

# Modeling language change and language acquisition

Anthony Kroch  
University of Pennsylvania

## 1 Introduction

Consider the following two models of syntactic change, which we can call the drift model and the imperfect transmission model. Under the drift model, change is driven by a gradual shift in the frequencies with which linguistic forms are used. If, after some time, the frequency of some crucial form drops below the threshold of learnability, the grammar changes. The frequency changes are driven by factors like random variation, gradual bleaching of meaning, phonological erosion of weakly accented syllables or other properties of usage. This model is explicitly or implicitly assumed by many generative and non-generative linguists. Its advantage is that the cause of grammatical reanalysis is obvious: Once the frequency of a crucial grammatical form is low enough, language learners are no longer able to learn grammatical properties that depend on it. The problem with this approach is that the drift in frequencies that it requires has no obvious motivation. If we assume that frequencies are learned by children as part of their knowledge of the language, why should they be learned accurately in through a period of stability and then no longer so learned when a language starts to change? Similarly, if the frequencies are indirect reflections of other factors, as is more commonly believed, then why should these factors become unstable at certain times?

Under the imperfect transmission model, change begins with a change in grammar, and this grammatical change causes the frequencies with which certain linguistic forms are used to change, as the grammatical change propagates through the speech community. The change in grammar is caused by an inaccuracy in language transmission; that is, by mislearning. The mislearning may be by adult second-language learners, as in the case of certain kinds of language contact, or by children. The character of changes will be determined by the sorts of errors that learners are likely to make. This way of thinking about language change is also quite widespread, and it is not always remarked that it is a quite different model from the first one. Either it can be invoked only in cases of contact and second language acquisition, where we know that learning is inaccurate or it will require that we loosen the standard idealization of synchronic generative linguistics that language learning is accurate; that is, that given a set of primary data, all learners will converge on the same grammar. Instead, we will have to say that success in learning is probabilistic. Given a population of learners, there

will be errors with a probability distribution that depends on the strength of the evidence for given grammatical properties and on variation in the primary data that learners are exposed to. The advantage of appealing to imperfect learning as a source of change is that it eliminates the need for uncaused drifts in usage frequencies that occur prior to and independently of grammar change. The chief problems of such an approach is that we do not know whether learning errors are sufficiently common to serve as an important cause of change, nor do we have a model of the mix of the propagation and suppression of error that would be needed to account for the mix of stability and change that languages present when examined diachronically.

In what follows, I will argue for a version of the imperfect transmission model, making my argument in two parts. First, using grammatical and quantitative analyses of the historical record, I will show that, in cases of syntactic change, diagnostic evidence for an innovative grammar appears in the record early in the evolution the frequency distribution of surface forms. Of course, the historical record is never as rich as we would like, which limits certainty of our conclusions; nevertheless, it seems that diagnostic evidence for grammatical innovation appears as early in the time course of a change as one can expect, given the size and quality of the record in each case. Hence, we conclude that grammatical change precedes frequency change. Clearly, if this conclusion is correct, the drift model is undermined. Second, I will address the problem of how grammatical innovation can appear “out of nowhere” in the record by giving evidence that the general stability of languages over time, far from being a simple fact, masks a complex homeostatic system of variation and regulation which often slips, allowing grammar change to enter into a community of speakers, where, under favorable circumstances, it may spread and become a recognized language change. The character of the homeostatic system arises out of the generally imperfect character of language learning, even by children acquiring their first language. This imperfection leads to the need for a mechanism of error correction, a mechanism which itself is only approximately effective.

In specifying a model of language learning, I will rely on proposals by Clahsen and others (Clahsen 1991, Penner 1992) to the effect that learners set linguistic parameters in an irreversible way; that is, once the learner has set a parameter, s/he cannot unset it. I will assume further that this irreversibility of parameter setting is a property of the primary language acquisition device only, not of the whole linguistic system of the learner. If a learner sets a parameter erroneously, because certain data has been ignored or not yet encountered, s/he can still learn the language “correctly” but only in the way that s/he would learn another language or dialect, by adding knowledge rather than by substituting one setting for another. The result is bidialectalism with respect to the relevant parameter. If the learner is young enough and the second setting is conclusively better supported by the data than the first, the child will slowly shift toward using the second “dialect” at the expense of the first. Over time, this shift will become total, at least when evaluated on the basis of linguistic performance. Whether the mental state of such a learner is identical to that of someone who does not go through the process of correction is a more difficult question, though work on bilingualism by psychologists of language suggests that it may not be. If both

parameter settings are supported by the data, as when they correspond to different registers of the language – for example, one the home vernacular and the other the standard language used in school – then the learner will preserve active knowledge of both dialects. Bidialectalism due to mislearning and correction is the source, in the imperfect transmission model as developed here, of language change. For at least one reason the process of correction of misset parameters cannot be infallible. If one child hears language produced by another that has misset a given parameter, then the listening child will actually get positive evidence for the incorrect setting and may be induced to itself set the parameter incorrectly. Any bias in the system that favors the incorrect setting may, under appropriate circumstances, favor the eventual spread of the originally incorrect setting. For example, in the spirit of much recent work, the weakening of morphological evidence for V-to-I raising may leave the word order in sentences with adverbs as the only evidence for verb movement. If, in the spirit of Clark and Roberts (1993) and Bobaljik (2002), the raising analysis is seen as syntactically more complex than the verb *in situ* analysis, we might expect verb raising to disappear, once the morphology simplifies, because mislearning will be common and the correction process will not be effective enough to maintain language stability.

