} (laugh-PST; Taylor 1994)
  {*jimmies, sprinkles*} (candy topping)

"Principle of Contrast"

- A strategy that children use in acquiring language: assume that two forms have two meanings (or uses)(Clark, 1987, 1990, *inter alia)*.
- Children hypothesize that novel words also refer to novel objects (as in Markman and Wachtel, 1988, among many other replications of the effect).

The Principle of Contrast

- A strategy that children use in acquiring language: assume that two forms have two meanings (or uses) (E. Clark 1987, *inter alia*).
  - Synonyms should only be acquired as a last resort.
- Demonstrated in experiments such as Markman and Wachtel (1988).
  1. 20 children
  2. 6 pairs of one familiar item (banana, cow, cup, plate, saw, spoon) and one unfamiliar item (cherry pitter, odd shaped wicker container, lemon wedgepress, radish rosette maker, studfinder, tongs).
  3. **Control**: "Show me one"
  4. **Test**: "Show me the X" (X = nonsense syllable)
- Control children pick the unfamiliar object at chance levels, but test children choose unfamiliar objects significantly higher than chance.

Blocking = Contrast + Selection

- A doublet is two variants competing for finite resources, as in e.g. biological evolution.
  - Instead of competing for something like food, they are competing for use (time in the mouths/brains of speakers).
  - **Selection** operates on the number of times a variant is heard (and accurately analyzed) by an acquirer.
- Either one variant has an advantage, and so replaces the other (following a logistic function; Nowak, 2006).
- Or neither variant has an advantage (or much of one), in which case random walk and drift.
- But in linguistic doublets, random walk cannot persist indefinitely because of the acquisition pressure of the Principle of Contrast.

Doublets = Competing Grammars (Kroch, 1994)

"Competing Grammars", general form: 2 variants are available to a speaker, in the relevant inventory of grammatical formatives, with overlapping functions (e.g. the same meaning).

- E.g. two featural versions of the same syntactic head.
- E.g. two different output mappings for the same phonological input.
- E.g. two different Spell-outs of a morpheme.

- A fact about language use: at some point in the derivation, the speaker reaches a **decision-point**.

  - The speaker has a choice between formatives to continue the derivation, and either will result in a grammatical utterance and a meaning close enough to the speaker's intention.
  - It is speaker "choice" more in the sense of an urn problem.

Doublets = Competing Grammars

- Necessary for the description of **any** linguistic change in a categorical dimension.
    - E.g. word-order parameters (Pintzuk, 1991; Santorini, 1992); a phonological rule like German final stop devoicing (Fruehwald, Gress-Wright, & Wallenberg 2009).
    - In any such case, a speaker in the middle of the change in progress (code-)switches between categorical variants (Kroch, 1989).

Summary: Blocking and Contrast

So, doublets are Competing Grammars, and the possible
historical outcomes are driven by the selection and the Principle
of Contrast.

Proposal: every case of categorical linguistic variation or
optionality can be reduced to Competing Grammars, leading to
one of these two outcomes.

This simplifies the grammatical architecture necessary to
account for both optionality and language change (in pursuit of
a Minimalist hypothesis).
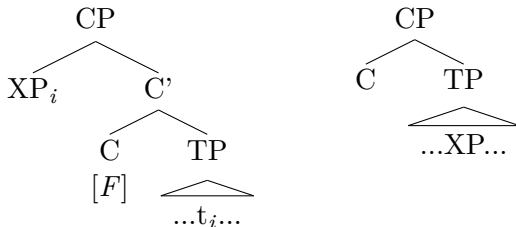
Example: English "Topicalization"

- Prince (1985, 1998, 1999): felicitous in two English discourse contexts, both of which require a certain type of contrast to appear on the fronted XP.

    (1) She's going to use three groups of mice. One, she'll feed them mouse chow, just the regular stuff they make for mice. Another she'll feed them veggies. And the third she'll feed junk food.

    (2) She was here two years. [checking transcript] Five semesters she was here.
        (Prince, 1999, 8,9)

- However, it is **never** obligatory.

Example: English Topicalization

- As long as the accent pattern is kept constant, both orders are felicitous:

  (3) She's going to use three groups of mice. One, she'll feed them mouse chow, just the regular stuff they make for mice. Another she'll feed them veggies. And **the third** she'll feed **junk food**.

