One Very Slow Change: the decline of relative clause extraposition

Joel C. Wallenberg
Newcastle University
joel.wallenberg@ncl.ac.uk

Bad Data Workshop

March 15, 2016

Some Challenges for Dealing with Bad (diachronic) Data

- Resolution of our observations
- Time-depth
- Sources of noise
- Cross-linguistic comparison
- Acquisition's link to diachronic patterns of incrementation, diffusion, etc.

Some Challenges for Bad Data

- **Our hypotheses about the problems of language change are becoming sufficiently precise that they are increasingly difficult to test with the data at hand, even as data becomes better.**
- Weinreich, Labov, & Herzog (1968) identified, among others, the following problems for a theory of language change, which are also questions about acquisition:
  - **actuation** of new variants.
  - **constraints** on possible changes (and possible stability).
  - **transition** between language states (maybe reducible to the others?).
  - **embedding** of the change in larger linguistic and extralinguistic systems.

Production data for linguistic research (e.g. parsed corpora)

- **Linking Hypotheses for linguistic structure:** link linguistic theory to the quantitative patterns of language use, by finding clear quantitative predictions of hypotheses about linguistic structure.
- **Statistical Reasoning:** in probabilistic (stochastic) data, find the patterns underlying the noise (without accidentally modeling the noise).
- Example of a successful linking hypothesis is the "Constant Rate Hypothesis" (Kroch, 1989, and many subsequent):
  - If one syntactic parameter has a number of different surface realizations, and that parameter changes over time, then changes in the different surface contexts will occur at the same rate.

Production data for linguistic research

Linking Hypotheses for acquisition: link our theory of acquisition to the quantitative patterns of language use, by finding clear quantitative predictions of specific hypotheses about the acquisition process.

Outline

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 object by another.
2. Optionality (within a grammar?)
   - Diachronically stable variation between grammatical processes.

Introduction

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

- We expect variation (apparent optionality) between two grammatical forms to be diachronically unstable.
- True(st) optionality = stable(st) variation, and its difference from usual language change must be explained by some aspect of language use outside of the linguistic variable itself (**embedding**), which slows the change.
- We argue that it depends on the mathematical character of some dimension (i.e. external to the variable's grammar) with which the variation interacts (**embedding**).
    - Partial specialization of categorical variants along a continuous dimension.

Blocking and Contrast

"Blocking Effect" (Aronoff, 1976)

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

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

"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).

Blocking = Contrast + Evolutionary Dynamics

- 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).
- Either one variant has an advantage, and so **replaces** the other (often following a logistic function; Nowak, 2006).
- Or neither variant has an advantage (or much of one), in which case random walk and drift (which can also lead to **replacement**).
- But in linguistic doublets, there is always the counter-pressure to resolve the doublet through **specialization**, because 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.
- E.g. two different adjunction sites.

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 (the resulting meanings overlap).
- 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 Contrast plus evolutionary dynamics yield the outcomes **replacement** and **specialization**.

Whether replacement occurs depends on whether specialization can arrest it.

Consequence: Blocking and Contrast

- If specialization occurs, its effect depends on the **domain of specialization**. A change can be:
  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 **imperfectly 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 be slowed by **imperfect specialization**.

A Very Slow Change

- One consequence of our overall hypothesis is that some things that didn't look like change turn out to be.
- Relative clause extraposition is a change in progress, but a very slow one (Wallenberg, Accepted, under revision, 2013b; Fruehwald and Wallenberg, In preparation).
  - It has been mischaracterized as syntactic optionality.
- The study used the same coding query (with minor adaptation) on 7 parsed diachronic corpora (4 language histories).
- Both the time-depth and cross-linguistic dimensions were necessary in order to discover the change.
- Only because we had both dimensions were we able to observe (and confirm) the slowest syntactic change discovered to date.

Case Study: Relative Clause Extraposition

French

(1) 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

(2) All had now been tried [which either threats or promises, forbearance or fatherly chastisement, could effect].
(PPCMBE, FROUDE-1830,2,2.20; date: 1830)

Hypotheses for the diachronic study

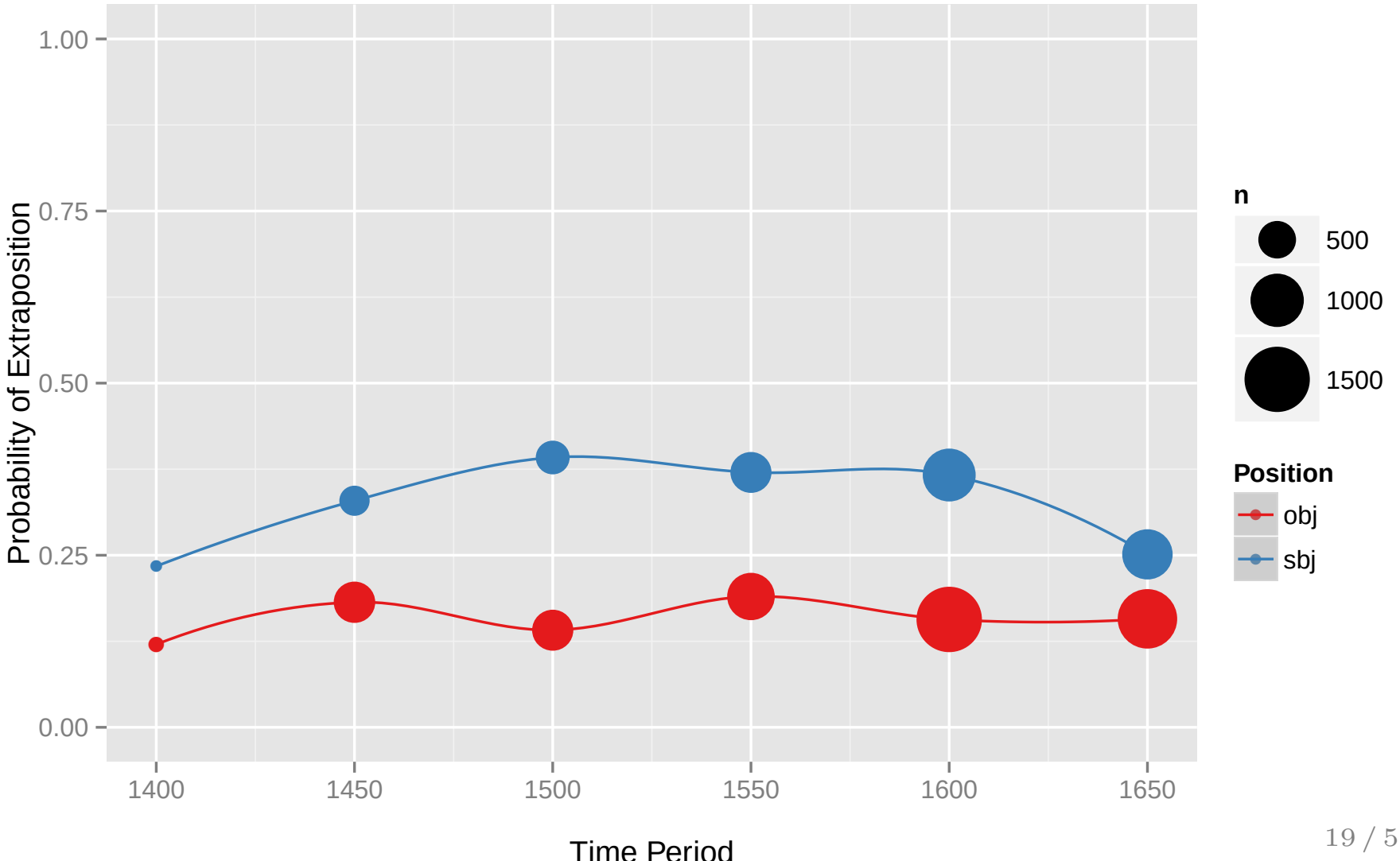
- **Hypothesis 1:** Relative clause position (a binary variable) is specialized along a continuous dimension, weight, and so it should be **nearly** stable, but not entirely stable.
- **Hypothesis 2:** All IE relative clauses derive historically from clause-adjoined relatives (Kiparsky, 1995).
- **Hypothesis/Suggestion 2':** The old clause-adjoined kind are still around, in the form of extraposition, and the change Kiparsky proposed hasn't finished yet (Wallenberg, Accepted, under revision).