Once language change is in progress, the learner will hear evidence in the speech community for mutually inconsistent settings and learn both. Over historical time, one setting is likely to win out over the other, for reasons that have yet to be well understood. There may well be a weak but systematic cognitive bias against bilingualism that, in the absence of countervailing forces, pushes speech communities toward grammatical uniformity. In any case the time course of change is easily enough observed. In all members of the community, one setting gradually wins out over the other. Initially, forms that give evidence for both settings appear with more than a negligible frequency and everyone learns and uses both settings. In the course of time, however, the relative frequencies of the two forms shift until eventually the one form occurs too rarely to support learning. In the historical data that I will discuss, because it comes from written records, we will be looking at competing settings that probably reflect different stylistic registers, since written language is generally conservative and resists change. In such cases, the vernacular seems always to win out over a long enough span of time, probably because it is the dialect of unreflecting speech.

## **2 Evidence for grammar change before frequency change.**

In arguing for the imperfect transmission model of change, I want to begin with historical data and show that close inspection of the time course of change gives evidence that the frequency changes found in languages undergoing syntactic change reflect grammar competition rather than drifts in usage. This approach is explicitly contrary to the view of Hróarsdóttir in her publications on Icelandic (Hróarsdóttir 2002, 2003, 2004, and elsewhere), as well as Lightfoot in various discussions (see, for

example, Lightfoot 1993, 1999). Unfortunately, I am at present not in a position to address Hróarsdóttir’s treatment of Icelandic directly, though it certainly deserves careful consideration. Instead I propose to look at certain other cases that have been thought to involve drift and give evidence against that approach. Specifically, I propose to discuss three concrete cases that have been investigated in detail in which I think we can distinguish the two models on the basis of the quantitative predictions that they make and show that the evidence presently available favors grammar change before frequency change. These cases are: (1) the shift from OV to VO word order as discussed by Pintzuk; (2) Santorini’s discussion of the shift from Infl-final to Infl-medial phrase structure in Yiddish ; and (3) the loss of Verb-to-Infl movement in English as studied by myself and Chung-hye Han.

## 2.1 OV $\Rightarrow$ VO in English (Pintzuk 1997, 2002).

It is often said (Stockwell 1977, Canale 1978, Allen 2000, among many others) that the shift from OV to VO word order in English occurred as a result of a drift in the frequency of extraposition, which eventually led learners to reanalyze the grammar as VO. This approach has been criticized by Pintzuk, who makes the point that it predicts that clear cases of VO order, ones which cannot be analyzed resulting from extraposition should on the account only be found at the end of the change. In the English case, however, there are examples of head-initial VPs found at least by the 10th century, a period when OV was still by far the predominant order. Pintzuk cites the following example, in which a verbal particle appears after the main verb.

- (1) he wold adraefan ut anne æþeling  
 he would drive out a prince  
 ChronB(T) 82.18-19 (755)

Crucially, in this example there is a tensed auxiliary in second position so that the postverbal position of the particle cannot be attributed to movement of the verb to T or C.

As is known from the syntax of the other West Germanic languages and is confirmed in the first line of Table 1, particles do not postpose in OV sentences. Hence, Pintzuk points out, the existence of a non-negligible number of postverbal particles in the Old English corpus, as shown in line two of the table, is entirely unexpected on the story that increasingly frequent extraposition leads to VO order. Instead, Pintzuk says, the word order in (1) is better understood as an early instance of VO order. In support of this conclusion, she adds that 14 of the 15 examples of postverbal particles in her database have the main verb immediately after the auxiliary. One has an adverb between the two verbs. None has an argument or prepositional adjunct in that position. This is expected if the sentences are VO but not if they are OV plus some other process that shifts the particle to the right of the verb.

Clause type	Frequency after main verb
INFL-final	0/90 = 0.0%
INFL-initial	15/275 = 5.5%

Table 1: Frequency of post-verbal particles by clause type in Old English (from Pintzuk 1997).

Another striking piece of evidence that Pintzuk gives in support of the occurrence of VO word order in Old English is a comparison of word order frequencies in *Beowulf* with those in later texts. In earlier work (Pintzuk and Kroch 1989), we showed that *Beowulf* is an OV language with verb-second. There is no evidence of VO word order in the poem. In particular, there are no examples like (1). Because the poem is very early, this is not surprising and it allows Pintzuk to compare frequency patterns in *Beowulf* with those in later texts. What she finds is shown for DP complements in tables 2 and 3. PP complements and adjuncts behave similarly.

Clause structure and type	N	Frequency postverbal
I-initial main clause	68	.22
I-initial subordinate clause	32	.09
I-final main clause	38	.40
I-final subordinate clause	72	.25
Total	210	.24

Table 2: Frequency of postverbal DPs in *Beowulf* by clause type.