  (4) She's going to use three groups of mice. One, she'll feed them mouse chow, just the regular stuff they make for mice. Another she'll feed them veggies. And she'll feed **the third junk food**.

Topicalization in Minimalism

- Move is triggered by the feature content of some head.
- Given "Merge...preempts Move" (Chomsky, 2000), a feature cannot encode optional movement.
- Therefore, optional movement must involve a choice (for the **Numeration**) between two variants of a functional head, out of an inventory of possible heads:

```
        CP                          CP
       /  \                        /  \
     XP_i   C'                    C    TP
           /  \                       /  \
          C    TP                    ...XP...
         [F]   /  \
              ...t_i...
```

- This is the core case of morphosyntactic doublet (i.e. competing heads) described in Kroch (1994).

A Minimalist Hypothesis

Given that:

- these mechanics are necessary to encode syntactic optionality in a Minimalist system,
- the same mechanics are necessary to describe a change in progress,

Then, the system is simplest if no more machinery is added to deal with optionality/variation.

- Note: syntactic optionality represented as multiple formal (featural) versions of a functional head, competing for the speaker's choice in the speaker's inventory, is a logical consequence of the "Borer-Chomsky Conjecture" (Borer 1984, so named in Baker 2008).

Minimalist Theory of Variation

- **Prediction:** every case of syntactic optionality or variation is one of the following:
    1. A replacement change in progress (outright competition going to completion).
    2. A specialization change in progress (specialization for different functions going to completion).
    3. **"Stable" variation, or optionality:** variants have partially specialized along a continuous (or ordinal) dimension, e.g. style, prosodic weight.

- If categorical variants specialize along a categorical dimension, complete specialization should eventually result.

- If categorical variants specialize along a continuous or ordinal dimension, then complete specialization can **never** result (but replacement can still be arrested).

Example: Embedded Polar Questions

A quantitative study of embedded *yes/no*-questions in English and Icelandic, comparing the use of *whether* vs. *if*, and *hvort* vs *ef* found **specialization** in English, and **replacement** in Icelandic (Bailey, Wallenberg, & van der Wurff 2012).

(5) John wondered whether Mary was coming to the party.

(6) John wondered if Mary was coming to the party.

Bailey et al. (2012) suggest the Proto-Gmc dual pronoun cognate with *whether* is reanalyzed as a polarity wh-word in clauses containing a disjunction.

I asked *whether* ("which of two") he wants, A or B.

Example: Embedded Polar Questions

In all stages of English and in historical Icelandic, a disjunction favors *whether*.

English

Disjunction:

(7) I wonder {**whether**,if} John or Bill is bringing coffee.

(8) I wonder {**whether**,if} John is bringing tea or coffee.

Simple:

(9) I wonder {whether, **if**} Bill is bringing coffee.

Example: Embedded Polar Questions

Disjunction:

(10) eftir   því   **hvort** maður vill   heitt eða
     according it-DAT whether man   wants hot   or
     kalt
     cold
        "According to whether one wants hot or cold"
        (*Sagan Öll*, date: 1985, from IcePaHC)

Simple, (older) Icelandic:

(11) vér vitum eigi, **hvort**   vér tökum öndina
     We know not   whether we take   soul-the