Specialization within individuals

- 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.

Hypothesis: individual speakers treat weight as a continuous variable, with extraposition specialized imperfectly along it (as suggested by Ingason and MacKenzie, 2011).

All PCEEC, over time (N = 8073 clauses)



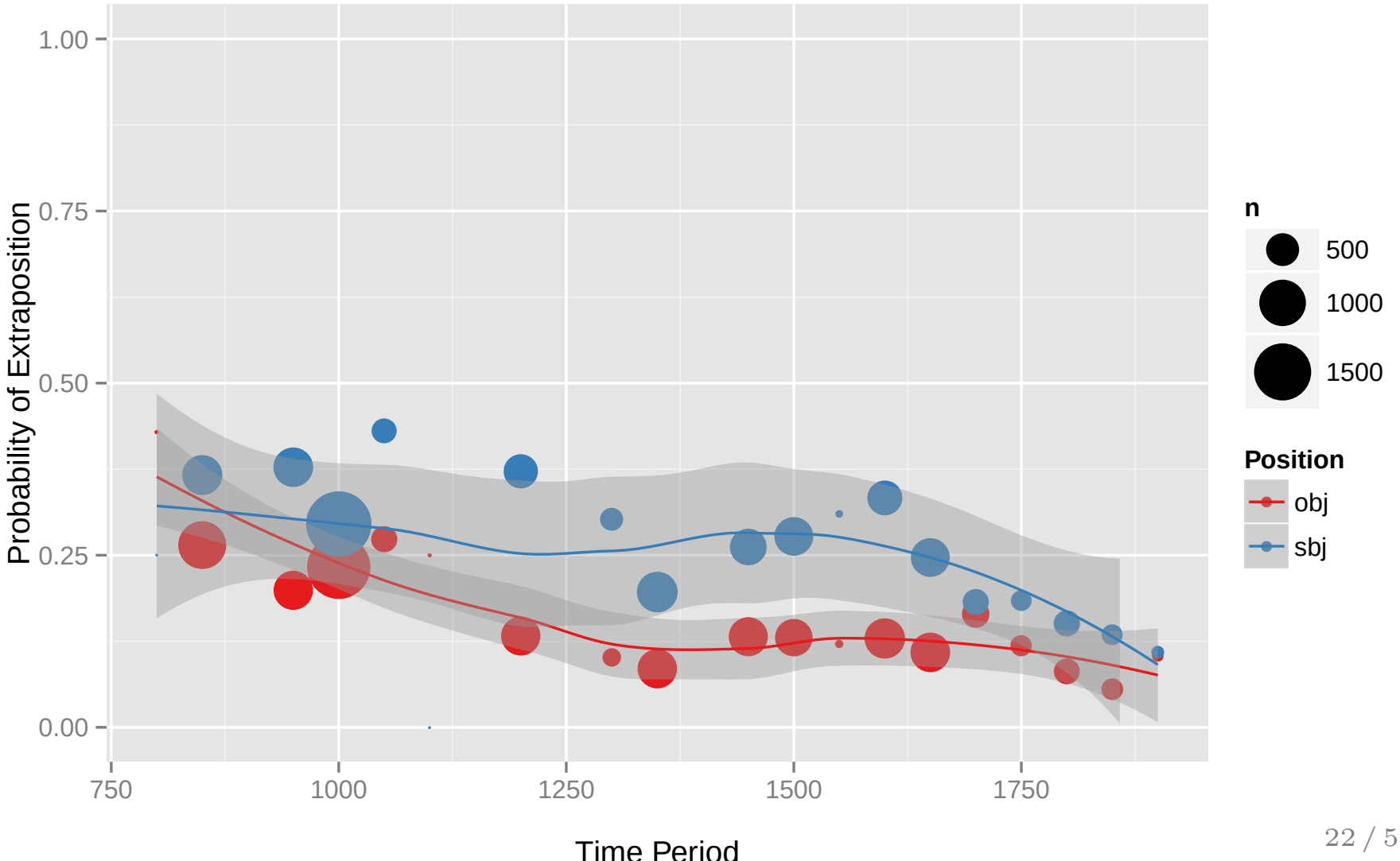
Diachronically, Crosslinguistically

- **Empirical Result:** relative clause extraposition is a change in progress.
  - Extraposition is dying, but **extremely slowly**.
- Same coding query (with minor adaptation) on 7 parsed diachronic corpora (4 language histories).
- Both the time-depth and cross-linguistic dimensions were necessary in order to discover the change.

Diachronically, Crosslinguistically

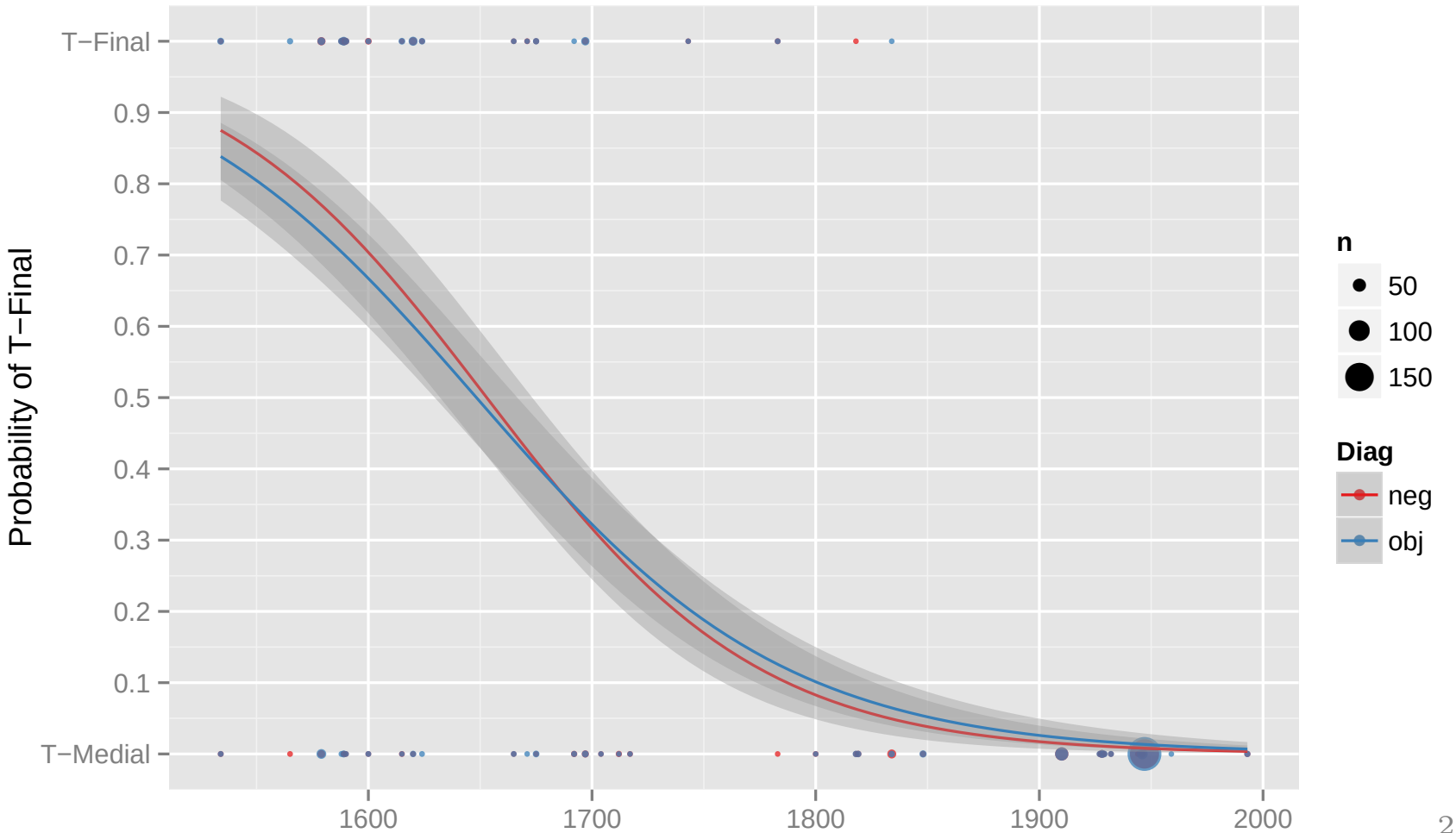
- **English:** YCOE (Taylor et al., 2003), PPCME2 (Kroch and Taylor, 2000), PPCEME (Kroch et al., 2005), PPCMBE (Kroch et al., 2010).
- **Icelandic:** IcePaHC (Wallenberg et al., 2011).
- **Old/Middle French:** MCVF Corpus (Martineau, Hirschbühler, Kroch, & Charles Morin, 2010).
- **Historical Portuguese:** Tycho Brahe Corpus of Historical Portuguese (Galves and Faria, 2010).