Clause structure and type	N	Frequency postverbal
I-initial main clause	245	.43
I-initial subordinate clause	230	.23
I-final main clause	40	.08
I-final subordinate clause	140	.12
Total	655	.27

Table 3: Frequency of postverbal DPs in later Old English by clause type.

There is a striking difference between *Beowulf* and the later texts in the effect of INFL-final word order on the likelihood of a complement appearing in postverbal position, where once again only sentences with auxiliary verbs are considered to avoid the the verb-second effect. In *Beowulf*, DP complements appear postposed after the verb slightly more often in INFL-final sentences than in INFL-medial ones, though this effect is not statistically significant. In the later texts, however, there is a strong effect in the reverse direction. Pintzuk points out that this pattern is what would be expected if VO surface order in later Old English, but not in *Beowulf*, was a

mixture of cases of DP extraposition and cases of underlying VO order. This pattern adds further strength to the case for the early appearance of VO order in English and against the drift plus reanalysis story. The database on which Pintzuk bases her conclusions is limited in size and we should learn more about the prevalence of VO order in Old English when the larger parsed corpus recently completed at the University of York is analyzed.

## 2.2 The rise of INFL-medial phrase structure in Yiddish (Santorini 1993).

Some time before Pintzuk's studies of Old English, Beatrice Santorini explored the history of Yiddish using similar methods. Santorini's work was one of the earliest studies to show that it was possible to use structurally ambiguous data in studies of the time course of language change when one could estimate the relative frequencies of the different structures in environments where the structures could be distinguished on the surface. In an important paper Santorini showed that the rate of increase in INFL-medial word order when estimated from unambiguous sentences like those in (2) – (6).

Unambiguously INFL-final sentences:

- (2) ven der vatr nurt doyts leyan kan  
if the father only German read can
- (3) ven du mir meyn kop ab shneydst  
if you me my head off cut
- (4) dz ikh reyn t<sub>i</sub> verde [PP fun der ashin]<sub>i</sub>  
that I clean become from the ash

Unambiguously INFL-medial sentences:

- (5) ven mn hibbt shme isral an  
when one lifts Shma Israel on
- (6) az di nshmh zal nit oys gin  
that the soul shall not out go

Santorini classified these sentences into two types, those with a single tensed verb and those with both a tensed auxiliary and a non-finite main verb, which she called the simple verb and complex verb cases respectively. The simple verb case consisted of those sentences in which the verb was final on the surface and there were enough preverbal constituents to ensure that the sentence was not structurally verb second (as illustrated in (3) and (4)) plus those cases that were surface verb medial and contained an element, like a particle, to the right of the verb that is known not to extraposed (as illustrated in (5)). Sentences with auxiliary verbs were always taken to be unambiguous because the relative position of the two verbs is known independently of the position of other elements. The possibility of West Germanic verb raising effects

in these cases was put aside for reasons that would take us too far afield to discuss here. The data from Santorini’s corpus is given in Table 4.

Date	Simple verb case			Complex verb case		
	INFL-medial	INFL-final	frequency	INFL-medial	INFL-final	frequency
1400-1489	0	27	.00	1	15	.06
1490-1539	5	37	.12	2	35	.05
1540-1589	13	59	.18	4	51	.07
1590-1639	5	81	.06	7	57	.11
1640-1689	13	33	.28	18	44	.29
1690-1739	15	20	.43	25	28	.47
1740-1789	1	1	.50	11	17	.39
1790-1839	54	3	.95	79	0	1.00
1840-1950	90	0	1.00	62	0	1.00

Table 4: Frequency of Yiddish INFL-medial phrase structure by verb type in unambiguous clauses.

When the rates of change for the two verb types in this data are estimated via logistic regression the results are as in Table 5.

	Slope	Intercept
Unambiguous simple verb case	1.01	-18.0
Unambiguous complex verb case	1.19	-20.9

Table 5: Rates for the rise of Infl-medial phrase structure in unambiguous contexts (logistic regression estimates).

The slopes and intercepts here are quite close and statistically not likely to be distinct. This is as expected, given the Constant Rate Effect that we have discussed in other work (Kroch 1989). The more interesting result that Santorini found involved ambiguous cases like those in (7) – (8).

(7) daz er hat eyn brudr  
that he has a brother  
(INFL-medial or INFL-final plus postposed object)

(8) da ihushe kam in arts isral  
when Joshua came into land Israel  
(INFL-medial or INFL-final plus postposed PP)

Because Yiddish allowed rightward extraposition of DP and PP complements and adjuncts, most sentences without an auxiliary are in fact ambiguous between an

INFL-medial and INFL-final analysis. The number of simple verb cases which contain diagnostics for INFL-medial or INFL-final structure is only about 40% of the total. Table 6 gives the data and Table 7 the regression results for this case.

Date	Ambiguous case		
	Possibly INFL-medial	INFL-final	frequency
1400-1489	3	9	.25
1490-1539	13	13	.50
1540-1589	58	39	.60
1590-1639	41	29	.59
1640-1689	32	21	.60
1690-1739	21	11	.66
1740-1789	5	2	.71
1790-1839	58	3	.95
1840-1950	69	0	1.00

Table 6: Frequency of Yiddish INFL-medial phrase structure by verb type in ambiguous clauses.

