HISTORICAL SYNTAX

Extraposition is disappearing

JOEL C. WALLENBERG

Newcastle University

This study describes a change in which relative clause extraposition is in the process of being lost in English, Icelandic, French, and Portuguese. This current change in progress has never been observed before, probably because it is so slow that it is undetectable without the aid of multiple diachronic parsed corpora (treebanks) with time depths of over 500 years each. Building on insights from Kiparsky (1995), the study shows that the change may date as far back as the innovation of Proto-Germanic and Proto-Romance relative clauses, as these varieties differentiated from Proto-Indo-European. It also shows that the unusually slow speed of the change is due to partial specialization of the construction along the dimension of prosodic weight, following the argument made at greater length in Fruehwald & Wallenberg 2016. Finally, the change is shown to have important consequences for the syntax of extraposition, supporting the adjunction analysis of Culicover and Rochemont (1990). The article also discusses the implications of Sauerland's (2003) analysis of English relative clauses, and while modern English data supports his analysis, the diachronic extraposition data is not yet fine-grained enough to bear on the 'raising' analysis of relatives in general. This is identified as an important question for further research on this change.*

Keywords: syntax, language change, Indo-European, evolutionary dynamics, treebanks

1. INTRODUCTION. This article presents important preliminary results demonstrating the existence of a syntactic change that is many times slower than any other that has previously been reported: relative clause extraposition is being lost in both Germanic and Romance languages.

Furthermore, this decline in extraposition has been underway for over a thousand years, and while it is nearing completion in modern Portuguese, the change appears to still be ongoing at present. Since the results are still preliminary and much of their analysis is still unclear, the main goal of this article is to carefully describe a type of syntactic change that presents problems for both synchronic syntactic theory and the theory of language change, but not to present a comprehensive analysis of every aspect of this change, or this construction, at the present time. I hope this will be the first of many studies of this interesting and important phenomenon, and I consider the suggestions here to be a starting point, not anything near the last word on the subject.

The construction at issue is exemplified in 1a. It is currently in decline, and possibly on a slow path to extinction.

(1) a. Another method … may be employed which is less open to the above objection.
    (John Strutt's scientific papers, date: 1890; STRUTT-1890,3,356.396 in PPCMBE, Kroch et al. 2010)

  b. Another method which is less open to the above objection may be employed.
    (constructed)

In order to define the scope of enquiry more precisely, this study considers the extraposition of finite relative clauses, that is, CP-sized clauses containing a gap, where the gap is interpreted as referring to some NP antecedent. (The antecedent is standardly re-

ferred to as the 'relative head', even though it is not a head at all, in the technical sense.)
The extraposed relatives this study considers are ones where the entire finite clause
containing the gap is displaced from the relative head, whether this is analyzed as
movement or something else, surfacing at the right edge of the *v*P or higher TP. (A more
detailed description of how these are located in the historical corpora can be found in
§2.1.) Extraposed relatives have often been analyzed as an instance of rightward move-
ment followed by right-adjunction, but they have also been analyzed (following Kayne
1994) as being the result of the leftward movement of other material (see Baltin 2006
for an overview of various analyses and their shortcomings). Our particular concern
here is with extraposed relatives that are interpreted with respect to an NP in the subject
of the higher clause, as in 1a, though some data on extraposition from objects will also
be presented.

   The goal of this article is to present some potentially important preliminary findings,
and some possible interpretations of those findings; the goal is not to present the defin-
itive analysis of extraposition or of the diachronic change. That being said, the simple
result this article presents, that extraposed relatives are in decline, has the following
consequence: any analysis of relative clauses (and extraposition) must take into account
the fact that the extraposed and in-situ variants are, in some sense, 'competing gram-
mars' (in the sense of Kroch 1989, 1994, inter alia); they must be in competition in lan-
guage use, as one variant is in the process of replacing the other. The fact that such slow
changes exist, and that their effects on synchronic language use can be demonstrated,
should cause all researchers in synchronic syntax to question whether all purported
cases of 'optional' movements are in fact changes in progress. Perhaps 'optional' has no
coherent meaning in grammar, outside of the meaning of change in progress.

   The article is organized as follows. First, I report quantitative data that relative clause
extraposition is being lost in the following languages: English, Icelandic, French, and
Portuguese, using seven different parsed diachronic corpora (i.e. treebanks). The changes
are surprisingly similar across the languages, particularly in Icelandic and English, and I
pose the question as to what a crosslinguistic 'constant rate effect' (Kroch 1989) might
mean. Next, in §3.1 I discuss the idea that the extraposed and in-situ variants are 'com-
peting grammars', syntactic variants that are mutually exclusive and competing in lan-
guage use, and suggest some consequences this has for the analysis of relative clauses.
In particular, I consider whether the competing grammars we see could be the competing
structures for relative clauses identified by Sauerland (2003), and/or whether they are
competing adjunction sites, building on Culicover and Rochemont (1990). I conclude
that while Sauerland's analysis is relevant for understanding the change, his two relative
clause analyses are not the relevant competing grammars. In §3.3, I discuss the fact that
the extraposed variant is specialized for particularly heavy relative clauses, and why this
specialization slows the change down to such a degree that it is only observable over a
period of a thousand years. Finally, §3.4 builds on Kiparsky (1995) and presents the pos-
sibility that this change began back in prehistory, at a time when true relative clauses were
first innovated in Romance and Germanic, and has been underway ever since.

2. DECLINE IN RELATIVE CLAUSE EXTRAPOSITION. The data in §2.2 below shows a re-
markably slow decline in the frequency of relative clause extraposition from subject
and object DPs in four languages: English, Icelandic, French, and Portuguese. I origi-
nally collected this data to test the hypothesis that the frequency of relative clause ex-
traposition was stable over time in a number of languages; as is shown below, though
the change is slow enough to be unobservable without data from a period of many hun-
dreds of years, this hypothesis can be soundly rejected.

2.1. METHODS. It is worth noting that the type of data presented below is entirely unobtainable without diachronic parsed corpora (i.e. treebanks), ones that have been annotated at a very high level of accuracy (i.e. hand-corrected by expert syntacticians, not just automatically parsed). For this study, the following parsed diachronic corpora were used: for English, the York-Toronto-Helsinki Corpus of Old English Prose (YCOE; Taylor et al. 2003), the Penn-Helsinki Parsed Corpus of Middle English 2 (PPCME2; Kroch & Taylor 2000), the Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME; Kroch et al. 2005), and the Penn Parsed Corpus of Modern British English (PPCMBE; Kroch et al. 2010); for Icelandic, the Icelandic Parsed Historical Corpus (IcePaHC; Wallenberg et al. 2011); for Old and Middle French, the MCVF Corpus (Martineau et al. 2010); and for historical Portuguese, the Tycho Brahe Corpus of Historical Portuguese (Tycho Brahe; Galves & Faria 2010).

Finite clauses that contain a relative clause modifying a subject or object were extracted from the corpora using CorpusSearch coding queries (Randall 2000/2013) and were coded for the following variables: the response variable of whether or not the relative clause was extraposed, whether it modified a subject or object DP, its weight in number of words, whether or not the clause occurred in reported speech, and the date of the text the clause appeared in.[1] These variables were controlled for in all subsequent statistical analysis. Extraposed finite relatives are annotated in the same way across all of the corpora above, as a 'CP-REL' inside of an NP, which dominates a particular kind of trace, labeled 'ICH'. The trace is coindexed with another 'CP-REL' node that occurs displaced from the trace position, to the right of some other clause-level material. This annotation is only given when the relative unambiguously has a gap corresponding to some relative head NP, and when the relative is clearly displaced from that NP in the string. (Please see the annotation manuals for the corpora, and my actual queries, for more detail.)

All coding queries and the extracted data sets are available at the following public *git* directory: https://github.com/joelcw/tyneside/tree/master/extraposition/.

2.2. RESULTS. For the four languages I investigated—English, Icelandic, French, and Portuguese—there is a slow but steady decline in the rate of relative clause extraposition over the course of each language's written history (given the corpora we currently have), compared to the use of in-situ relative clauses. This is shown in Figures 1–4.[2] Modeling with mixed-effects hierarchical logistic regression confirms this informal impression of the data.