English, over time (N = 18530 clauses)



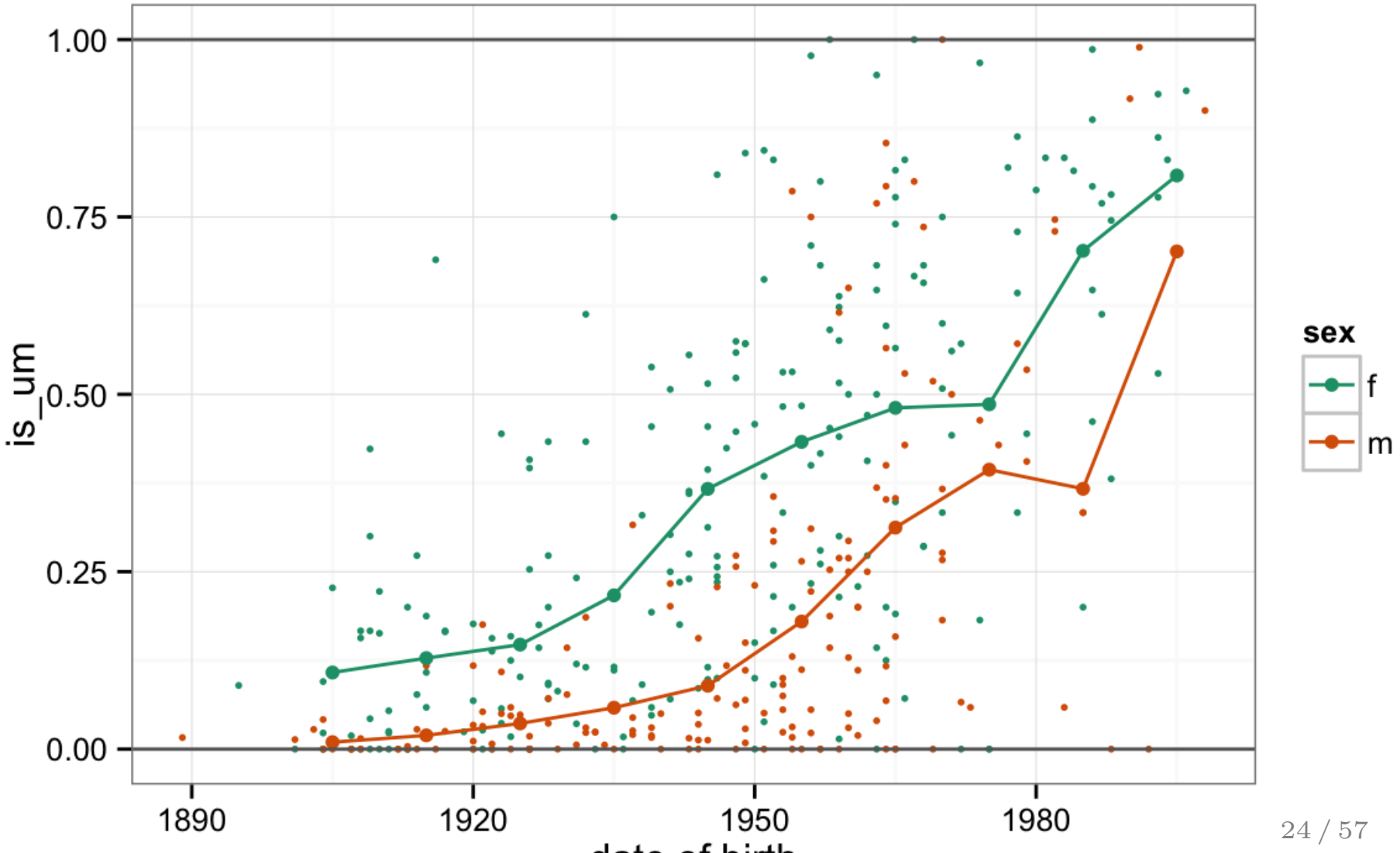
cf. Yiddish Tense, N = 1030 clauses

Wallenberg (2013a), building on Santorini (1993)