	Slope	Intercept
Unambiguous cases combined	1.11	-19.5
Ambiguous (simple verb) case	.36	-5.5

Table 7: Rate for the rise of Infl-medial phrase structure in the ambiguous context (logistic regression estimate).

It is clear from the regression results that the ambiguous case does not behave like the others. Santorini claimed that this was because the ambiguous data was a mixture of INFL-medial and INFL-final clauses, as the grammar allows. She reasoned further that if it was possible to estimate the proportions of this mixture, then the data could be apportioned between the two types and the rate of increase in INFL-medial structure could be recalculated. To estimate the proportions of the mixture, she used data on the extraposition of DPs and PPs in unambiguously INFL-final clauses, assuming that the rate of extraposition would be independent of the clause type in which it occurred. This independence assumption has been supported by other studies (Kroch 1989, Taylor 1994, among others). In any case, the result of these calculations brought the behavior of the ambiguous case in line with the other, as can be seen in Table 8.

	Slope	Intercept
Unambiguous cases combined	1.11	-19.5
Ambiguous case corrected	.97	-16.9

Table 8: Rate for the rise of Infl-medial phrase structure in the ambiguous context, corrected for postposition.

The data on extraposition that Santorini used to rectify her ambiguous cases is given in Table 9.

Date	DP postposing			PP postposing		
	Postposed	Not postposed	frequency	Postposed	Not postposed	frequency
1400-1489	1	12	<b>.08</b>	9	12	<b>.43</b>
1490-1539	7	19	<b>.27</b>	13	16	<b>.45</b>
1540-1589	7	24	<b>.23</b>	52	21	<b>.71</b>
1590-1639	10	40	<b>.20</b>	39	23	<b>.63</b>
1640-1689	4	19	<b>.17</b>	17	30	<b>.36</b>
1690-1739	1	5	<b>.17</b>	6	3	<b>.67</b>
1740-1789	1	2	<b>.33</b>	8	7	<b>.53</b>
1790-1839	0	1	<b>.00</b>	1	1	<b>.50</b>
1840-1950	no INFL-final data		—	no INFL-final data		—

Table 9: Frequencies of DP and PP postposition by time period.

This data is of the greatest interest for our present purposes. The same intuition that has led students of the history of English to claim that the transition from OV to VO word order was a consequence of a drift in the frequency of extraposition could be applied to Yiddish. Why not say that Yiddish underwent a similar drift, which led to a similar result in that language? Santorini’s data, however, shows that this story is unlikely in the Yiddish case. Inspection of Table 9 shows that there is no measurable drift in the frequency of extraposition for either DPs or PPs throughout the history of the language, right up to the period when INFL-final word order is lost and the measurement of rates of extraposition becomes impossible. We have no data comparable to this Yiddish data for Old English and difficulties in dating the texts may make it hard to replicate the Yiddish study, but certainly the Yiddish data is entirely consistent with the Old English situation as Pintzuk has described it. In both cases, the evidence cuts against the drift model and provides indirect support for the alternative model of early grammatical change that I am proposing.

### 2.3 The loss of verb-to-INFL raising in English.

One of the most interesting and well-known syntactic developments in the history of English is the loss of so-called “verb to INFL raising.” In Middle English, the tensed

verb invariably appears in a functional head position called “INFL” that precedes negation and certain adverbs. In modern English, by contrast, the tensed verb appears in this position only if it is an auxiliary. When a sentence of modern English contains no auxiliary verb, the adverb or negation appears before the verb. In the case of negation, however, the pleonastic auxiliary *do* must also be inserted into INFL. The contrast between the Middle English and modern English forms is illustrated in examples (9) and (10):

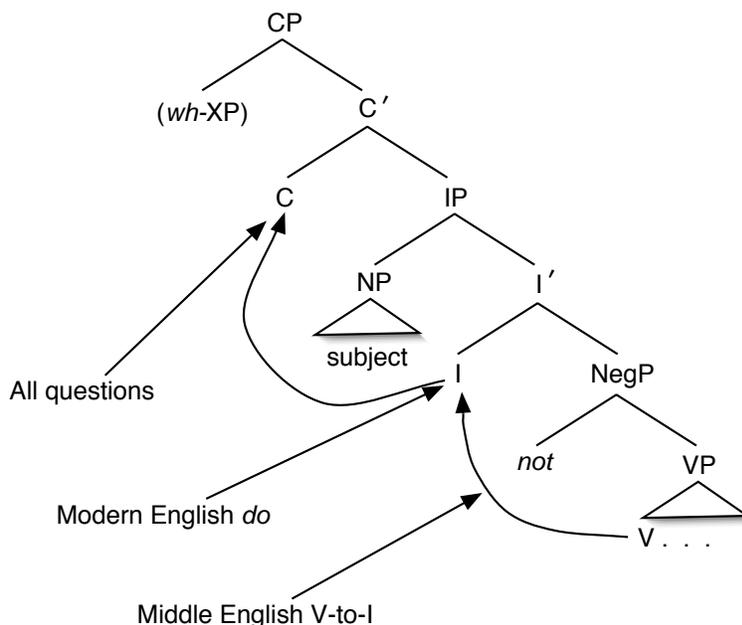
- (9) a. but there faces were backward So that they **sawe not** there fathers  
nakydnes. (TYNDOLD,IX,20G.322)
- b. And when mens judgments are once byassed they **do not believe** accord-  
ing to the evidence of things ... (TILLOTS,2.2,419.42)
- (10) a. One is, in the beste pasture where the rammes **goo alwaye** with theyr  
ewes ... (FITZH,44.218)
- b. He **never seemed** to understand foreign affairs well ...  
(BURNETCHA,1.1,169.51)