(12) og spurðu, **ef** hann væri Kristur
     and asked if he   were Christ
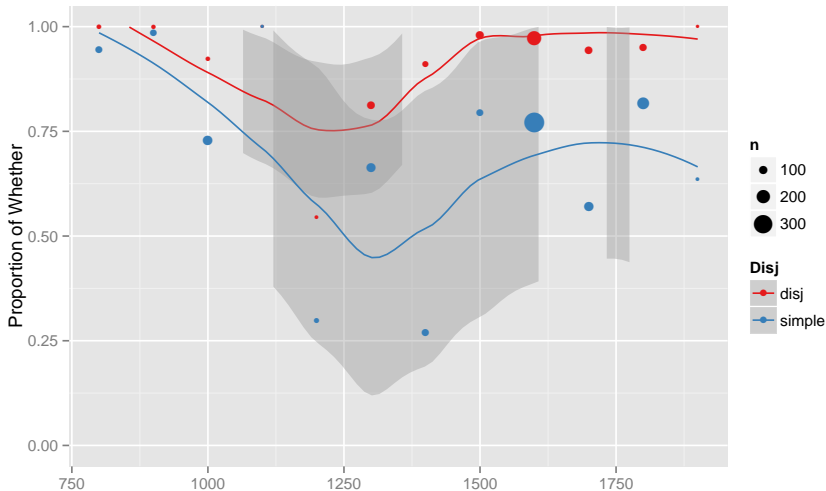   (*Icelandic Homilies*, date: 1150, from IcePaHC)
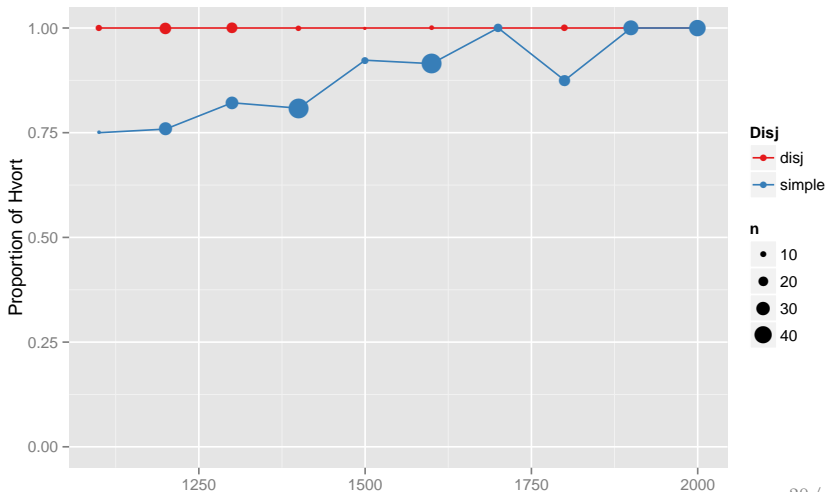
Specialization in English (N = 1929 clauses)

Parsed Corpora: YCOE, PPCME2, PPCEME, PPCMBE

Replacement in Icelandic (N = 397 clauses)

IcePaHC (Wallenberg et al., 2011)

Stable Variation

Hypothesis: Stable variation, i.e. optionality, results from categorical variants specializing along a continuous dimension.

There are many possible continuous dimensions, including language internal dimensions like

- weight (word length)
- prosodic accent (number of aligned prosodic peaks, degree of stress clash between two positions)

and language external dimensions like

- style
- speech rate

Example: English Topicalization

- Is the frequency stable over time? Possibly since Late Middle English (Speyer, 2010). (Though this is subject to revision).

- Is it specialized for different speech styles (registers)? Not that we know of.

- Is it sensitive to prosody? Definitely (Speyer, 2008, 2010).

Prosodic Sensitivity

- Speyer (2008, 2010) shows experimentally that prosodically ill-formed topicalization is subject to prosodic repair.

(13) **The first group** she'll feed mouse chow, **the second** she'll feed veggies, and **the third** she'll feed junk food.

(14) ? **The first** Caitlin will feed, **the second** Joe will feed, and **the third** Maggie will feed.

(15) ?? **Joel**$_i$ Caitlin will pay t_i, **Bob**$_j$ Joe will pay t_j, and **Ann**$_k$ Maggie will pay t_k.

(16) ??? **Joel**$_i$ Caitlin will pay t_i 10 dollars, **Bob**$_j$ Joe will pay t_j 15 dollars, and **Ann**$_k$ Maggie will pay t_k 20 dollars.

Example: -ɪn∼-ɪŋ

(17) John has been {singing/singin'}.

(18) {Dunking/Dunkin'} Donuts.

- Is the frequency stable over time? Probably, as the variation has its roots in OE morphology (Houston 1985), and both variants were present in Middle English texts (Labov, 1989).

- Is it specialized for different grammatical contexts? Yes, in part, along a nominal↔verbal dimension (Labov, 1989).

- It it specialized for different speech styles? Yes, in part, along a continuous dimension of formality.

Example: -ɪn∼-ɪŋ

A proof of concept simulation shows that plausibly, under minimal acquisition assumptions:

- Variants specialize along a continuous dimension like style.
- For a continuous dimension, the process will stabilize at **partial** specialization.

Gen 0: -m∼-ŋ doublet is innovated, with no stylistic conditioning. **Gen 0** picks a style to speak in, and produces a variant, repeats.