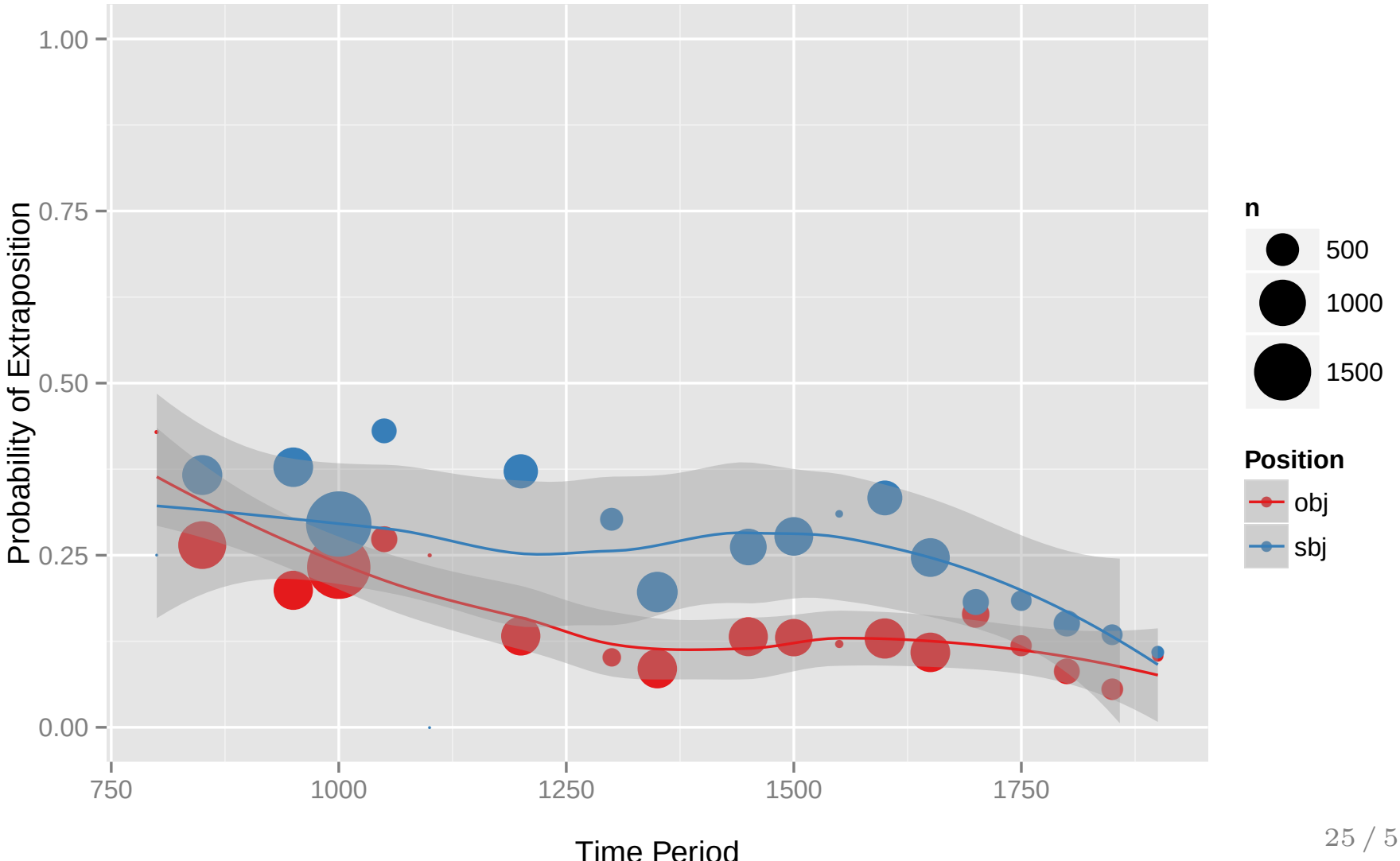


Pause-fillers, *um* vs. *uh*, $n_{speakers} = 308$

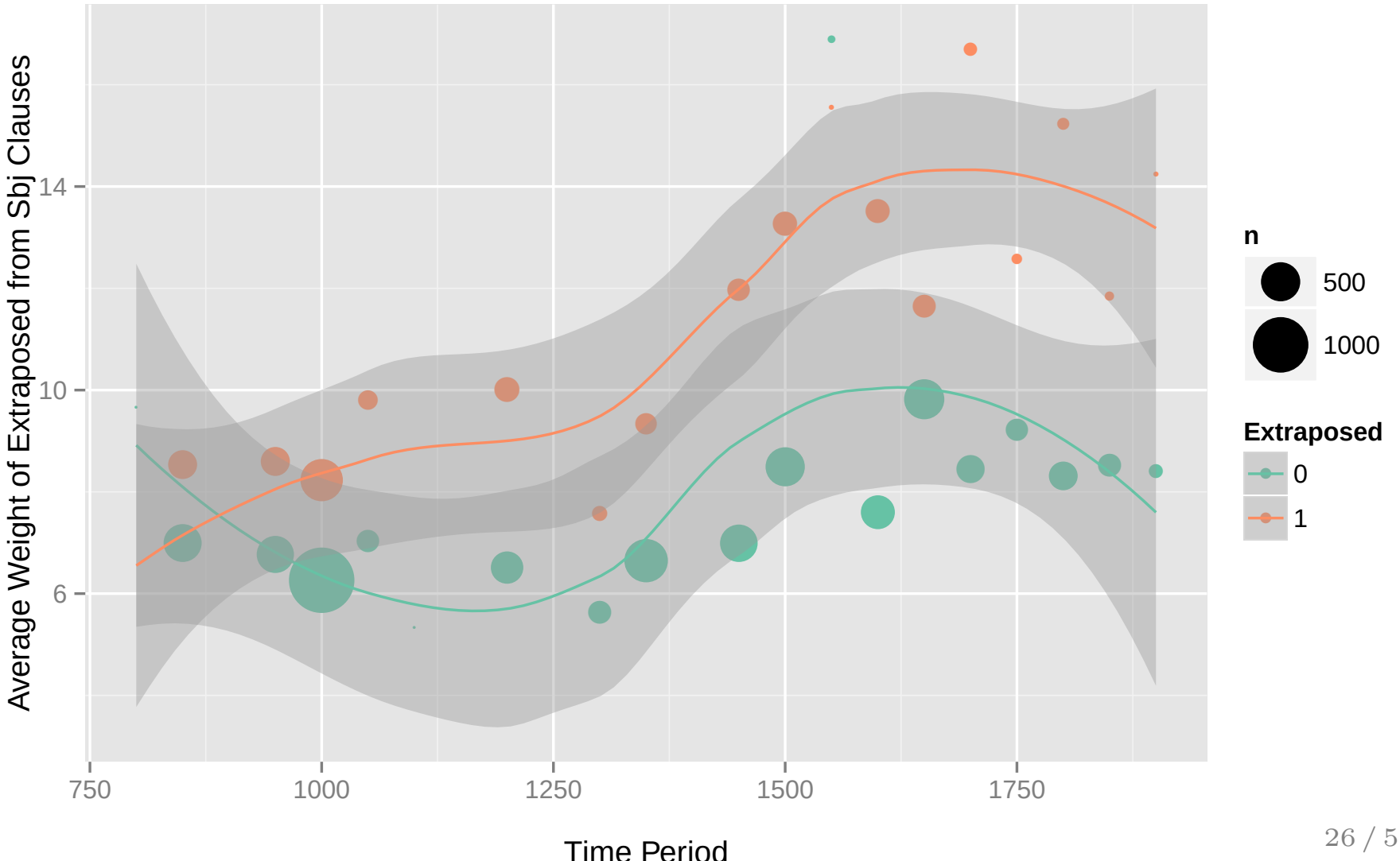
(Fruehwald 2015., using *Philadelphia Neighborhood Corpus*)



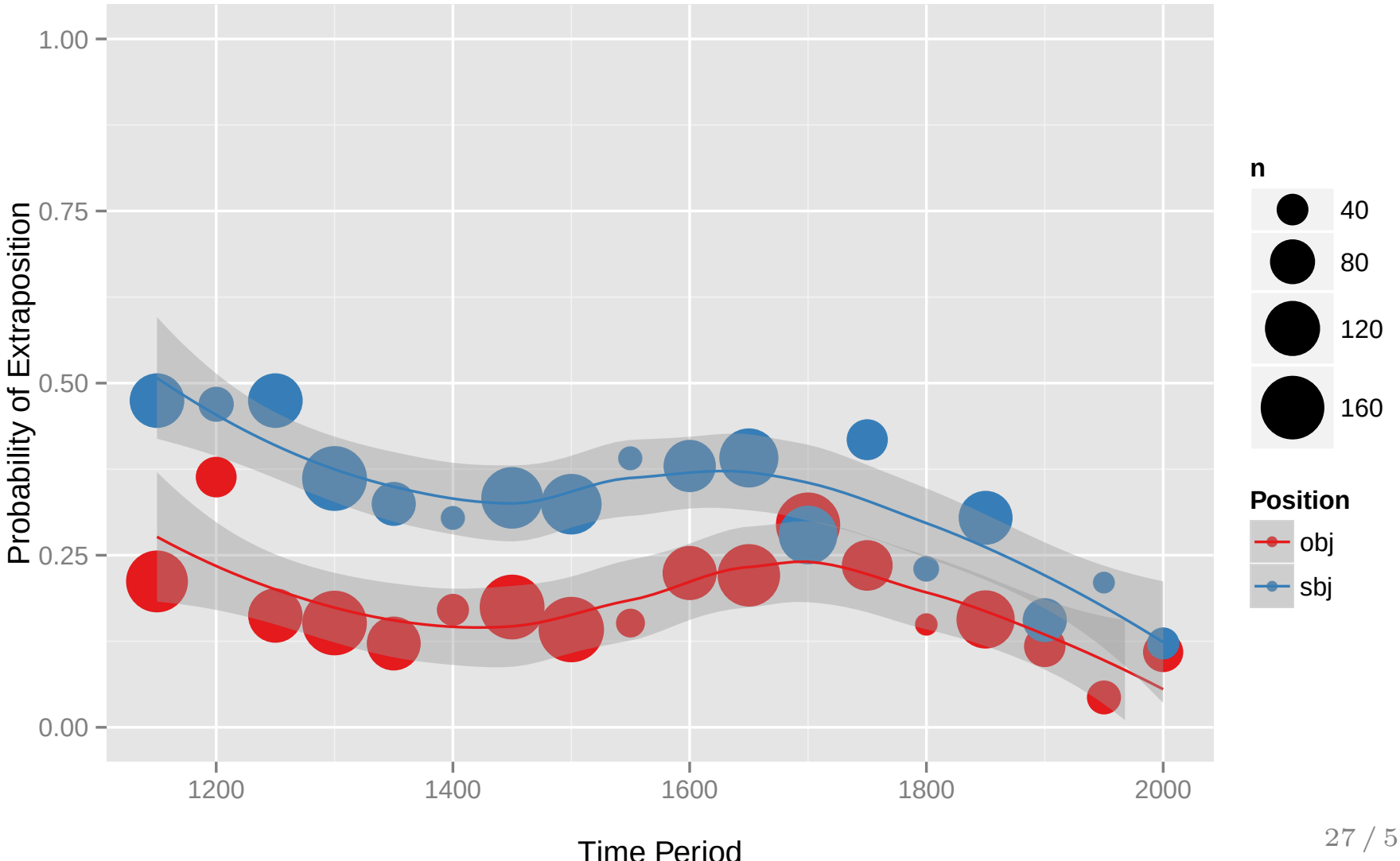
English, over time (N = 18530 clauses)