Pleonastic *do* is also obligatory in questions in Modern English, while in Middle English it was not used in this environment either. The contrast in this case is illustrated in (11):

- (11) a. Wherefore **toldest thou** me not that she was thy wife? (TYNDOLD,XII,1G.365)
- b. How longe **dost thou make** vs doute? (TYNDNEW,X,20.907)

In the Germanic languages, the tensed verb in questions moves to the front of the sentence, to a head position that is standardly thought to be the same as that of the complementizer (“COMP”) that introduces subordinate clauses. It is also thought that this movement occurs in two steps, first a movement to the INFL position and then a further movement to the COMP position. If we characterize the change from Middle to Modern English as loss of the movement to INFL, then we can say that *do* is inserted as a substitute for that movement in both negatives and questions. The alternative derivations are illustrated in the tree diagram in (12):

(12)



Among the many issues raised by this approach is, of course, why the presence of negation requires V-to-I movement or the insertion of *do* while the presence of an adverb does not. Several interesting answers have been given to this question and others but discussion here of the matter would take us too far from our diachronic interests so we will put it aside.

David Lightfoot (1993, 1999) and others have argued that V-to-I movement was lost from English only in the 18th century, while we have claimed that the loss occurred in the middle of the 16th century. This difference of opinion is not a difference over the facts. It is well known that the characteristic V-to-I word order continued to be possible in negative sentences well into the 18th century while with questions this order had largely disappeared by the end of the seventeenth century. We have claimed, however, that there was a grammatical reanalysis in the middle of the 16th century, so that the grammatical character of the V-to-I raising option was different after that point than it had been before. According to Lightfoot, on the other hand, the syntactic changes involving English V-to-I movement occur at the end of Middle English, with the introduction of periphrastic *do*, and in the 18th century, with the final loss of the old word order. Between these two points in time, the frequency of periphrastic *do* drifts upward and this drift, along with other changes in the language eventually leads to the grammatical reanalysis that makes the old word order no longer possible.

The difference between Lightfoot's account and ours is another instance of the general opposition between descriptions based on drift followed by reanalysis and those based on grammar change followed by competition between grammars. In this case, we are fortunate to have a quantitative study of great scope and insight that may help us to evaluate the competing stories. In 1953, the Swedish linguist Alvar Ellegård pub-

lished a doctoral dissertation on the rise of periphrastic *do* which has ever since served as a model for quantitative diachronic syntax. The data in his study is extremely extensive, enough so as to support reanalysis by modern statistical techniques and our work, pending the completion of the early Modern English corpus we are currently constructing, and of other early Modern English corpora under construction elsewhere, has been based on Ellegård's data. Let us look at it in some detail to see what has led us to our conclusions.

Table 10 and the corresponding graph in Figure 1 gives us a summary of the quantitative evolution of periphrastic *do* in negative declaratives and in questions.

Date	Negative declarative		Negative question		Affirmative question	
	% do	N	% do	N	% do	N
1400-1425	0	177	11.7	17	0	1
1426-1475	1.2	903	8.0	25	2.6	38
1476-1500	4.8	693	11.1	27	12.5	40
1501-1525	7.8	605	59.0	78	25.5	55
1526-1535	13.7	651	60.7	56	46.2	26
1536-1550	27.9	735	75.0	84	47.6	84
1551-1575	38.0	313	85.4	48	73.1	67
1576-1600	23.8	629	64.8	128	76.7	202
1601-1625	36.7	278	93.7	95	86.4	381
1626-1650	31.7	344	84.2	38	88.8	89
1651-1700	46.0	274	92.3	52	89.3	122
1700-1750	87.1	70	100.0	16	100.0	37

Table 10: Frequency of periphrastic *do* by context (all data from Ellegård 1953).