Gen 1: **Gen 1** learns an estimate for the style value of -m, ŋ, as soon as she hears the first tokens of each from **Gen 0**. She adjusts this estimate as she gets more data from **Gen 0**.

Gen 2: **Gen 1** picks a style, produces one of the variants with a probability weighted by how far her style estimates are from the current style, repeats. **Gen 2** learns style estimates for variants as above.

Instructive Observations from Simulation

1. We did not build the Principle of Contrast into the model. We just left the learner the possibility of contrasting the variants.

2. There is no default style value for the variants. Their initial values are the first styles the learner hears from the last generation.

3. The variants do specialize, but not all the way to the limits – they stabilize at the quartiles.

4. Given 20000 tokens in each generation to learn from, the specialization stabilizes in about 4 generations.

Open Question: can we predict when specialization will occur successfully (or when replacement will happen instead)?

When Specialization vs. Replacement?

- If neither variant has a selectional advantage, then specialization should occur (modulo **drift**).
- But if there is an advantage, then we have a race; who wins?
- First, the speaker must identify some **salient domain of specialization**.
- Specialization is driven by the Principle of Contrast, which is a fact about child cognition (i.e. should be as stable as the human brain).

When Specialization vs. Replacement?

- Suppose, then, that specialization always occurs at the same rate per token of the linguistic variable that the child encounters.
  - **Factor 1:** the size of selectional advantage for one variant vs. another, compared with the fixed rate of specialization.
  - **Factor 2:** the frequency, in speech, of the context for the linguistic variable in question.
  - Even if Factor 1 is small compared with rate of specialization, replacement could still occur if Factor 2 is small enough.

- The specialization stabilizes at 20000 tokens, but with 1000 tokens it never does – implies that the frequency of the variant could be a determining factor.

Case Study: Relative Clause Extraposition

French

(19) mais l'heure vient que je ne parleray plus a
      but the time comes that I NEG speak-FUT more to
      vous en proverbes
      you in proverbs
         "The time approaches when I will no longer speak to
      you in parables"
   (MCVF, 1523-NEW-TESTAMENT-P, A5V.2491)

English

(20) none lives that more loves you
      (PCEEC, TIXALL, 53.019.369, date: 1619)

Within Speakers, and Diachronically

- Relative clause extraposition in the Parsed Corpus of Early English Correspondence (PCEEC; Taylor et al. 2006).

- Allows us to look at reasonable samples from individual speakers (letter-writers), as well as an historical sample from 1400–1700.

- Coded for prosodic weight of the relative clause, in number of words, from 0–50.

- Also coded for extraposition from Subject vs. Object.

Hypothesis: individual speakers treat weight as a continuous variable, with extraposition specialized imperfectly along it, in roughly the way the acquisition model predicts (following on Ingason and MacKenzie, 2011).

All PCEEC, over time (N = 8073 clauses)
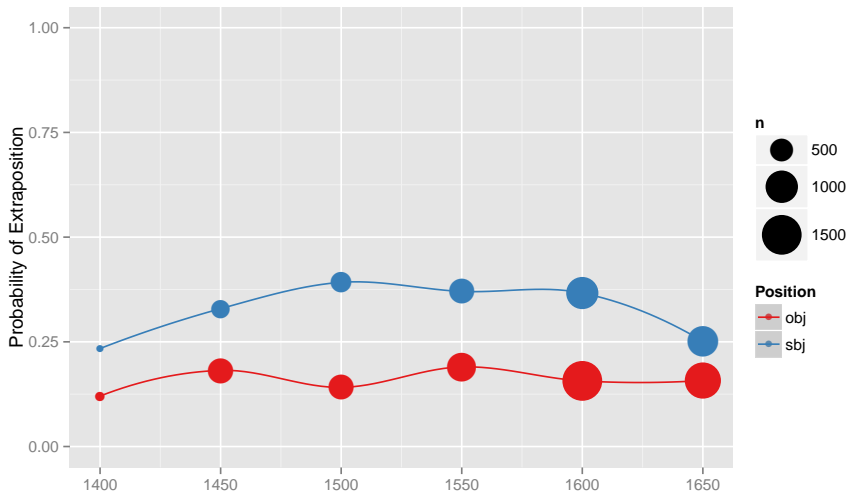
Diachronically, Crosslinguistically