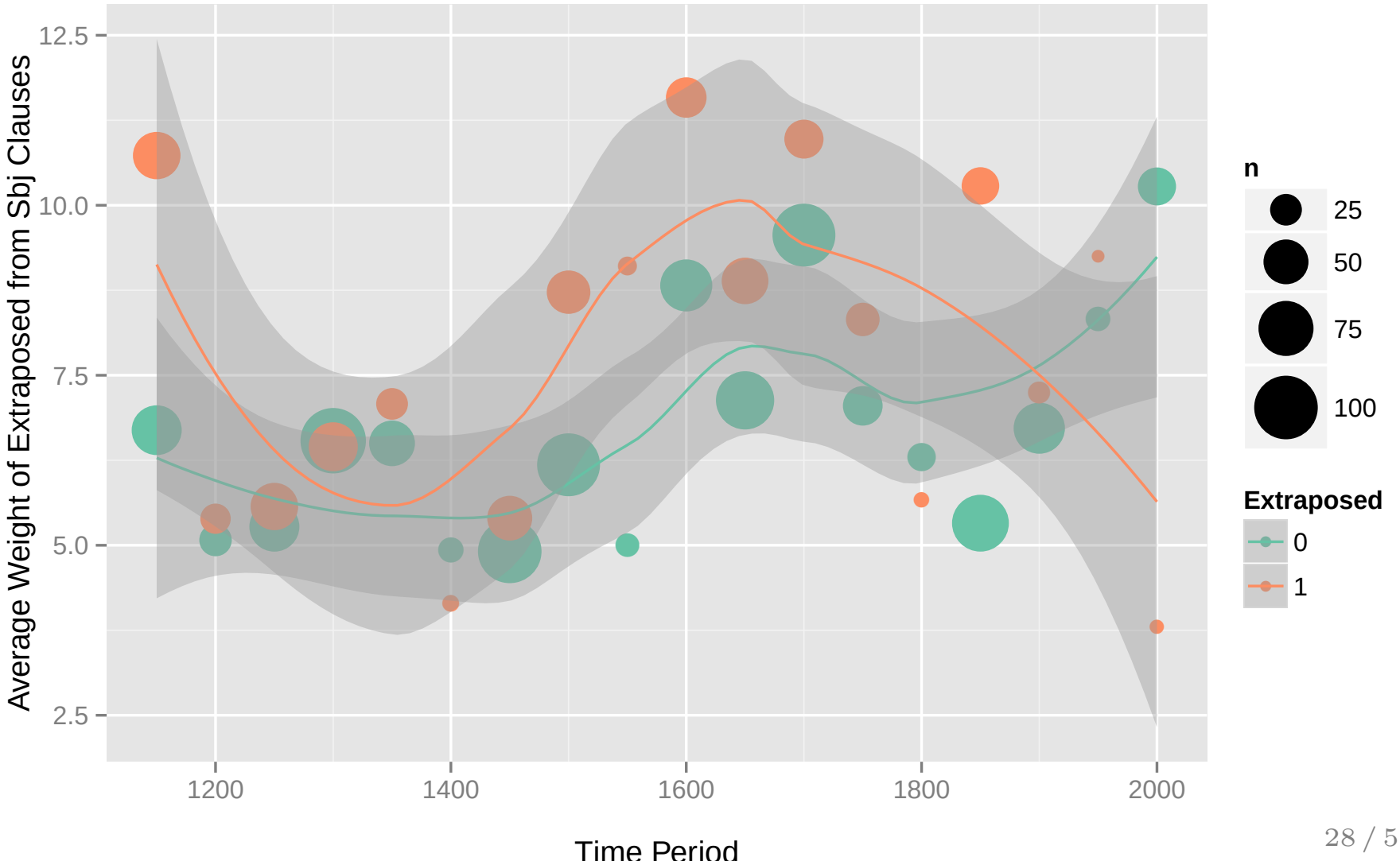
English, average weight over time (N = 18530)



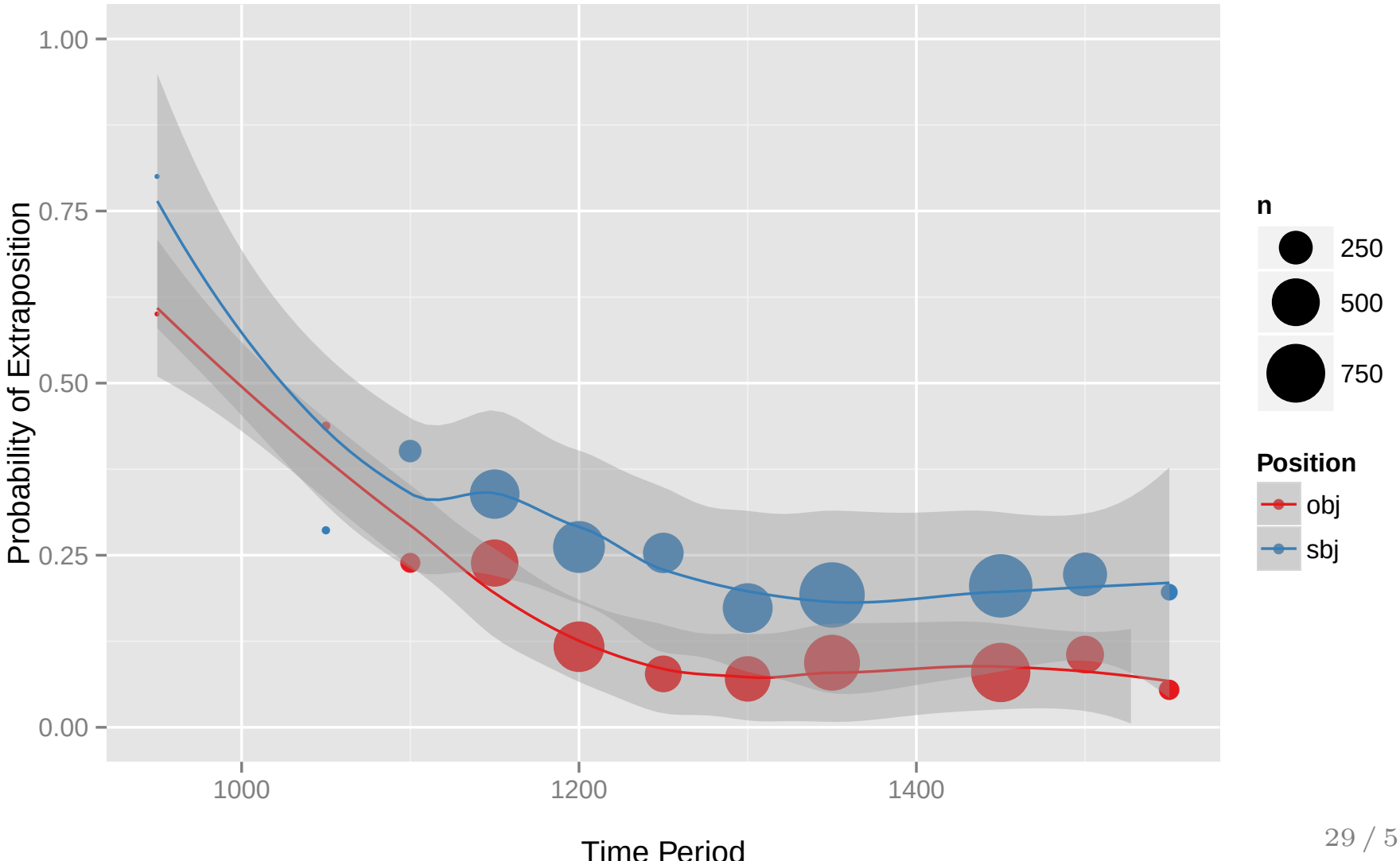
Icelandic, over time (N = 3486)