As can be seen in the figures for English (Fig.1) and Icelandic (Fig. 2), the decline in extraposition is so slow that it could easily have been missed if the corpora of English (collectively) and Icelandic were not at sufficient time depth (over 1,000 years of data for each); any time window of only a few hundred years would likely be too small to detect such a change at all.

For comparison, I conducted the same study on the Parsed Corpus of Early English Correspondence (PCEEC; Taylor et al. 2006), which covers a period from the fifteenth

---

[1] If no precise date of composition was available, as was the case for many Old English and Old Icelandic texts, I used an estimated date of composition wherever possible, or a manuscript date in cases where there is no clear consensus on the date of composition. In all cases, I have followed the dates provided by the documentation of the diachronic corpora, following the philological sources they cite.

[2] For data sets in the form of lists of codes extracted from the corpus, please see files ending in 'fixed.ooo' in the github directory linked above—that is, https://github.com/joelcw/tyneside/tree/master/extraposition/queriesandoutput/*fixed.ooo. The code lists are easily importable into R, and the same directory contains the queries, ending in '.c', that produced the codes and say how to interpret them.
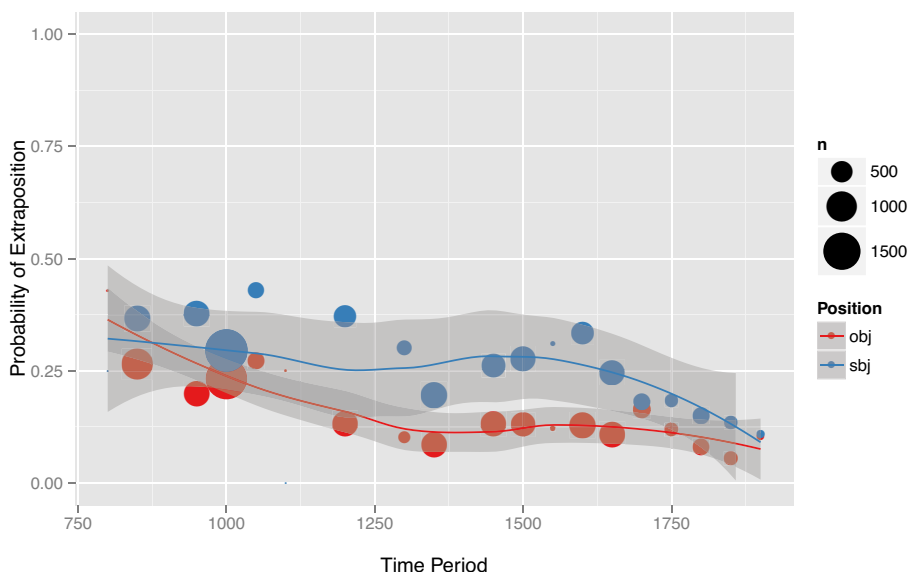
FIGURE 1. Declining proportion of relative clause extraposition (vs. in situ) from subject and object positions, early Old English prose through Modern English. $N = 18,530$ relative clauses. Lines are LOESS.
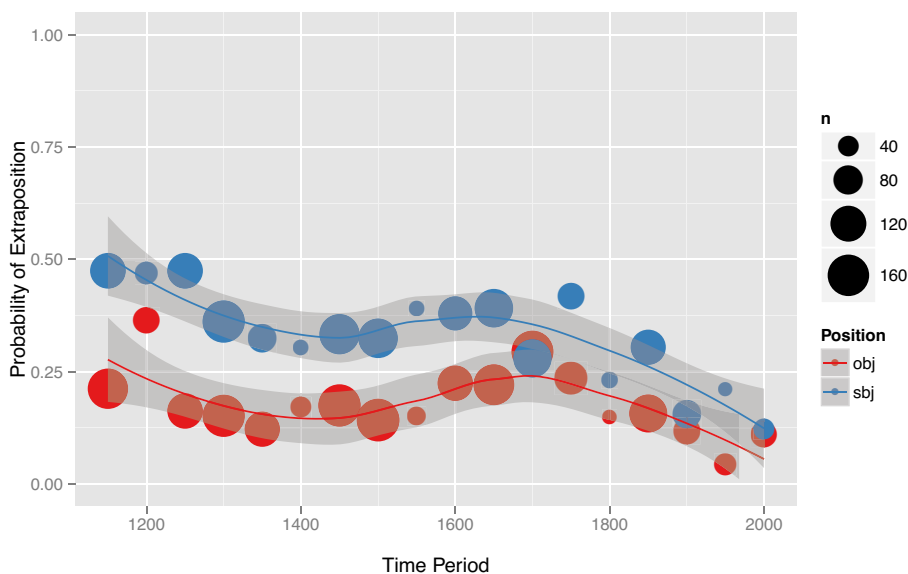

FIGURE 2. Declining proportion of relative clause extraposition (vs. in situ) from subject and object positions, early Old Icelandic prose through modern Icelandic. $N = 3,486$ relative clauses. Lines are LOESS.

through seventeenth centuries. The results of that study make the frequency of extraposition appear to be stable over time, as seen in Figure 5.

The results in Fig. 5 match the results from the larger English corpora for the same time period (compare Fig. 1), but miss the change entirely; simply by accident, the 250-year time slice of the PCEEC happens to not reflect the change, which then becomes visible once we observe a larger diachronic window.
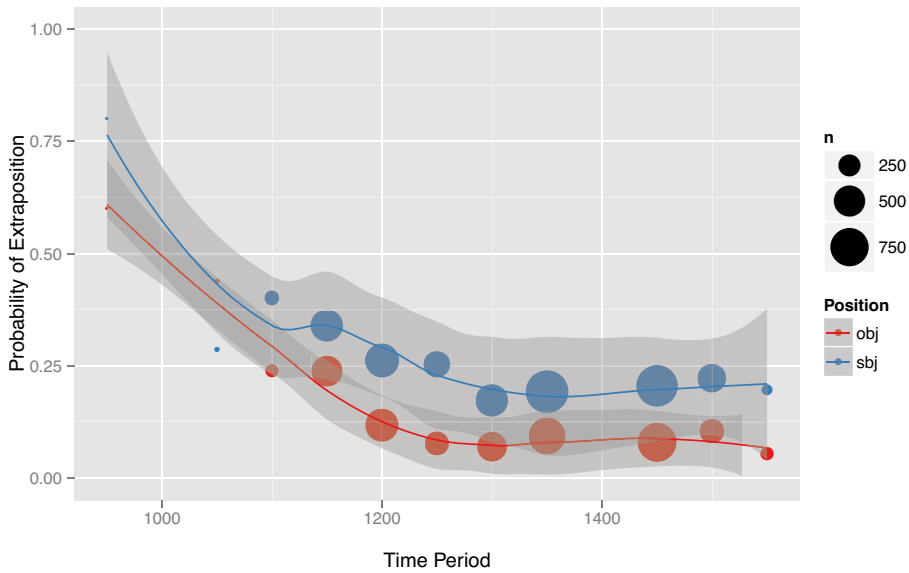
FIGURE 3. Declining proportion of relative clause extraposition (vs. in situ) from subject and object positions, Old and Middle French. $N = 8{,}207$ relative clauses. Lines are LOESS.
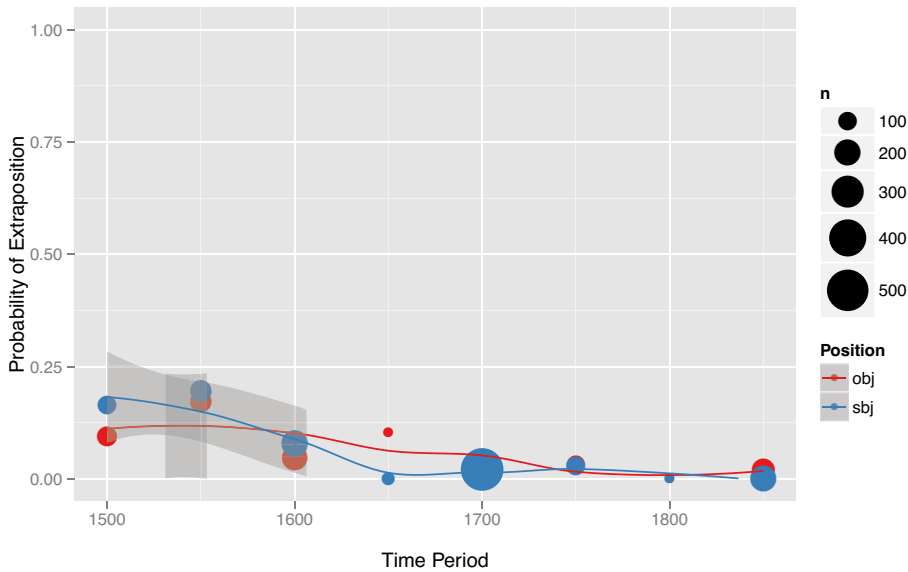
FIGURE 4. Declining proportion of relative clause extraposition (vs. in situ) from subject and object positions, fifteenth–nineteenth c. Portuguese. $N = 2{,}398$ relative clauses. Lines are LOESS.