This data gives little indication of any reanalysis in the mid-16th century and appears to support the view that grammatical changes occurred only at the beginning of Modern English, with the appearance of *do* and later with the use of *do* becoming obligatory. There is, however, one odd feature of this graph. While the use of *do* in questions rises along the expected S-shaped curve that is characteristic of gradual linguistic change, the S-curve in negative declaratives appears to be interrupted in the middle of the 16th century and not to rise again for another 100 years. It should be mentioned in this context that last time period in Ellegård's table is based on a small number of tokens and that other studies have shown its estimate for *do* in negative declarative to be too high. Logistic regression applied to this data shows that the rate of increase in the use of *do* is the same in negative declaratives and in questions until the mid-16th century but that afterward the rates diverge. This divergence is contrary to the Constant Rate Hypothesis, which we have proposed and which is supported by several other studies, according to which different linguistic contexts in which a single grammatical change is manifested should change at the

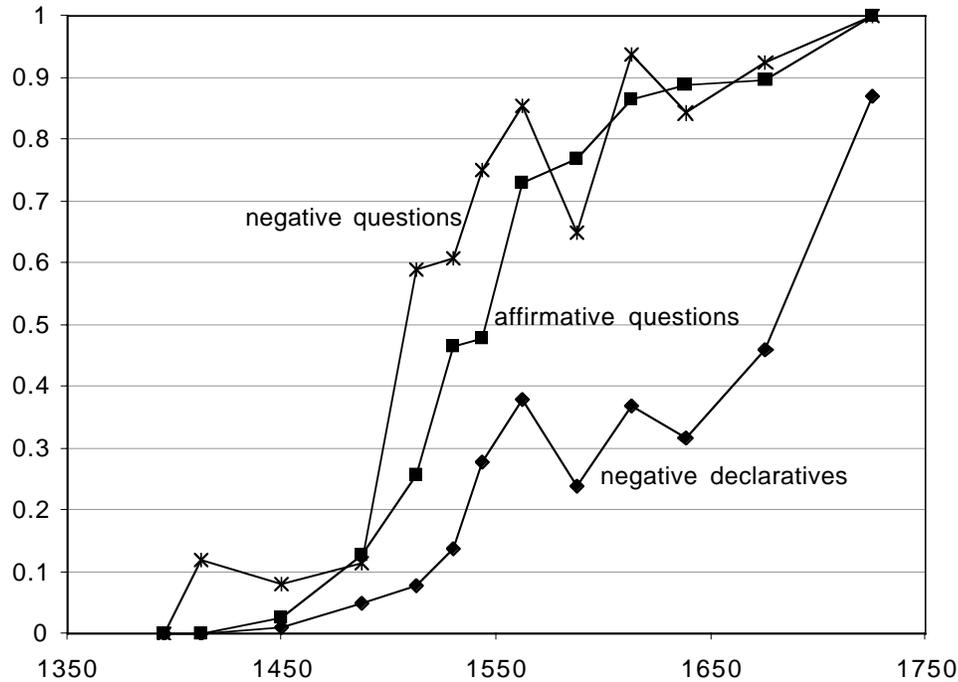


Figure 1: Frequency of periphrastic *do* by context.

same rate. This violation of the Constant Rate Effect leads us to wonder whether the grammars in play after the mid-16th century are the same as those in play earlier.

Consider now Tables 11 and 12 and the augmented graph in Figure 2. Here we see two other contexts that are relevant to the loss of V-to-I movement, affirmative declarative sentences and sentences with medial adverbs. Ellegård gives us data for the common adverb *never*, which is meant to stand for all similar elements. In early Modern English, periphrastic *do* was possible in ordinary affirmative declaratives like (13):

- (13) for whip and whurre. The olde prouerbe **doth say**, neuer made good furre.  
(UDALL,L.263.148)

In the modern language *do* is only possible in ordinary affirmative declaratives in cases of *verum* focus. The fact, therefore, that the rate of use of *do* in this context rises along with other contexts until the middle of the 16th century and declines thereafter is a further indication of grammatical change. Postverbal appearance of an adverb, as in (14), is a clear indication of V-to-I movement and, again, the fact that it occurs in only 10% of cases after the middle of the 16th century is also suggestive.

- (14) þou **wost** **neuer** what, ... (CMCLOUD,17.75)  
thou knowest never what ...

Dates	% do	N
1400-1425	0.23	4,324
1426-1475	0.27	42,770
1476-1500	1.78	56,024
1501-1525	1.37	26,884
1526-1535	2.27	17,672
1536-1550	7.05	18,048
1551-1575	8.13	13,724
1576-1600	4.59	16,920
1601-1625	2.07	7,426
1626-1650	1.43	6,768
1651-1700	0.92	7,426

Table 11: Frequency of *do* in affirmative declaratives.

Dates	% never-V	N
1426-1475	23.5	154
1476-1500	34.7	186
1501-1525	69.4	109
1526-1535	88.8	170
1536-1550	89.8	152
1551-1575	89.2	88

Table 12: Frequency of *never-V* word order in sentences with tensed main verbs.

The most interesting change in the quantitative patterns associated with the loss of V-to-I movement that occurs in the mid-16th century is what happens with negative imperatives. In Old and Middle English, as in modern German, imperatives move to the COMP position at the front of the clause. This fact is, of course, visible only in examples with overt, usually 2nd person, subjects, as in the examples in (15) and (15b):

- (15) a. Helpe thou me.  
help you me  
(The Earliest Prose Psalter 150.2290)
- b. Naske ye of cunseil.  
not-ask you of counsel  
(Ancrene Riwe 58.569)