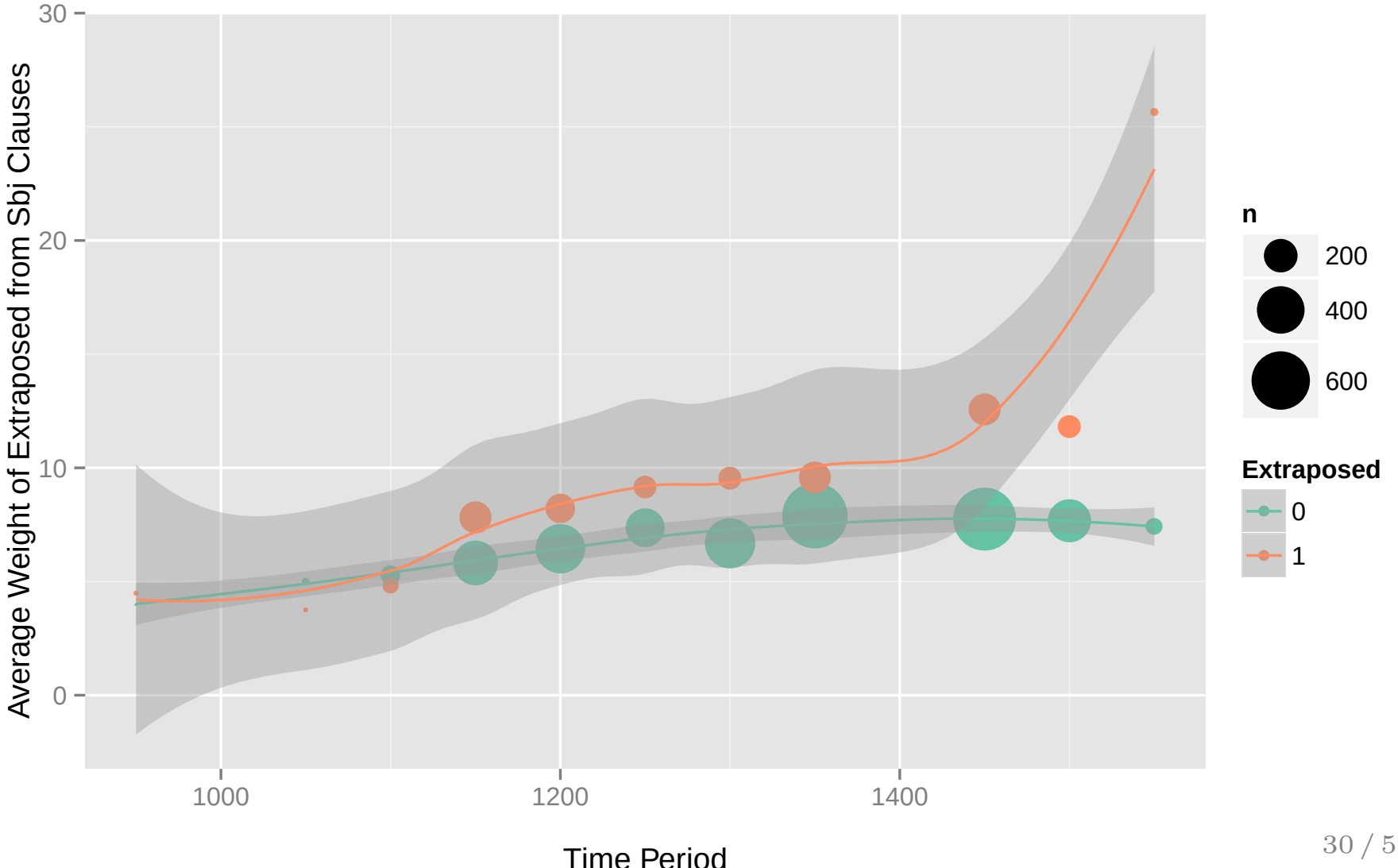
Icelandic, average weight over time (N = 3486)



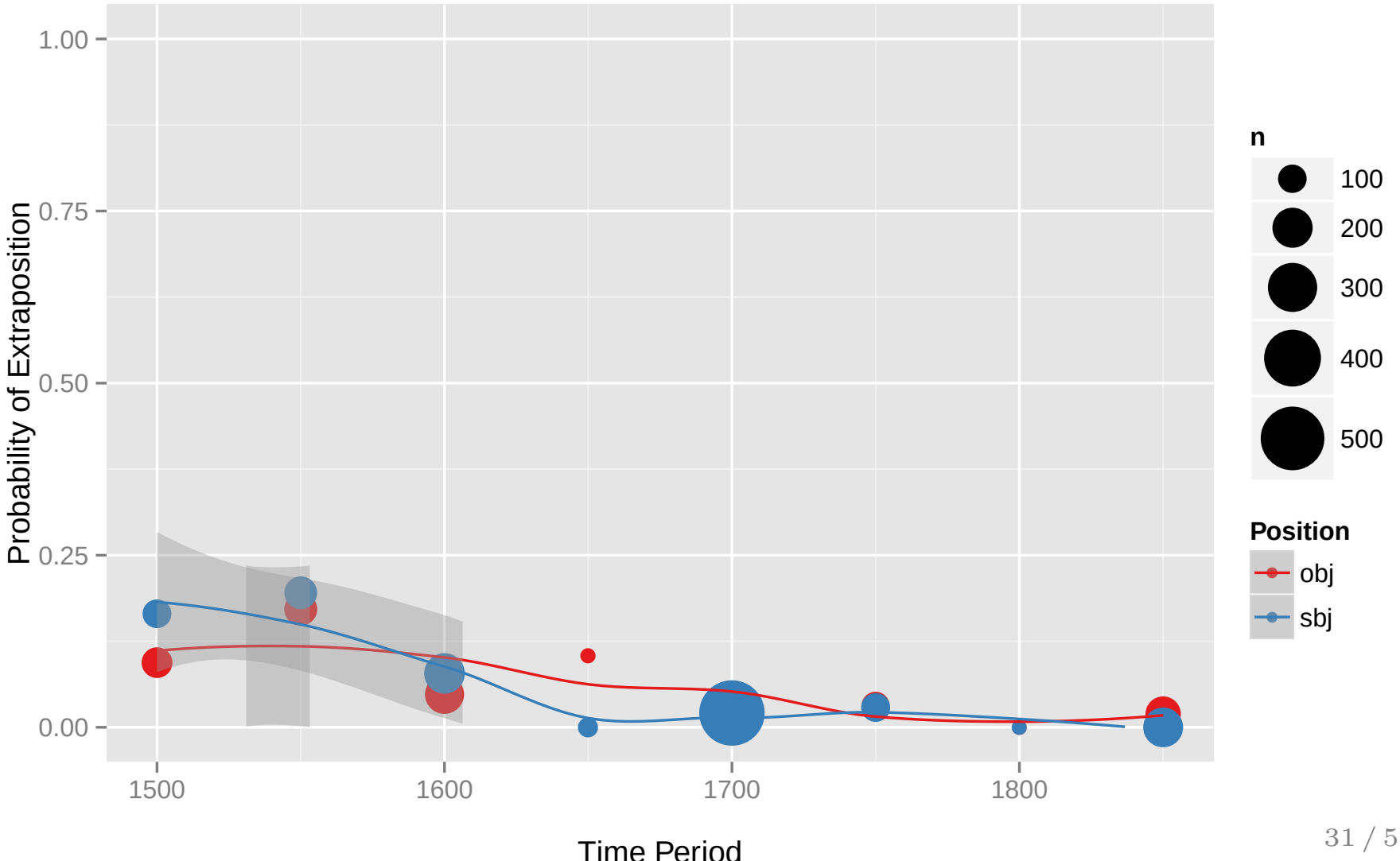
Old/Middle French, over time (N = 8207)