Nevertheless, fitting a mixed-effects logistic regression model, with random slopes for individual texts, shows that the decline in extraposition is statistically significant in each of the languages;[3] text Year of composition (estimated in many cases, depending on manuscript evidence) is a significant factor in the model, showing that an interaction between text year and extraposition significantly improves model fit ($p < 0.00001$ for

[3] Models were fit using the package lme4 (Bates et al. 2011) and function glmer in R (R Core Team 2014).

FIGURE 5. Proportion of relative clause extraposition (vs. in situ) from subject and object
positions in the PCEEC. $N = 8{,}073$ relative clauses.

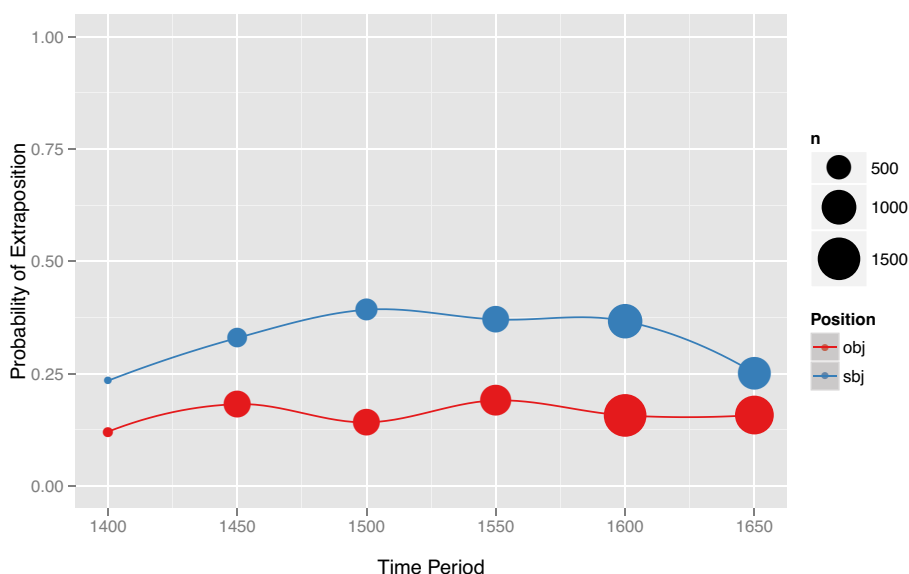all languages, AIC difference < 3, BIC difference < 7 for all languages except Por-
tuguese, where the model with Year had a lower AIC by 20 and a lower BIC by 8.59[4]),
controlling for these other variables: matrix/subordinate clause-type, simple text vs. re-
ported/quoted speech,[5] subject vs. object position, and weight of the subordinate clause
(i.e. length measured in number of words).[6] In other words, model comparison between
a model that contains a term for Year of text and a model that does not shows that in-
cluding Year substantially reduces deviance of the model for each of the languages,
which is a strong indication that the change I am suggesting is a real one, given statisti-
cal noise and other relevant factors.

For the Icelandic and French models, it was also possible to test whether there was a sig-
nificant interaction between Weight and Year, and there was not (for Icelandic: $p = 0.49$,
AIC difference = 1.6, BIC difference = 7.6; for French: $p = 0.89$, AIC difference = 1.9, BIC
difference = 1). This indicates that the effect of Weight on the probability of a speaker ex-
traposing the given relative clause was plausibly constant over time, even as the overall

---

[4] Because of the AIC and BIC, we should exercise some caution in interpreting the Portuguese results, be-
cause the model with Year loses significantly less information than the simpler model. However, the chi-
square comparison and associated p-value suggests there is a change, and the Portuguese data set is smaller
than the others.

[5] Clauses containing quoted speech in the texts are annotated with a special tag in these corpora.

[6] The mixed-effects models also controlled for possible interactions between Year and Position and Year
and Clause type (main or subordinate), except for in the Portuguese model, where including the interaction of
Year and Position would not allow the model to converge, for computational reasons. Including additional in-
teractions would not allow a mixed-effects model to converge, but I did fit a standard logistic regression with
all possible interactions, and the model comparison for including Year yielded the same results for all lan-
guages. For the Icelandic and French models, it was possible to include one additional interaction, Year and
Weight, but as I note below, this was not significant. In the saturated standard logistic model for Icelandic, I
also included Genre in the model; the way genre is balanced across time in IcePaHC makes it easier to con-
trol for statistically than in the other models. Note also that the numeric predictors of Year and Weight were
converted to Z-scores (i.e. scaled) in the mixed-effects model, which helped the algorithm to converge.

probability of extraposition was changing. This might indicate that the specialization of extraposition for heavier clauses had stabilized by the period we can observe in the corpus. Or, if extraposition is continuing to specialize throughout the time period, it must be doing so slowly enough that we cannot observe it statistically. (It was not possible to test this hypothesis for English and Portuguese, simply because the mixed-effects models would not converge with this additional interaction for computational reasons. Standard logistic regression did show this interaction to be insignificant, though.)

Looking a little more closely at the fitted mixed-effects logistic models, the terms for Year are estimated to have various negative coefficients, showing that the models estimate the probability of extraposition to be in decline over time, as one would expect. These are the model's estimates for how much the log-odds of extraposing a relative clause has declined per year. For the models discussed above, the estimated slopes for the effect of Year on the probability of extraposition were as follows: English: -0.42 ($p < 10^{-16}$), Icelandic: -0.11 ($p = 0.23$), French: -0.52 ($p < 10^{-16}$), and Portuguese: -1.04 ($p < 10^{-9}$). They all are associated with low p-values, aside from the slope from the Icelandic model—so that number should not be overly interpreted. The first thing that will strike the reader is that these numbers are small; compare, for example, Kroch's (1989) original estimates for the slopes of the rise of *do*-support (or loss of V-to-T) in English, which were all between -3 and -4.5. This change is more than three times slower than the *do*-support change in English. To take another example, the loss of Tense-final phrase structure in Yiddish is a slower change than the rise of English *do*-support, but it is still twice as fast as the loss of relative clause extraposition in any of the languages, except possibly Portuguese: Santorini (1993) estimates the loss of Tense-final clauses in Yiddish at approximately -1 (depending on the diagnostic). (And since Portuguese is the smallest of the data sets and is also at the very end of its change in the time period covered by the corpus, one might question how representative that slope is.)

Note that the slopes above are estimated for extraposition from subject and object positions together, and it is worth considering extraposition from subject alone. Analyzing subject and object positions in the same model could easily be problematic, since the effect of weight on relative clause extraposition (see §3.3) means that there could plausibly be an interaction between relative clause extraposition and HEAVY NP SHIFT (HNPS) of the object, which could affect slope estimates. Furthermore, any different behavior of HNPS, or the syntax of objects in general, in these four languages would interfere with the crosslinguistic comparison. HNPS is also likely to have interacted with independent phrase structure changes in the various languages (as Wallenberg 2015 shows for English), such that comparing the diachrony of the loss of relative clause extraposition across languages becomes problematic. Therefore, for the sake of caution and completeness, the slopes of the loss of relative clause extraposition from subject position, when modeled alone, are as follows: English: -0.36 ($p < 10^{-9}$), Icelandic: -0.37 ($p < 0.001$), French: -0.32 ($p < 0.001$), and Portuguese: -1.24 ($p < 10^{-6}$). As in the models above, model comparison shows a significantly better fit when including Year, indicating that a slow change is indeed taking place, even looking at extraposition from subject position alone. Note also that these slopes, aside from the Portuguese, are very similar to each other, suggesting that the change may be occurring at the same rate in three of the languages. This is difficult to test, however, as mixed-effects models fit across all of the language data sets will not converge.