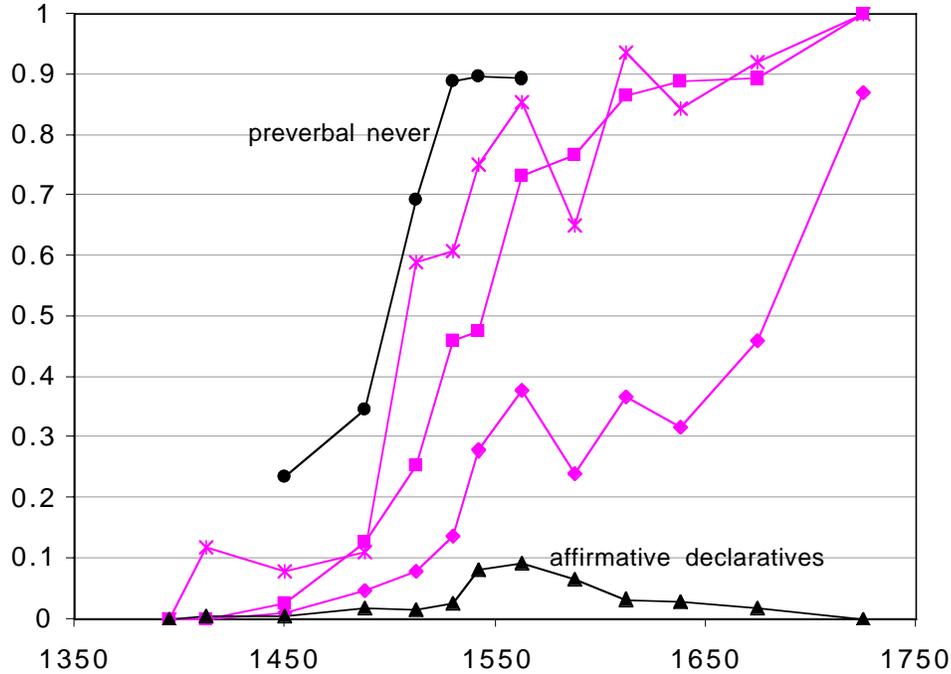


Figure 2: Frequency of preverbal *never* and affirmative declarative *do* compared to *do* in questions and negative declaratives.

This pattern was the same in positive and negative imperatives. In early Modern English the parallelism between positive and negative imperatives breaks down. Main verbs stop moving to COMP. In positive imperatives the loss of verb movement means that these sentences no longer manifest subject-verb inversion. Hence, we have modern English examples like (16):

- (16) You come here, child.

The development in negative imperatives is different. In this case, there is still verb movement to COMP, but it is movement of periphrastic *do*, as is seen in (17):

- (17) a. Don't you worry.  
 b. Don't anybody move.

For present purposes, the crucial fact about this split between positive and negative imperatives is that it occurred in the middle of the 16th century. This can be seen in Table 13 and Figure 3, which give the frequency of *do* in negative imperatives.

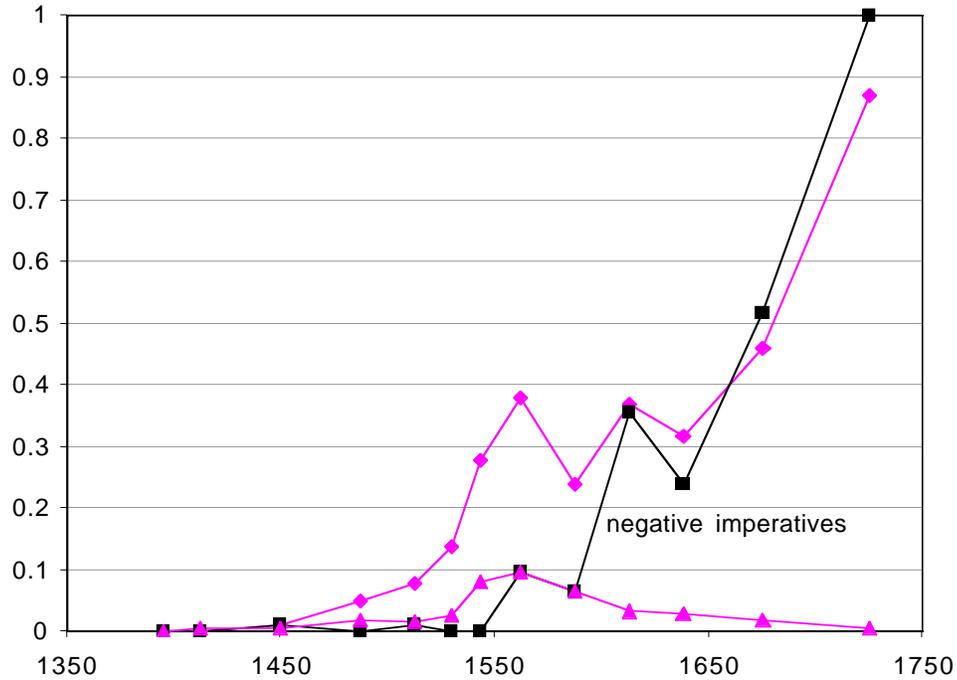


Figure 3: Frequency of *do* in negative imperatives compared to negative and affirmative declaratives.