- **English:** YCOE (Taylor et al., 2003), PPCME2 (Kroch and Taylor, 2000), PPCEME (Kroch et al., 2005), PPCMBE (Kroch et al., 2010).
- **Icelandic:** IcePaHC (Wallenberg et al., 2011).
- **Old/Middle French:** MCVF Corpus (Martineau, Hirschbühler, Kroch & Charles Morin, 2010)
- **Historical Portuguese:** Tycho Brahe Corpus of Historical Portuguese (Galves and Britto, 2002/2013).
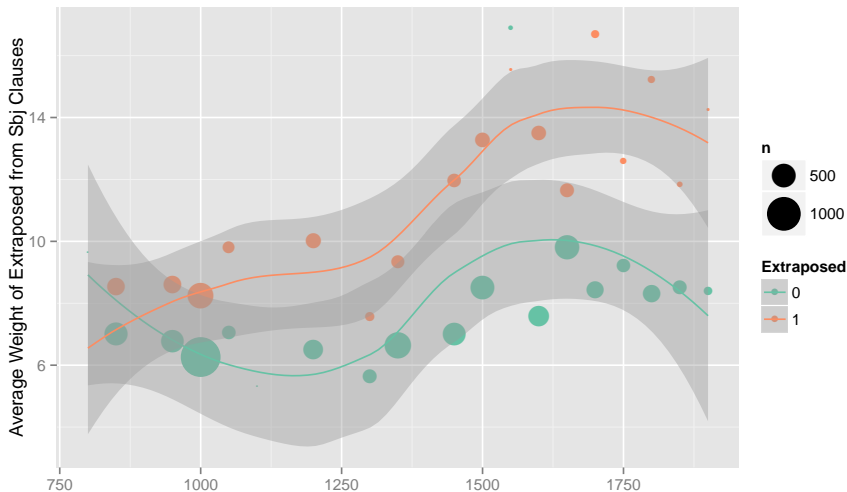
Hypothesis 1: the specialization for weight has stabilized, so the effect of weight will be constant over time. **(Confirmed!)**
Hypothesis 2: the overall rate of relative extrapolation will be stable over time. **(Rejected!)**
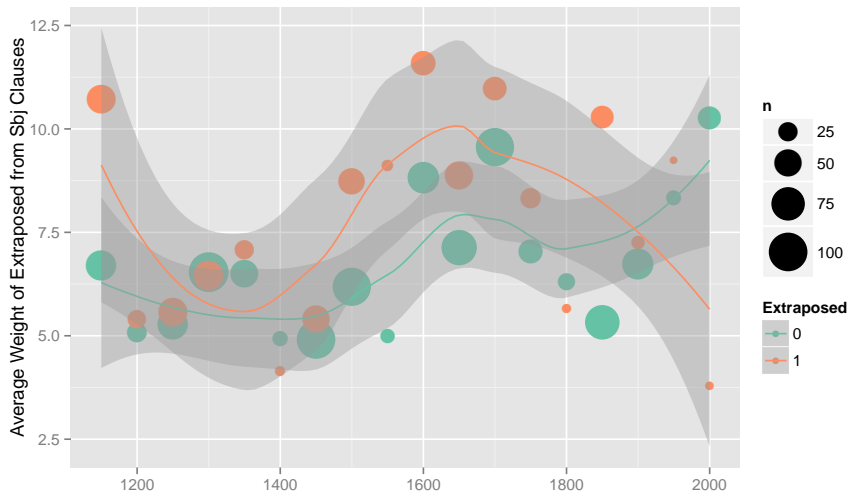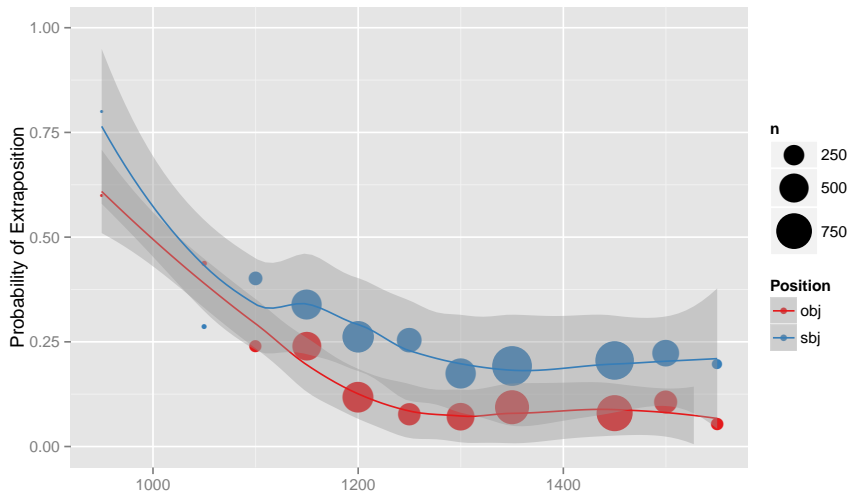
English, over time (N = 18530 clauses)
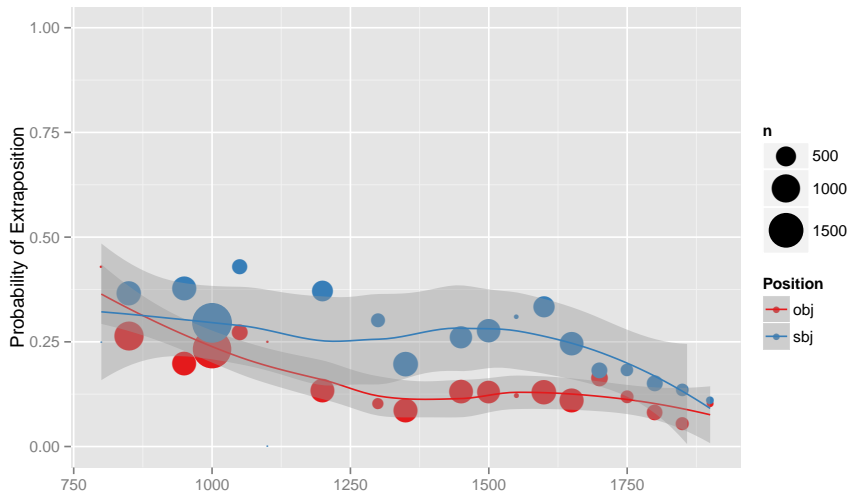
English, average weight over time (N = 18530)
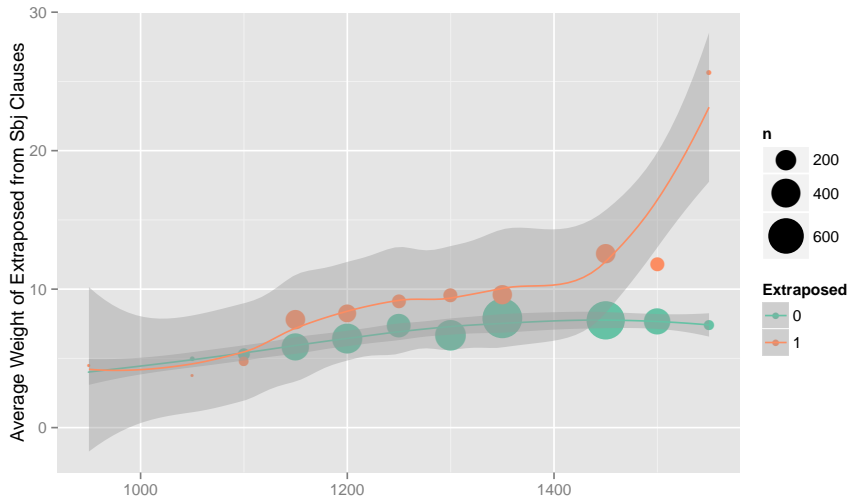
Icelandic, over time (N = 3486)