However, if we fit a standard logistic model across English and Icelandic, constraining it to have the same slope parameter for the decline in both (controlling for the same factors as above), this crosslinguistic model does not do significantly worse than a

model that allows different slope terms for the two languages ($p \approx 0.47$).[7] This is reminiscent of the CONSTANT RATE EFFECT (Kroch 1989, Pintzuk 1991, Santorini 1993, among others), which states that the different surface realizations of the same underlying linguistic object will all change with the same slope as that underlying linguistic object changes. Though the situation here is not one of comparing different linguistic contexts, but rather different populations of speakers (i.e. Icelandic and English), perhaps the same basic logic applies: the same trajectory of change could indicate that the same forces are underlying the change in both populations, both in terms of what is driving the change, and in terms of which basic linguistic object is undergoing change. But of course, it is tricky to interpret a constant rate effect across populations of speakers, since in order for it to be interpretable, the population dynamics of those speech communities would have to be comparable.

Any standard logistic crosslinguistic model that includes the French and Portuguese data fits significantly worse when it assumes the same slope for the change over time across the languages. This is true if French and Portuguese are modeled together, separately from Icelandic and English, or if any combination of languages including either French or Portuguese are modeled together ($p \leq 0.01$ in all cases). This either means that these languages show a similar-looking but ultimately different change (at least at the time points we are observing), or that some other feature of those data sets is obscuring what should be the same slope as in Icelandic and English. Since the text-sampling methods were designed to control for genre effects over time in the Icelandic and English corpora, but this was not done in the same way for the Tycho Brahe and MCVF corpora (and the latter is the only corpus that contains significant amounts of poetry as well as prose), we cannot rule out the possibility that the estimated slopes for French and Portuguese have been obscured by factors that are extraneous to the issue under investigation. It is also interesting to note that once random slopes were added for individual texts in the mixed-effects models, the estimated slope for the French change in extraposition from subject position was very similar to the English and Icelandic slopes for extraposition from subject position, even though we could not test this more rigorously for technical reasons.

3. DISCUSSION.

3.1. EXTRAPOSITION AS COMPETING GRAMMARS. One consequence of the study above is that extraposition, which is usually described as an optional movement process producing multiple output options within a single grammar, is in fact behaving diachronically more like two coexisting linguistic objects, one replacing another over time just as in the classic cases of 'competing grammars' (Kroch 1989, 1994, Santorini 1992, and subsequent). As Kroch (1994) discusses at length with respect to the 'blocking effect' (an expanded understanding of the principle first proposed in Aronoff 1976), situations of competing grammars are inherently diachronically unstable for general cognitive reasons. In all four languages (with the possible exception of modern Portuguese), even though the loss of extraposition began at least a thousand years ago, the change is still in progress at the present time. Thus, the seeming optionality between extraposed and nonextraposed relative clauses is really most simply described as the code-switching that is characteristic of a change in progress (the definition of 'competing grammars').

---

[7] As I mentioned above, it was not possible for a mixed-effects model to converge on the data when all languages were included. But note that the slope estimates for Year from the mixed-effects models in the previous paragraph are nearly identical for English and Icelandic, and very close for French.

Once two structures (or two versions of a functional head, following the Borer-Chomsky conjecture; Borer 1984, Kroch 1994, so named in Baker 2008) with the same meaning or function are both available to a speaker, that is a situation of competition for language use: one of two possible forms is chosen by the speaker for use in a given utterance. As becomes important later, this is true even if the two structures do not have exactly the same meaning, function, or context of use, but just overlap in these areas; there will still be some context where the speaker is forced to choose between the two, putting them in competition for the finite resource of use, or speaker-time.
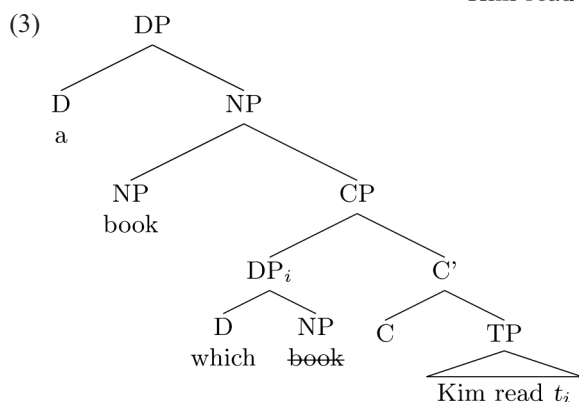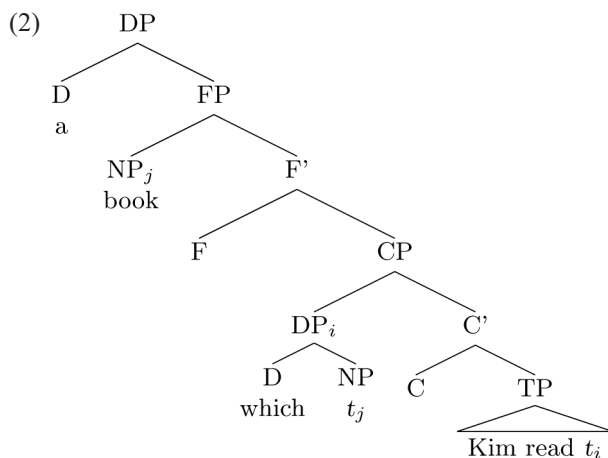
At this point, nothing more needs to be said in terms of grammatical apparatus in order to explain diachronic instability. Competition in use between functionally equivalent linguistic forms will be unstable diachronically by virtue of the fact that human brains prefer not to store or learn synonymous forms, whenever possible; indeed, this is probably a child language acquisition strategy, reducible to the independently observed 'principle of contrast' (Clark 1987, 1990, inter alia). In other words, competing syntactic forms result in an unstable situation because they are subject to the same cognitive 'blocking effect' as morphological doublets like Middle English *lough*, *laughed* ('laugh-PST'; Taylor 1994), in the way suggested by Kroch (1994). (For much more discussion of the BLOCKING EFFECT and how it can be reduced to independent principles, see Fruehwald & Wallenberg 2016.)

But this is only true under a simple, restrictive model of grammar: if the grammar is not specifically engineered to stably produce and maintain optionality as one of its design features, then it will not. The grammar simply produces derivations by putting atomic pieces together, and if the brain needs to choose between two pieces for the same function at some point in the derivation, that creates apparent optionality on the surface, and also diachronic instability. If, by contrast, a grammar is specifically designed to accommodate optional or stochastic processes as part of the derivation, such as STOCHASTIC OT (Boersma & Hayes 2001) or NOISY HARMONIC GRAMMAR (Boersma & Pater 2008, Pater 2008, and much subsequent work), then something additional needs to be said in order to explain the fact that even cases as close to diachronic stability as extraposition are in fact changes in progress where one structure replaces another.

3.2. CONSEQUENCES FOR THE ANALYSIS OF RELATIVE CLAUSE EXTRAPOSITION. This basic generalization stands, regardless of the specific analysis one proposes for relative clause extraposition, and must be accommodated in any reasonable analysis: the diachronic instability means that somewhere in the derivation, the speaker makes a (probably unconscious) choice between two syntactic formatives that are functionally equivalent in many contexts. (Note that they need not be ENTIRELY functionally equivalent for change to be imperative; I return to this point in §3.3.) While various analyses could accommodate this generalization, there is currently an analysis of English relative clauses, proposed for independent reasons, that treats relative clauses as being in a state of competing grammars: Sauerland 2003. Given the parsimony between Sauerland's independent arguments, the theory of competing grammars, and the empirical result above, I suggest the evidence lines up in favor of the Sauerland analysis of relatives. It is very tempting to explain the entire loss of relative clause extraposition as the loss of Sauerland's 'matching' type of relative and the slow ascendency of the 'raising' type of relative. However, this analysis of the change is not without problems; I point out some of these below and conclude that while Sauerland's analysis is important for understanding the change, something additional about the syntax of relative clauses and how the 'matching' type adjoins must still be said. Ultimately, we would require even more

refined data in order to determine how much of a role the Sauerland analysis plays in this change.