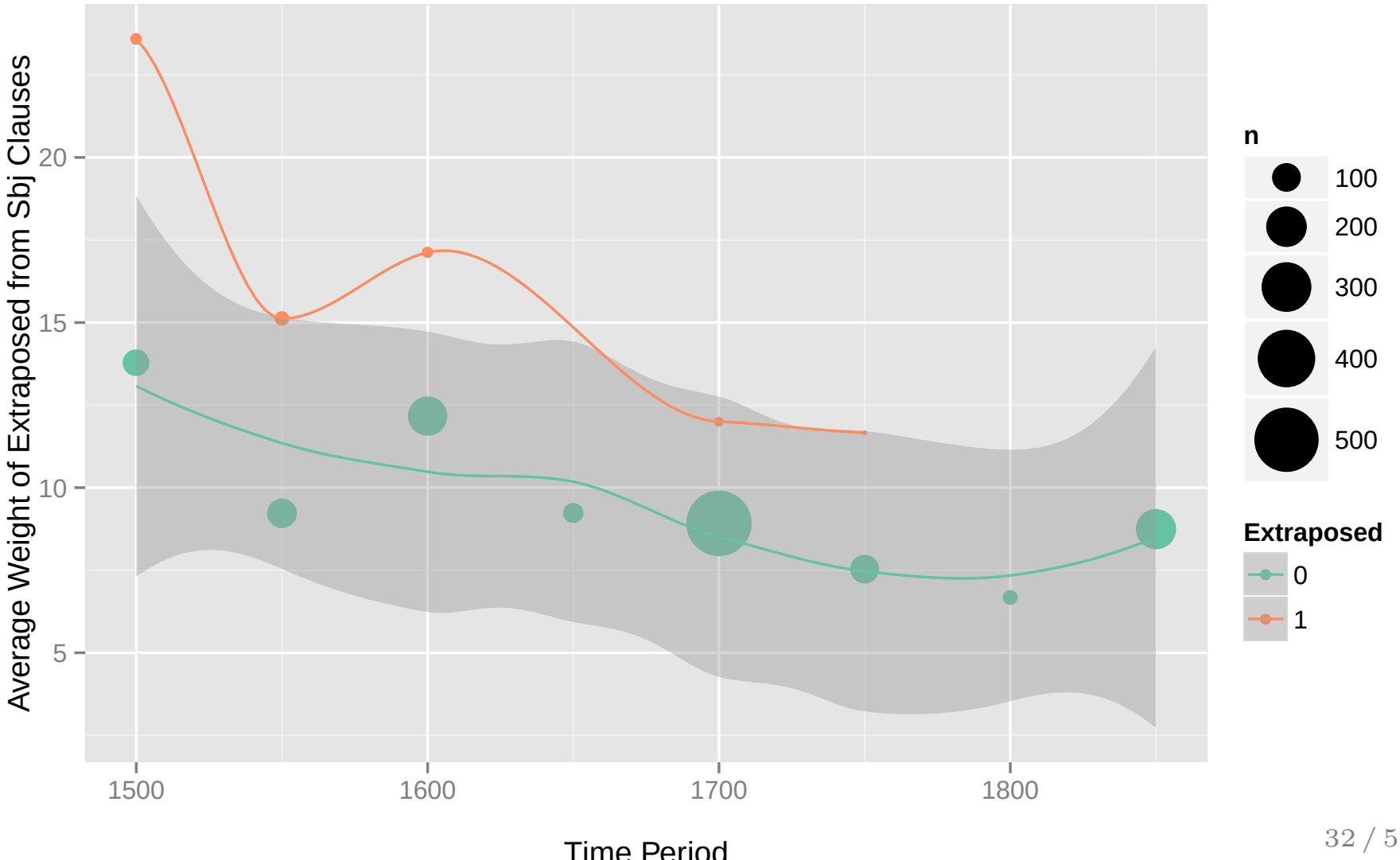
French, average weight over time (N = 8207)



Portuguese, over time (N = 2398)



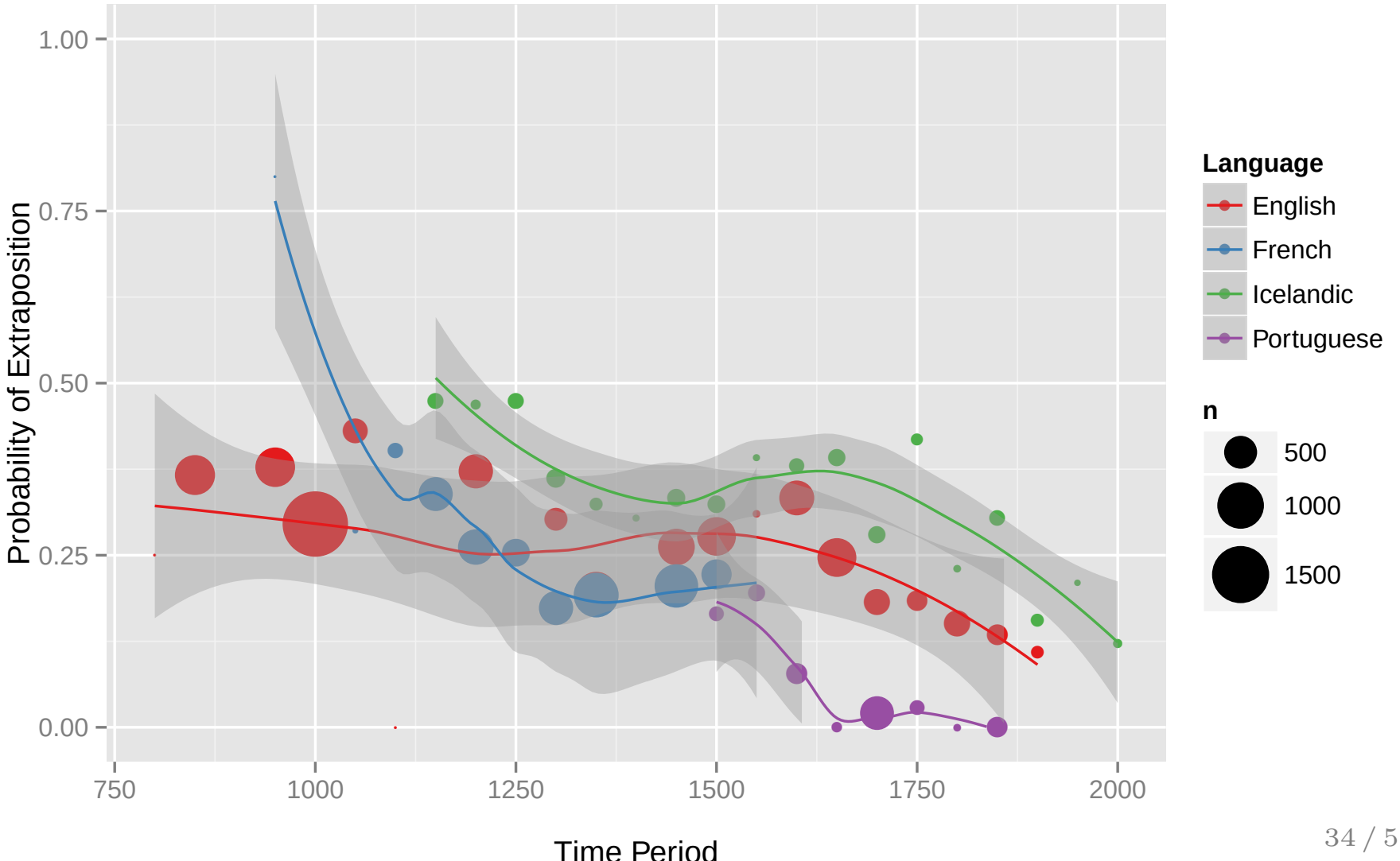
Portuguese, average weight over time (N = 2398)



Statistical characteristics of the change

- The slope of the decline over time is close to zero for English and Icelandic, but nonzero ($p < 10^{-6}$ for model comparison), and very small (e.g. -10^{-2}) for French and Portuguese (based on logistic regression controlling for weight and other factors). (**noise** and **resolution**)
- Weight has a significant effect in each language, but the effect doesn't change over time.
- At least Icelandic and English show the same rate of change (a kind of Constant Rate Effect?), $p = 0.47$:

Four Languages (Subj Ex), over time



What is in competition?