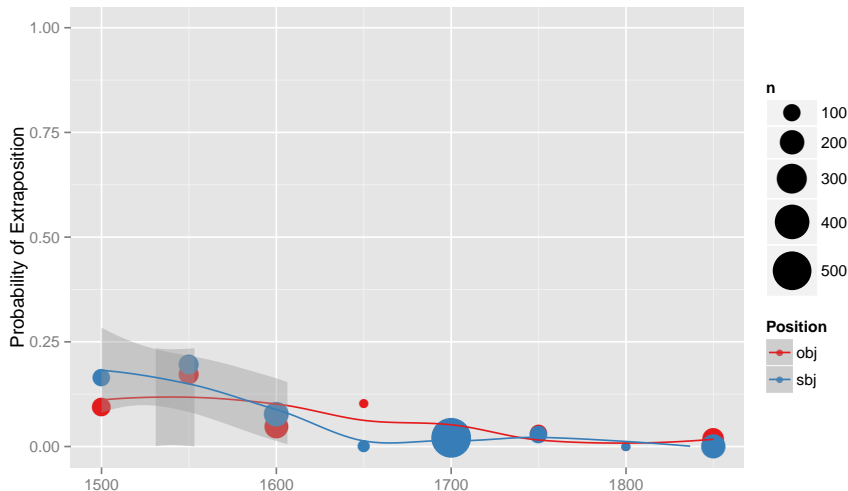# Icelandic, average weight over time (N = 3486)