COMPETING STRUCTURES FOR RELATIVE CLAUSES. Sauerland (2003), building on a proposal in Carlson 1977, proposes that English relative clauses are usually ambiguous between two possible structures. These two structures, the MATCHING STRUCTURE and the RAISING STRUCTURE, are used interchangeably in most cases, though their syntax is subtly different and circumstances can arise that force a choice between them.[8] According to Sauerland, the raising structure is one along the lines of Kayne's (1994) analysis, where the 'head' of the relative clause is moved from a position internal to the relative clause, as I schematize below in 2, and then the D heading the DP containing the relative clause takes a clausal complement, rather than an NP complement (shown as 'FP' just because I am not committed to the particular version of the split-CP hypothesis one adopts to make this analysis work). The matching analysis has a similar structure internal to the relative clause, but with ellipsis of the NP rather than movement of the NP to Spec(FP), as seen in 3.

(2)

```
                    DP
                  /    \
                 D      FP
                 a     /   \
                   NP_j      F'
                   book     /   \
                           F     CP
                               /    \
                            DP_i     C'
                           /    \   /   \
                          D    NP  C     TP
                        which  t_j      /   \
                                   Kim read t_i
```

(3)

```
                    DP
                  /    \
                 D      NP
                 a     /   \
                    NP      CP
                   book    /   \
                        DP_i     C'
                       /    \   /   \
                      D    NP  C     TP
                    which  ~~book~~   /   \
                                 Kim read t_i
```

For the matching-type relative, Sauerland does assume that the relative clause modifies an NP 'relative head', which is its own constituent outside the relative clause.

---

[8] Sauerland's proposal is intended to resolve the longstanding controversy in the relative clause literature over whether relative clauses involve empty operator movement, as was standardly assumed in most treatments prior to Kayne 1994, or whether they result from the NP 'head' of the relative itself moving from the gap site inside the relative clause.

Though he does not give the details of how the relative relates syntactically to the up-stairs NP, I will assume that the relative clause is right-adjoined to the NP in the upstairs clause, as in the classical analysis of relative clauses (as I show in 3). This is in contrast to the raising-type relative in 2, in which the relative clause is itself the complement of D. I have shown this in the structures above.

Sauerland's analysis is actually a competing-grammars analysis of relatives, though he does not use that term. Combined with the hypothesis presented above, his analysis predicts that there should be some sort of change in progress in relative clauses. Note that the matching type of relative is entirely compatible with relative clause extraposi-tion, since it adjoins freely to maximal projections. As with other kinds of modification, the only restriction as to where it adjoins is that it must be interpretable in its adjoined position: it must be interpreted with respect to some antecedent 'relative head' NP and, in Sauerland's account, trigger ellipsis of the NP. This is very similar to Culicover and Rochemont's (1990) analysis of relative clause extraposition, where it does not involve movement per se, but rather choice of adjunction sites for the relative clause and a structural restriction on how far from a potential antecedent the relative can be. In fact, the locality restriction on interpretation that Culicover and Rochemont propose proba-bly makes more sense as a restriction on ellipsis. Thus, the matching relative is an ad-joined modifier with some limited freedom of adjunction.

Sauerland's raising type of relative, by contrast, is incompatible with extraposition, and so the loss of relative clause extraposition can be viewed as a side effect of the gradual out-competition of the matching structure by the raising structure. First, Baltin (2006) lists a number of problems with accounting for extraposition under the raising analysis, which would apply to raising-type relatives under Sauerland's analysis. A full discussion of these issues is beyond the scope of this article, but suffice it to say that the challenges Baltin catalogues for the raising analysis fall away if there is another possi-ble structure, the matching structure, which can accommodate extraposition as simply an additional adjunction site.

Second, there is independent evidence from binding that extraposed relatives are al-ways matching relatives. Sauerland discusses a number of binding examples that are only grammatical under one or the other structure and so force a particular structure to emerge in the mind of the speaker. Among these are condition A, which forces the rais-ing structure when an anaphor inside the relative head is bound by a DP inside the rela-tive clause; such sentences can only be grammatical if the NP containing the reflexive has raised from inside the relative, and so can reconstruct and be bound in that position. Similarly, Sauerland discusses variable binding, where a quantifier inside the relative can only bind a variable inside the relative head NP if the NP can reconstruct to a posi-tion inside the relative. Examples of these two with relative clauses in situ are shown below in 4a and 6a.

Sauerland claims that when condition A and quantifier binding are not possible in a sentence, it is because only the matching analysis is available, and so the necessary re-construction cannot take place. Note that the extraposed versions of 4a and 6a, namely 4b and 6b, are ungrammatical. (I have also included parallel examples of extraposition without the binding in 5 and 7, just to show that there is no independent problem with extraposition in these examples.)[9]

---

[9] Note that the examples below all contain demonstratives in the DP containing the relative, or which the relative is interpreted with respect to. This is because, as Sauerland discusses, the raising analysis is not pos-sible inside a DP headed by an indefinite determiner. He suggests that this may be true for all indefinites, though the data is not clear (see his footnote 9). In order to avoid this confound, I have avoided indefinites in

(4) a.   That picture of himself that John likes a lot was just published.
    b. ?*That picture of himself was just published that John likes a lot.
(5) a. That picture that John likes a lot was just published.
    b. That picture was just published that John likes a lot.
(6) a.   That draft of their paper that every student produced at the beginning of
          the semester was also graded. (*their* intended with singular reference)
    b. *?That draft of their paper was also graded that every student produced at
          the beginning of the semester.
(7) a.   That draft of the paper that the students produced at the beginning of the
          semester was also graded.
    b. That draft of the paper was also graded that the students produced at the
          beginning of the semester.

The parallel examples with potential condition C violations, however, are grammatical,
as they should be under the matching structure.

(8) That picture of John was just published that he likes a lot.
(9) That draft of Mary's paper was also graded that she produced at the begin-
      ning of the semester.

   The fact that the binding in three examples becomes ungrammatical under extraposi-
tion shows that the reconstruction required for the binding cannot take place in an ex-
traposed relative. This fact is consistent with the idea that extraposition necessarily
involves a matching structure and cannot contain a raising structure. The reconstruction
thus cannot take place because the NP in the relative head, which needs to be bound, has
never moved.

   Given that extraposition always involves the matching structure, but the in-situ rela-
tives are in a state of variation between the matching and raising structures, then the loss
of extraposition could be due to a general replacement over time of the matching struc-
ture by the raising structure. If the speaker's syntactic apparatus makes the decision about
which structure to use before or simultaneously with the decision to extrapose, then the
gradual replacement of matching by raising will result in less extraposition over time. It
would be elegant if the change could be described so simply. It would also be in line with
the view I mentioned in §3.1, that grammar competition is really competition between
multiple versions of a functional head, in this case the D selecting either an NP or a CP.
If this were always true, then it would be evidence for the Borer-Chomsky conjecture.
Unfortunately, there are also some shortcomings of the matching vs. raising approach to
the historical facts in this case.

   This analysis of the diachronic change has four potential problems. First, speakers
may be able to choose the matching structure IN ORDER TO extrapose. If this were the
case, then even if raising replaces matching in the in-situ case, the historical result
should be that raising becomes specialized for the in-situ case, while matching becomes
specialized for the extraposed case, and some other factor would have to explain the
loss of extraposition. It is hard to say if this specialization has taken place, since only
matching is compatible with extraposition, but both analyses are currently compatible
with the in-situ position. Second, Sauerland (and references he cites) observes that the

my examples, but extraposition is degraded from definite DPs (Baltin 2006 and references therein) unless
they are headed by a demonstrative. I am tempted to suggest that the fact that extraposition is degraded from
definite DPs somehow relates to the presence of the raising analysis in exactly these cases, but I do not feel
that the data on definiteness and matching vs. raising is clear enough to make such a claim at present, partic-
ularly since the binding facts show that the matching structure is available in situ in definite contexts.

raising structure is degraded in indefinite contexts, which is also where extraposition occurs most naturally. Thus, replacing matching with raising would have little effect on the rate of extraposition (unless this is actually one of the reasons the change is so slow, and/or perhaps part of the specialization of raising for in situ?). Third, since the matching structure can occur in more contexts than the raising structure can (i.e. indefinites and extraposition), there might be more evidence to the learner for matching than for raising. According to models like that of Yang (2000), this should result in matching winning the competition over time, rather than losing it. And there is a fourth problem, which I discuss below.