- If we take the view of Culicover and Rochemont (1990), then relative clauses right-adjoin to various phrasal categories, and extraposition is adjunction to a higher phrasal category.
- A principle of interpretation ("Complement Principle") determines how deeply embedded the interpretation can be.
- The doublet, the competing grammars, are actually **competing adjunction sites with the same interpretation**.

The origin of the slow change?

(3) By God's blessing I calculate that the Spirit of Dishonesty shall not get dominion over me; nor the Spirit of Despondency, nor any other evil spirit; **in which case all will and must be well**.
(Letter by Thomas Carlyle, date: 1835; ID CARLYLE-1835,2,266.176 in PPCMBE)

(4) Nowadays, however, flowers can be arranged in various styles – some flat, some slightly raised, some bunched boldly in certain places and forming the piece de resistance of the whole work – **all of which variations depend upon the artistic perceptions of the operator**.
(*Commercial gardening...*, date: 1913;ID WEATHERS-1913,1,9.217 in PPCMBE)

The origin of the slow change?

(5) **yó mártyaḥ śíśīte áty aktúbhir**,
which mortal sharpen-Mid-Sg overly nights-INSTR,
má naḥ sá ripúr īśata
not us-GEN that trickster dominate-Subj3Sg
"As for the mortal who makes himself too sharp by night, may that trickster not gain power over us"
(RV 1.36.16, cited in Kiparsky, 1995, 156)

Competition? Sauerland (2003)

- Sauerland argues that English relatives are often ambiguous between a **raising** structure (raising of the "head" NP) and an adjoined **matching** structure, based on binding Principles A and C.
- Interestingly, relative clause extraposition seems to be only compatible with the **matching** structure, the one that does not allow A binding of the "head" NP.

(6) That picture of John that he likes a lot was just published.

(7) That picture of himself that John likes a lot was just published.

(8) ?* That picture of himself was just published that John likes a lot.

Competition? Sauerland (2003)

Can the loss of extraposition be competition between the matching and raising structures, with matching slowly losing and taking extraposition with it?

Competition: matching vs. raising?

- Can speakers choose the matching structure in order to extrapose?
  - If so, then raising should specialize for in situ, matching for extraposition...why does extraposition decline?
- Sauerland claims raising is bad with indefinite NPs, which is where extraposition is most natural, so raising replacing matching would have little effect on extraposition...unless this is part of the slowness, and/or part of the specialization of raising for *in situ*?
- And there's still competing adjunction sites within the matching structure. A 3-way competition?
- Is there any evidence to the learner for the matching and raising structures?

Summary: Change in Extraposition

- Cross-linguistic Constant Rate Effect: clear for English and Icelandic changes ($p = 0.47$)
  - Not so clear for Romance, though data is much sparser and the corpora are far less stylistically balanced
- Seeing the change in all the languages makes us more confident that the effect is real in each one.
- The change is not at all observable without considerable time-depth.
- **Why the change?** After actuation, extraposition and *in situ* are competing variants in use, so there can't not be a change, even with partial specialization.
  - Specialization can only be partial along the (continuous) weight dimension.
- Perhaps Kiparsky (1995) identifies a change that goes beyond Proto-Germanic, though hard to test.

More Nuanced Questions

- Any role for Sauerland's competing relative clause structures? How would we observe them diachronically?
- Relative clause extraposition is restricted in its syntactic context in modern Portuguese (Cardoso, 2011, 2012); specialization? cause and effect?
- How likely is a change like this to result from fixation by drift in a finite population (of utterances) and speakers through random death? (Moran 1958; see Nowak 2006 for an overview and references).

Specialization in Acquisition: active or passive?

1. Identify a domain of specialization:
   - **Actively**, creatively, by the child innovating *de novo*?
   - **Passively**, though random sampling of finite populations of utterances?
2. For categorical variants along categorical dimensions, uncouple tracked frequencies of variants.
3. Specialization completes in a categorical dimension:
   - **Actively**, by punishing one variant when the other is rewarded?
   - **Passively**, by allowing whatever evolutionary dynamics hold in the different contexts play out, whether the outcome is different or not?

Specialization in Acquisition: active or passive?

1. For categorical variants along continuous dimensions, uncouple tracked means of variants in the dimension of specialization.
2. Specialization completes in a continuous dimension:
   - **Actively**, by moving the means of variants away from each other?
   - **Passively**, by allowing the means to potentially move away from each other?

Specialization in Acquisition as a Challenge for Bad Data

- This model of specialization is an extension of Yang (2000, 2002).
- Can we observe specialization, and the choosing of domains of specialization diachronically?
- Can we observe it in acquisition?
  - Production data may not be good enough here.
- Can we observe either well enough to test active vs. passive hypotheses?

Conclusion: embedding and transition

- Relative clause extraposition is in decline, and we can just about link it to:
  - a larger theory of syntactic representations (**grammatical embedding**)
  - a theory of language specialization in acquisition (**extragrammatical embedding**).
- There is a level of syntactic detail that is currently beyond our/my ability to test (?).
- There is a level of detail in the theory of specialization that is currently beyond our/my ability to test (?).
- The information on **transition** is based on very high-quality, expensive data (parsed diachronic corpora, beyond the ability of an individual to produce).
- It is still noisy, and we would like finer resolution as we tackle questions above, and e.g. selection vs. drift.

Acknowledgements

Thank you first to Josef Fruehwald for working out many of these ideas with me, and to Anthony Kroch and Betsy Sneller for much discussion of these issues. Thanks to Anton Karl Ingason for use of his CS queries. Also Aaron Ecay, Ricardo Bermúdez-Otero, Caitlin Light, Laurel Mackenzie, George Walkden, and anonymous reviewers.

Extraposition Study:
`github.com/joelcw/tyneside/tree/master/extraposition`

References I

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

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.

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

References II

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

Culicover, Peter W., and Michael S. Rochemont. 1990. Extraposition and the Complement Principle. *Linguistic Inquiry* 23–47.

Fruehwald, Josef. 2015. Social meaning and information theory. Poster Presentation.

Fruehwald, Josef, Jonathan Gress-Wright, and Joel C. Wallenberg. 2013. Phonological Rule Change: The Constant Rate Effect. In *Proceedings of the 40th Annual Meeting of the North East Linguistic Society*, ed. Seda Kan, Claire Moore-Cantwell, and Robert Staubs. GLSA (Graduate Linguistic Student Association).

References III

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* .

Fruehwald, Josef, and Joel C. Wallenberg. In preparation. Optionality is Stable Variation is Competing Grammars.

Galves, Charlotte, and Pablo Faria. 2010. 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.

References IV

Kiparsky, Paul V. 1995. Indo-European origins of Germanic syntax. In *Clause structure and language change*, ed. Adrian Battye and Ian G. Roberts, Oxford Studies in Comparative Syntax, chapter 6, 140–169. Oxford University Press.

Kroch, Anthony, Beatrice Santorini, and Ariel Diertani. 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.

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.

References V

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.

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. Size ∼ 994000 words. University of Ottawa, `http://www.arts.uottawa.ca/voies/voies_fr.html`.

References VI

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.

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

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

References VII

Santorini, Beatrice. 1993. The rate of phrase structure change in the history of Yiddish. *Language Variation and Change* 5:257–283.

Sauerland, Uli. 2003. Unpronounced heads in relative clauses. In *The interfaces: Deriving and interpreting omitted structures*, ed. Kerstin Schwabe and Susanne Winkler, volume 61 of *Linguistik Aktuell*, 205–226. Amsterdam: John Benjamins.

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

References VIII

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. `http://www-users.york.ac.uk/~lang22/YcoeHome1.htm`.

Wallenberg, Joel C. 2013a. Scrambling, LF, and Phrase Structure Change in Yiddish. *Lingua* 133:289–318.

References IX

Wallenberg, Joel C. 2013b. A unified theory of stable variation, syntactic optionality, and syntactic change. In *Presented at the 15th meeting of the Diachronic Generative Syntax (DIGS) Conference*. University of Ottawa.

Wallenberg, Joel C. Accepted, under revision. Extraposition is disappearing. October, 2014 Accepted, under revision.

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`.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for historical linguistics*, ed. W. Lehmann and Y. Malkiel, 95–195. Austin, TX: University of Texas Press.

References X

Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12:231–250.

Yang, Charles D. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.