Old/Middle French, over time (N = 8207)

French, average weight over time (N = 8207)

Portuguese, over time (N = 2398)

Portuguese, average weight over time (N = 2398)

Change in Relative Clauses

- Why the change?

- Relative clause extraposition has become severely restricted in modern Portuguese (Cardoso, 2011, 2012); cause and effect?

- Perhaps a small selectional advantage has asserted itself in the usage-overlap, but can only do so slowly because the overlap is limited.

- Drift; the population (of utterances) is more likely to fixate on the majority variant over time through random death, given enough time (Moran 1958; see Nowak 2006 for an overview and references).

- **Note:** topicalization in English has slowly declined from 1750 through the 20th c. (A. Kroch, p.c.).

Conclusions

Conclusions

- Within syntax, only one formal account of optionality is available, the same one that accounts for language change: Competing Grammars (i.e. decision points in the selection of grammatical formatives).

- All categorical variation/optionality/change = Blocking Effect, Competing Grammars

- Blocking Effect = selection, P. of Contrast (and a domain of specialization)

- Thus, all categorical variation/optionality/change is reduced to interactions of Competing Grammars, selection, Principle of Contrast, and inherent mathematical properties of possible domains of specialization.

Conclusions

Conclusions

- Competing Grammars result in replacement, specialization, or "stable variation".

- The latter is (only) the result of mapping categorical variation onto a continuous dimension of specialization.

- An acquisition simulation shows how stable variation can emerge under a minimal assumptions about the Principle of Contrast.

- Even so, the most stable variation may not be entirely stable.

- It is possible and desirable to extend this formal account to other domains of variation, like morphology and phonology.

Conclusions

Further Work

- We have a hypothesis about which factors influence specialization vs replacement, as they do not **appear** to be deterministic; find a way to test it!

- Look into other data sets for within-speaker categorical variation specialized along a continuous dimension.

- Look at the set of very slow changes in more detail (where we expected stability), to control for as many factors as possible.

- Figure out how to appropriately parameterize phonological variation.

Acknowledgements

References I

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Bailey, Laura, Joel C. Wallenberg, and Wim van der Wurff. 2012. Embedded yes/no questions: reanalysis and replacement. Paper presented at Presented at the 2012 Annual Meeting of the Linguistics Association of Great Britain (LAGB), University of Salford.

Baker, Mark. 2008. The macroparameter in a microparametric world. In *The Limits of Syntactic Variation*, ed. Theresa Biberauer, 351–374. John Benjamins.

Borer, Hagit. 1984. *Parametric syntax: case studies in Semitic and Romance languages*. Dordrecht: Foris Publications.

References II

Cardoso, Adriana. 2011. Variation and change in the syntax of
relative clauses: new evidence from Portuguese. Doctoral
Dissertation, University of Lisbon.

Cardoso, Adriana. 2012. Extraposition of restrictive relative
clauses in the history of Portuguese. In *Parameter Theory
and Linguistic Change*, ed. Charlotte Galves, Sonia Cyrino,
Ruth Lopes, Filomena Sandalo, and Juanito Avelar, number 2
in OSDHL, 77. Oxford University Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework.
In *Step by step: Essays on minimalist syntax in honor of
howard lasnik*, ed. Roger Martin, David Michaels, and Juan
Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

References III

Clark, Eve. 1987. The Principle of Contrast: A constraint on language acquisition. In *Mechanisms of language acqusition*, ed. Brian MacWhinney, The 20th Annual Carnegie Symposium on Cognition. Lawrence Erlbaum Associates, Inc.

Clark, Eve. 1990. On the pragmatics of contrast. *Journal of Child Language* 17:417–431.