A WRINKLE: COMPETING ADJUNCTION SITES. If we accept Sauerland's analysis of relatives as involving two structures, a raising structure and an adjoined matching structure, then extraposition can be straightforwardly analyzed as an instance of the matching type. Since this type presumably adjoins in the way a modifier does, I adopt Culicover and Rochemont's (1990) conclusion that (matching) relatives adjoin to any maximal projection where they can receive an interpretation, and there is a general principle (their 'complement principle'[10]) governing possible antecedents for the matching relative. If we combine the approaches of Culicover and Rochemont (1990) and Sauerland (2003), then this principle becomes a constraint on the ellipsis of the 'matched' NP. According to this view, the extraposed version of the relative clause in 10b is not derived by movement, but rather represents an alternative adjunction site for the relative, higher up in the structure.

(10) a. A man that no one knew came into the room.
     b. A man came into the room that no one knew.

(Culicover & Rochemont 1990:23)

Under this analysis, matching relatives are modifier CPs that can adjoin in at least two positions in the phrase structure, without interpretive difference. This creates a situation of competing grammars, even if both adjunction sites are hosting instances of the matching structure. According to the definition of competing grammars I gave above in §3.1, in any situation where multiple structures are available to a speaker in the same context, there is competition for use between the two structures. Even though an extraposed relative is not equivalent to an 'in-situ' relative in all ways (especially prosodically, as I discuss below in §3.3), there is still overlap in use: there are still situations where a well-formed utterance can be produced with either the extraposed or the nonextraposed variant. In this area of functional overlap, a speaker is nonetheless forced to choose one or the other, and this competition in use should be diachronically unstable according to the theory I have laid out above, following Kroch (1994).

By hypothesis, this situation of competition means that even within the matching structure, there is potentially diachronically unstable variation. The two adjunction positions are in competition within the matching structure, independently of the competition between the matching and raising structures (which also exists, if Sauerland's (2003) analysis is correct). This means that the loss of extraposition could be the gradual ascendancy of the raising structure, as I suggested above, or it could be the ascendancy of the low, NP-adjunction position within the matching structure. Even if the raising vs. matching competition is indeed taking place, this additional variation would still exist and could alone be the diachronic loss of extraposition. It could even be the case that the raising vs. matching competition is resolving in some other way; for example, matching could be winning over raising, but the low adjunction site is winning over the high one within matching, leading to the loss of extraposition.

---

[10] The name is historical: their proposal does not apply only to complements.

The high-adjunction vs. low-adjunction view of the change is much more difficult to see as competition between functional heads, as Kroch (1994) suggests based on the Borer-Chomsky conjecture. But this really highlights a problem with the very idea of phrases freely adjoining to multiple sites in the phrase structure, and indeed with the idea of syntactic optionality in general: if a derivation is an algorithm, a sequence of functions to be applied to an input, then if the output of one of those functions is indeterminate, the derivation will simply never converge on an output. (It would actually be an ill-defined function, mathematically.) Optionality, including multiple adjunction sites, can only mean that there are two possible derivations, one employing, for example, function 1 (or head 1), and the second employing function 1′ (or head 1′).

So, perhaps the best way to understand the variation in relative clause structures is to view it as a three-way competition between one raising structure, which is always 'in situ', as it contains the relative head, and two matching structures, each of which specifies, perhaps on its C head, either low adjunction or high adjunction. It could also be that adjunction is the wrong description of the modification relation, and that the CP is actually hosted by a silent modifier head, but a different version of the head for high attachment to the main clause than for low. (A full discussion of the details of this possibility would be beyond the scope of this article.) In either case, it is the high-attached relative clause that is the adjunction case, losing out in its competition with two 'in-situ' structures. It will be the subject of future work to more carefully determine whether there really is a three-way competition here.

Since the competition between adjunction sites for relative clauses (if that is the right way to view the two positions), in the Culicover and Rochemont sense, is most likely a necessary part of how and why the loss of extraposition takes place, I continue the rest of the discussion in this article in terms of competition between adjunction sites. This is not to dismiss the idea that the competition between the raising and matching structures may also be implicated in the change, but merely to say that it is unlikely to be the entire story of the change for the reasons I give above, and that distinguishing the two-way vs. three-way competition scenarios will be the subject of future work.

3.3. WHY SO SLOW? At this point in the discussion, readers are likely to have in a mind an obvious criticism of the proposal in §3.1, that extraposed and nonextraposed variants are in competition for speaker use (whether we think this is due to matching vs. raising, or competition between adjunction sites within the matching structure). It is well known that the extraposed and in-situ relative clause variants are not really functionally equivalent, since extraposition is more likely to occur if the relative clause is prosodically heavier.[11] Indeed, this is a crucial point, because it is this specialization of the two variants for different prosodic weights that accounts for the exceedingly slow speed of this change: the functional difference between the two variants removes much of the competition between them, because it keeps at a minimum the contexts in which the two overlap in use. However, the specialization does not remove ALL of the overlap in use, so the competition still exists, and the change still continues.

The specialization of the two variants for different prosodic weights can never be complete, in fact, because extraposed vs. in situ is a categorical (binary) variable, while prosodic weight is a continuous one. For a full discussion of this complex and interesting issue, I refer the reader to Fruehwald & Wallenberg 2016, where this and other cases

---

[11] Of course, I am not claiming that prosodic weight is the only factor that affects the probability of extraposition, but the existence of other effects does not detract from the main point being made.

are discussed in much greater detail and a full theory of linguistic specialization is developed (see also Fruehwald & Wallenberg 2013, Wallenberg 2013).

For the present discussion, I just briefly mention one result that is also presented there. A study of individual authors in the PCEEC shows that their probability of extraposing a given relative clause from subject position depends on the prosodic weight of the relative clause (measured in number of words) and does so in a continuous manner: as the number of words increases incrementally, so does the probability of extraposition. For instance, speakers do not split the weight dimension into categories of 'heavy' and 'light', so that it matches the binary dimension of extraposition, but rather allow extraposed and nonextraposed tokens to fall along the entire scale. This is shown for the systems of individual PCEEC authors in Figure 6.[12]
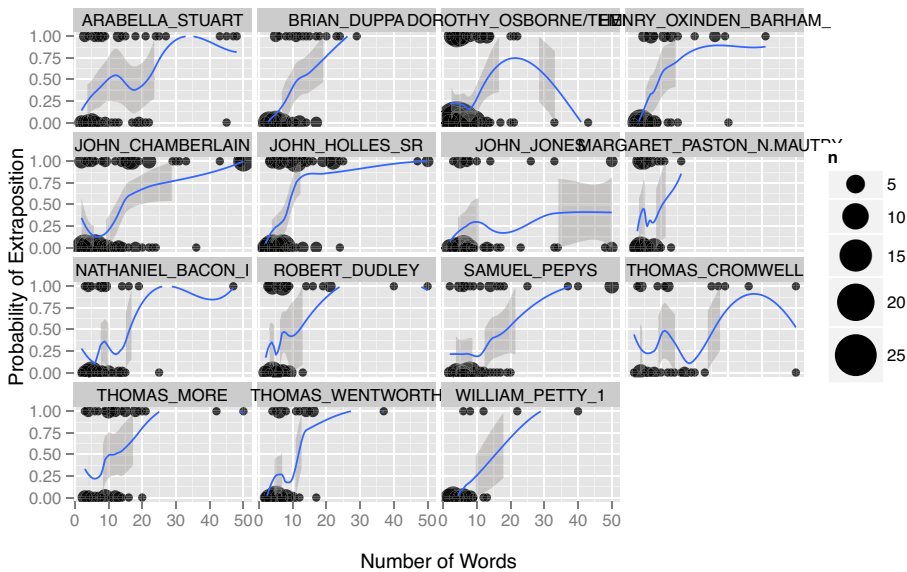

FIGURE 6. Effect of prosodic weight on the probability of relative clause extraposition from subject position, for individuals in the PCEEC. Lines are LOESS.