Fruehwald, Josef, Jonathan Gress-Wright, and Joel C. Wallenberg. 2010. Phonological Rule Change: The Constant Rate Effect. In *Proceedings of 40th Meeting of the Northeast Linguistic Society (NELS) 40*.

Fruehwald, Josef, and Joel C. Wallenberg. 2013. Optionality is Stable Variation is Competing Grammars. *Presented at 25th Scandinavian Conference of Linguistics, Formal Ways of Analyzing Variation (FWAV) Workshop* .

References IV

Galves, Charlotte, and Helena Britto. 2002/2013. The Tycho Brahe Corpus of Historical Portuguese. Department of Linguistics, University of Campinas (Unicamp), Online Version, http://www.tycho.iel.unicamp.br/~tycho/.

Ingason, Anton Karl, and Laurel MacKenzie. 2011. "heaviness" as evidence for a derive-and-compare grammar. Poster presented at The 19th Manchester Phonology Meeting, May 19–21 2011.

Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2005. Penn-helsinki Parsed Corpus of Early Modern English. Size 1.8 Million Words.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.

References V

Kroch, Anthony S. 1994. Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K. Beals et al et al.

Kroch, Anthony S., Beatrice Santorini, and Ariel Diertani. 2010. Penn Parsed Corpus of Modern British English. Size ∼ 950000 words.

Kroch, Anthony S., and Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English. CD-ROM. Second Edition. Size: 1.3 million words.

Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1:85–97.

References VI

Markman, Ellen M., and Gwyn F. Wachtel. 1988. Children's
   use of mutual exclusivity to constrain the meanings of words.
   Cognitive Psychology 20:121–157.

Martineau, France, Paul Hirschbühler, Anthony Kroch, and
   Yves Charles Morin. 2010. Modéliser le changement : les
   voies du français, Département de français. University of
   Ottawa, http://www.arts.uottawa.ca/voies/voies_fr.html.

Moran, Patrick Alfred Pierce. 1958. Random processes in
   genetics. In *Mathematical Proceedings of the Cambridge
   Philosophical Society*, volume 54, 60–71. Cambridge Univ
   Press.

Nowak, Martin A. 2006. *Evolutionary dynamics: exploring the
   equations of life*. Cambridge, MA: Belknap Press of Harvard
   University Press.

Introduction   Blocking and Contrast   Competing Grammars   Stable Variation   Conclusions   **References**
000000   000   00   00
      00   0
      00000   0000
      0000000000000

References VII

Pintzuk, Susan. 1991. Phrase structures in competition: Variation and change in Old English word order. Doctoral Dissertation, University of Pennsylvania.

Prince, Ellen. 1985. Fancy syntax and shared knowledge. *Journal of Pragmatics* 9:65–81.

Prince, Ellen. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. *Syntax and semantics* 281–302.

Prince, Ellen. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. In *Texas linguistics forum*, chapter 8. University of Texas, Austin: Citeseer.

Santorini, Beatrice. 1992. Variation and Change in Yiddish Subordinate Clause Word Order. *Natural Language & Linguistic Theory* 10:595–640.

References VIII

Speyer, Augustin. 2008. Topicalization and clash avoidance: On
    the interaction of prosody and syntax in the history of
    English with a few glimpses at German. Doctoral
    Dissertation, University of Pennsylvania.

Speyer, Augustin. 2010. *Topicalization and stress clash
    avoidance in the history of english*. Topics in English
    Linguistics. Berlin/New York: Mouton de Gruyter.

Taylor, Ann. 1994. Variation in past tense formation in the
    history of english. *University of Pennsylvania Working
    Papers in Linguistics* 1:143–159.

References IX

Taylor, Ann, Arja Nurmi, Anthony Warner, Susan Pintzuk, and Terttu Nevalainen. 2006. York-Helsinki Parsed Corpus of Early English Correspondence. Compiled by the CEEC Project Team. York: University of York and Helsinki: University of Helsinki. Distributed through the Oxford Text Archive.

Taylor, Ann, Anthony Warner, Susan Pintzuk, and Frank Beths. 2003. The York-Toronto-Helsinki Parsed Corpus of Old English Prose.

Wallenberg, Joel C., Anton K. Ingason, Einar F. Sigurðsson, and Eiríkur Rögnvaldsson. 2011. Icelandic Parsed Historical Corpus (IcePaHC). Version 0.9. Size: 1 million words. URL http://www.linguist.is/icelandic_treebank.