To summarize, because the extraposed and in-situ relative clause variants have specialized for different contexts (different prosodic weights), a large amount of the competition between them in use has been removed, and so the replacement of one by the other has slowed considerably. Because the domain of specialization is continuous, however, the two variants are unable to specialize completely, and so some of the competition remains. Thus, the extraposed variant continues to be lost over time, but very slowly.

3.4. ORIGIN OF RELATIVE CLAUSES. The next portion of this article is by far the most speculative, as suggestions regarding the actuation of a particular linguistic change often are. Nevertheless, it may be interesting to explore some implications of the very slow rate of change shown in the data above. The change is slow enough that it may well have been in progress for some thousands of years, and so quite possibly predates the differentiation of some of the languages under investigation into mutually unintelligible varieties. In particular, it may be the case that the change in extraposition relates

---

[12] Only authors with 100 tokens or more of relative clauses were included.

to the origins of Indo-European relative clauses, and that the change actually began when relative clause embedding was first innovated in Indo-European.

   If we follow the nonmovement adjunction analysis of relative clause extraposition in Culicover & Rochemont 1990, then the 'extraposed' version of a relative clause is actually the historical form of all Indo-European relative clauses, and the 'in-situ' version is a later innovation. Kiparsky (1995), following a number of previous researchers (see his references), proposes that relative clauses in Proto-Indo-European (PIE) only existed as clausal adjuncts; they were not truly embedded within the matrix clause, forming a constituent with some NP within it, but rather were adjoined to the matrix clause and associated semantically with some NP element within it.[13] I will assume that the high, 'clause-adjoined' relatives that Kiparsky describes could represent TP- or CP-adjunction, and some could be *v*P-adjoined (though this is often difficult to determine);[14] these are the possible adjunction positions for 'extraposition' I mentioned above in §3.1, with regard to Culicover and Rochemont's analysis.[15] These high-adjoined relatives represent either the only attested system or the dominant system found in various Old Indo-European languages, such as Sanskrit, Hittite, and Old Latin (Kiparsky 1995:155, and references cited therein). According to Kiparsky, the low-adjoined relative clause (again, following an adjunction analysis), the in-situ variant that appears within a DP, is a later innovation, and comes to be the dominant form in Germanic as that family differentiates from the rest of PIE (presumably this would apply to Romance as well).

   Although Kiparsky only discusses the earliest and latest stages of the change, not the details of how the change progressed in the interim, it is straightforward to apply to this case our understanding of syntactic change in progress from other changes that have been studied in quantitative detail (going back to Kroch 1989, Santorini 1989). Starting from a Sanskrit-like system of clause-adjoined (*v*P-adjoined) relative clauses, as Kiparsky suggests, once a new, DP-embedded type of relative clause is innovated, it will come into competition for use with the old clause-adjoined system. This would create the kind of variation we have observed in all four languages, if we accept that 'extraposition' is actually just the remnant of the old clause-adjoined relative system. Then, in

---

   [13] Kiparsky's analysis is broadly compatible with theories positing a historical stage prior to this one in which there were not even adjoined relative subordinate clauses without gaps, but paratactic main clauses which were later reanalyzed as subordinate clauses (e.g. Haudry 1973). However, Kiparsky's analysis certainly does not depend on there being such a stage.

   [14] Relatives appearing at the right edge of the clause are often ambiguous between these positions. Correlatives, high-adjoined on the left in languages like Sanskrit, may always be TP-adjoined, but this also depends on whether one adopts an analysis of head-final clause structures that is underlyingly head-final. Left-adjoined correlatives could be *v*P-adjoined under an analysis of head-final clause structures as underlyingly head-initial, but with 'roll-up' movement of *v*P to Spec(TP), as in Biberauer 2003 and Wallenberg 2009. In either case, though, there will still be an asymmetry between the syntax of left-adjoined and right-adjoined high relatives, but I leave a full discussion of this issue as a topic for further research. See also the following footnote.

   [15] I here note a problem for my analysis, pointed out by a referee: the high-adjoined relatives of the Old Indo-European languages are often left-adjoined correlatives, as in 11, rather than clause-final and right-adjoined, as I am arguing for the modern extraposed relatives. If the old correlative/high-adjoined system is really the same system as modern extraposed relatives, then there is a step missing from my story: when and why did the left-adjoined relatives disappear? I do not currently have a complete answer to this, and it may be a point against the analysis presented here. However, it may also be that the left-adjoined relatives have remained syntactically possible throughout the histories of the languages in question, but once high-adjoined relatives began to specialize for prosodic heaviness, the prosodic systems of early Romance and Germanic increasingly required heavy relatives to appear clause-finally. In the modern languages, the prosodic specialization ensures that left-, high-adjoined heavy correlatives should occur with a probability near zero.

order for the DP-embedded system to become dominant, as Kiparsky suggests it has in Proto-Germanic, the DP-embedded ('in-situ') variant would have to spread at the expense of the clause-adjoined ('extraposed') variant over a number of generations. This is exactly the historical progression I have shown above (with the addition that the change was slowed by specialization for prosodic weight, as I have already discussed).

The idea that the high/clause-adjoined relative system has never really left Germanic and Romance, but is really simply the 'extraposed' variants (following Culicover and Rochemont (1990)), is supported by the fact that clearly clause-adjoined relatives are still possible in languages like modern English and modern Icelandic. Kiparsky argues that in cases such as the Vedic Sanskrit sentence in 11, where the head of the relative is contained inside the relative clause and the main clause contains a different head to which the relative refers, the only possible analysis is the clause-adjoined one.

(11) yó   mártyaḥ śíśīte   áty   aktúbhir, mā́ naḥ   sá ripúr
     which mortal   sharpen.MID.SG overly nights.INST not us.GEN that trickster
         īśata
         dominate.SUBJ.3SG
     'As for the mortal who makes himself too sharp by night, may that trick-
     ster not gain power over us'   (RV 1.36.16, cited in Kiparsky 1995:156)

The same is true for the nineteenth-century English example in 12 and twentieth-century English example in 13. (Note that the same construction would be grammatical in twenty-first-century English.)

(12) By God's blessing I calculate that the Spirit of Dishonesty shall not get do-
     minion 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)

(13) 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)

Similarly, the high-adjoined system is also evidenced by cases where the relative clause refers to the entire proposition and is, in the words of Kiparsky (1995:157), 'unlinked to any argument/modifier position' in the main clause, as in the modern Icelandic and English examples in 14 and 15, respectively, or the Middle French one in 16.

(14) hann hafði enn hvorki boðið mér vott né þurrt, sem   var óvenjulegt
     he   had   still neither offered me wet nor dry   which was unusual
         á þeim bænum
         in that   town
     'he had still not offered me a drink, alcoholic or non-alcoholic, which was
     unusual in that town.'
                          (*Ofsi*, date: 2008; ID 2008.OFSI.NAR-SAG,.485 in IcePaHC)

(15) [if a man] sees a certain amount of his colleagues socially—which is ab-
     solutely essential to general harmony …
              (*The schoolmaster ...* , date: 1908; ID BENSON-1908,73.313 in PPCMBE)

(16) Ma Sibile, je vous   ai   escrit il n   i   a   que deux jours,
     my Sibile I you.DAT have written there NEG it.DAT have but two days
         qui   m anpaichera de   vous inportuner de long discours
         which me prevents   from you bother   of long discussion

'My Sibile, I wrote to you only two days ago, which prevents me from bothering you with a long discussion'   (letter by Marguerite de Valois, date: 1573–1596; ID 158X-VALOIS-AUTOGRAPH-P,124.838 in MCVF corpus)

Thus, this study has been able to build on Kiparsky's argument that the older PIE system consisted of high/clause-adjoined relatives, but I see no reason to conclude that they have disappeared from modern Germanic or Romance, nor that extraposed relative clauses are a separate phenomenon from the high/clause-adjoined system. Rather, accepting Culicover and Rochemont's (1990) analysis of extraposition as adjunction suggests a natural historical continuity: extraposition is simply the remnant of the older relative clause system (clause-adjoined), which has been in competition with the new, DP-internal relative clause system for well over a thousand years. But just as Kiparsky suggested, the quantitative study above does show that the old system is dying out, albeit very slowly.

4. CONCLUSIONS. This article has shown that relative clause extraposition is in decline in four Indo-European varieties, and that the decline represents a previously undescribed class of language change (or at least a previously undescribed speed). The change is a remarkable case in that the two variants have specialized for different functions, but this specialization has been incomplete due to the mathematical mismatch between the binary variation (relative clause position) and a continuous domain of specialization (prosodic weight). For this reason, the in-situ relative system continues to replace the extraposed variant, but the change has slowed to a speed that is barely detectable. Fortunately, with diachronic parsed corpora of sufficient time depth, it is possible to observe the change and see that it is still in progress at the present date in four languages.

The fact that the change is still in progress has implications for the analysis of relative clause extraposition. I have suggested that the best analysis of relative clause extraposition, in the light of synchronic and diachronic facts, is Sauerland's (2003) 'matching' and 'raising' analysis, combined with Culicover and Rochemont's (1990) idea of alternative adjunction sites for the matching type. It is the latter type of relative that creates the extraposition phenomenon, by sometimes adjoining high but being interpreted lower in a matrix-clause-internal position. The high(er)-adjoined clauses have been in competition with 'in-situ' relatives for a very long time. Sauerland's analysis suggests there may be a three-way competition involved as well, though more research is needed to test this hypothesis. The analysis of competing adjunction sites (or competing 'matching' structures) has the additional benefit that it follows nicely from Kiparsky's (1995) argument that all PIE relative clauses were matrix-clause-adjoined, and lower relatives are an innovation. This suggests that the decline in extraposition could have begun as a result of the innovation of lower 'in-situ' relatives in Indo-European. In that case, this is the oldest change in progress yet reported.

REFERENCES

ARONOFF, MARK. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.
BAKER, MARK. 2008. The macroparameter in a microparametric world. *The limits of syntactic variation*, ed. by Theresa Biberauer, 351–74. Amsterdam: John Benjamins.
BALTIN, MARK R. 2006. Extraposition. *The Blackwell companion to syntax*, vol. 2, ed. by Martin Everaert and Henk van Riemsdijk, 237–71. Oxford: Blackwell. DOI: 10.1002 /9780470996591.ch25.
BATES, DOUGLAS; MARTIN MAECHLER; and BEN BOLKER. 2011. lme4: Linear mixed-effects models using S4 classes. Online: https://CRAN.R-project.org/package=lme4.

Biberauer, Theresa. 2003. *Verb second (V2) in Afrikaans: A minimalist investigation of word order variation*. Cambridge: University of Cambridge dissertation.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32.45–86. DOI: 10.1162/002438901554586.

Boersma, Paul, and Joe Pater. 2008. Convergence properties of a gradual learning algorithm for harmonic grammar. Technical report. New Brunswick, NJ: Rutgers University. Online: http://roa.rutgers.edu/article/view/1000.

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

Carlson, Greg N. 1977. Amount relatives. *Language* 53.520–42. DOI: 10.2307/413175.

Clark, Eve V. 1987. The principle of contrast: A constraint on language acquisition. *Mechanisms of language acquisition: The 20th annual Carnegie Symposium on Cognition*, ed. by Brian MacWhinney, 1–34. Mahwah, NJ: Lawrence Erlbaum.

Clark, Eve V. 1990. On the pragmatics of contrast. *Journal of Child Language* 17.417–31. DOI: 10.1017/S0305000900013842.

Culicover, Peter W., and Michael S. Rochemont. 1990. Extraposition and the complement principle. *Linguistic Inquiry* 21.23–47. Online: http://www.jstor.org/stable/4178658.

Fruehwald, Josef, and Joel C. Wallenberg. 2013. Optionality is stable variation is competing grammars. Paper presented at the 25th Scandinavian Conference of Linguistics, Formal Ways of Analyzing Variation (FWAV) workshop.

Fruehwald, Josef, and Joel C. Wallenberg. 2016. Optionality is stable variation is competing grammars. Edinburgh: University of Edinburgh, and Newcastle upon Tyne: Newcastle University, ms.

Galves, Charlotte, and Pablo Faria. 2010. The Tycho Brahe Parsed Corpus of Historical Portuguese. Campinas: Department of Linguistics, University of Campinas (UNICAMP). Online: http://www.tycho.iel.unicamp.br/~tycho/.

Haudry, Jean. 1973. Parataxe, hypotaxe et corrélation dans la phrase latine. *Bulletin de la Société Linguistique de Paris* 68.147–86.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kiparsky, Paul V. 1995. Indo-European origins of Germanic syntax. *Clause structure and language change* (Oxford studies in comparative syntax), ed. by Adrian Battye and Ian G. Roberts, 140–69. Oxford: Oxford University Press.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1.199–244. DOI: 10.1017/S0954394500000168.

Kroch, Anthony S. 1994. Morphosyntactic variation. *Chicago Linguistic Society (Parasession on variation in linguistic theory)* 30.2.180–201.

Kroch, Anthony; Beatrice Santorini; and Ariel Diertani. 2005. Penn-Helsinki Parsed Corpus of Early Modern English. Philadelphia: Department of Linguistics, University of Pennsylvania. Online: http://www.ling.upenn.edu/ppche/ppche-release-2016/PPCEME-RELEASE-3. (1.8 million words.)

Kroch, Anthony; Beatrice Santorini; and Ariel Diertani. 2010. Penn Parsed Corpus of Modern British English. Philadelphia: Department of Linguistics, University of Pennsylvania. Online: http://www.ling.upenn.edu/ppche/ppche-release-2016/PPCMBE2-RELEASE-1. (~950,000 words.)

Kroch, Anthony, and Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English. 2nd edn. CD-ROM. Philadelphia: Department of Linguistics, University of Pennsylvania. Online: http://www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4. (1.3 million words.)

Martineau, France; Paul Hirschbühler; Anthony Kroch; and Yves Charles Morin. 2010. Modéliser le changement: Les voies du français. Ottawa: University of Ottawa. Online: http://www.voies.uottawa.ca. (~994,000 words.)

Pater, Joe. 2008. Gradual learning and convergence. *Linguistic Inquiry* 39.334–45. DOI: 10.1162/ling.2008.39.2.334.

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

R Core Team. 2014. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: http://www.R-project.org.

Randall, Beth. 2000/2013. CorpusSearch 2: A tool for linguistics research. Online: http://sourceforge.net/projects/corpussearch/.

Santorini, Beatrice. 1989. *The generalization of the verb-second constraint in the history of Yiddish*. Philadelphia: University of Pennsylvania dissertation.

Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory* 10.595–640. DOI: 10.1007/BF00133331.

Santorini, Beatrice. 1993. The rate of phrase structure change in the history of Yiddish. *Language Variation and Change* 5.257–83. DOI: 10.1017/S0954394500001502.

Sauerland, Uli. 2003. Unpronounced heads in relative clauses. *The interfaces: Deriving and interpreting omitted structures* (Linguistik aktuell 61), ed. by Kerstin Schwabe and Susanne Winkler, 205–26. Amsterdam: John Benjamins. DOI: 10.1075/la.61.10sau.

Taylor, Ann. 1994. Variation in past tense formation in the history of English. *University of Pennsylvania Working Papers in Linguistics* 1.143–59. Online: http://repository.upenn.edu/pwpl/vol1/iss1/10.

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. Online: http://www-users.york.ac.uk/~lang22/PCEEC-manual/.

Taylor, Ann; Anthony Warner; Susan Pintzuk; and Frank Beths. 2003. The York-Toronto-Helsinki Parsed Corpus of Old English Prose. York: University of York. Online: http://www-users.york.ac.uk/~lang22/YcoeHome1.htm.

Wallenberg, Joel C. 2009. *Antisymmetry and the conservation of c-command: Scrambling and phrase structure in synchronic and diachronic perspective*. Philadelphia: University of Pennsylvania dissertation.

Wallenberg, Joel C. 2013. A unified theory of stable variation, syntactic optionality, and syntactic change. Paper presented at the 15th meeting of the Diachronic Generative Syntax (DiGS) Conference, University of Ottawa.

Wallenberg, Joel C. 2015. Antisymmetry and heavy NP shift across Germanic. *Syntax over time: Lexical, morphological and information-structural interactions*, ed. by Theresa Biberauer and George Walkden, 336–49. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199687923.001.0001.

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

Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12.231–50.

[joel.wallenberg@ncl.ac.uk]