Dates	% do	N
1400-1425	0	52
1426-1475	1.06	282
1476-1500	0	129
1501-1525	1.20	166
1526-1535	0	101
1536-1550	0	72
1551-1575	9.30	43
1576-1600	6.40	125
1601-1625	35.33	184
1626-1650	23.81	21
1651-1700	51.52	33
1701-1710	100.00	28

Table 13: Frequency of *do* in negative imperatives.

Figure 3 shows clearly that negative imperatives behave like ordinary affirmative declaratives until the middle of the 16th century and thereafter change their behavior abruptly and start behaving like negative declaratives. Once again the data suggest a grammatical reanalysis, but in this case our knowledge of the syntax of negation and of imperatives and of how they might interact can help us to work out fairly precisely

what the reanalysis must have been, which Chunghye Han and I did in a recent paper (Han and Kroch 2000).

The analysis proposed by Han and Kroch depends on two notational assumptions and one fact. The first assumption is that the INFL projection should be decomposed into three heads, as originally proposed by Jean-Yves Pollock and since widely adopted in transformation variants of generative grammar. In our version of this so-called “split INFL” hypothesis, the three heads are Tense, Mood, and Aspect, which appear in the syntactic tree for a tensed sentence in that order. The second assumption is that infinitives and imperative sentences lack a Tense projection, as has been argued by several researchers. Under this assumption the *to* that marks the English infinitive cannot occur in Tense and we follow a suggestion by Baltin that it occurs in Mood. This proposal makes sense of the fact that the English *to*-infinitive substitutes historically for the subjunctive mood. The fact we rely on is that English has two positions for negation, one above Mood and one below. This fact is easily seen in the example in (18), where the two negatives can actually cooccur:

- (18) Bill believes [<sub>IP</sub> Mary (not) *to* (always) (not) agree ].

In Late Middle English, these two positions for negation were both available, as in the modern language, but the lower negative did not occur between *to* and the verb. Instead the verb moved past the negation into the Aspect node. The possibility of the higher negation is shown in (19) and of the lower negation in (20):

- (19) *not-to-verb*

- a. that sche wuld vwche-save nowth to labowre agens yw in this  
 that she would promise not to labour against you in this  
 matere tyl ye kom hom  
 matter until you come home  
 (Paston Letters 221.310)
- b. that they that ben sike of hir body ben worthy to ben hated but  
 that they that are sick of their body are worthy to be hated but  
 rather worthy of pite wel more worthy nat to ben hated  
 rather worthy of pity even more worthy not to be hated  
 (Chaucer’s Boethius 449.C2.379)

- (20) *to-verb-not*

- a. to sorow noght for hys syn as he sulde do  
 to sorrow not for his sin as he should do  
 (Rolle’s Form of Living 99.260)
- b. And herfore monye men vson wel to come not in bedde with  
 and therefore many men are-used well to come not in bed with  
 schetis, but be hulude aboue the bed  
 sheets but be covered above the bed  
 (Wycliffite Sermons I,479.641)

The order *to* > *not* > verb was not possible in Late Middle English and arises in modern English rather late in the loss of verb movement. The examples in (21) and (22) show that low negation is not adjoined to right of the verb phrase since it appears on the surface between the verb and its complement. Under the assumptions of a transformational analysis, examples like these are sufficient to demonstrate the existence of verb movement in infinitives.

(21) *to-verb-not-participle*

- a. and said mayster parson, I praye you to be not displeasyd ...  
 and said master parson I pray you to be not displeased ...

(Caxton's Prologues and Epilogues 88.176)

- b. Ha! What it es mykell to be worthi lovyng and be noght loved!  
 ha what it is much to be worth loving and be not loved

(Rolle's Form of Living 88.52)

(22) *to-verb-not-direct object*

- a. to conforme noght his will to Gods will, to gyf noght entent till hes  
 to conform not his will to God's will, to give not heed to his  
 prayers  
 prayers

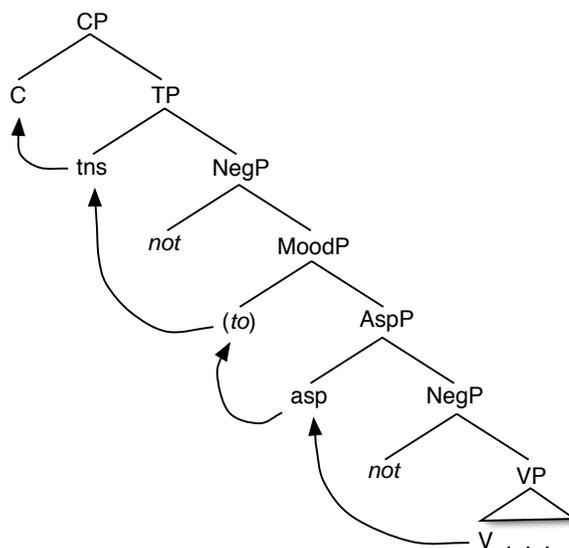
(Rolle's Form of Living 99.263)

- b. and to spille not oure tyme, be it short be it long at Goddis  
 and to waste not our time, be it short be it long at God's  
 ordynaunce.  
 ordinance

(Purvey's Prologue to the Bible I,56.73)

The model of verb movement we are assuming for Middle English is illustrated in the tree in (23):

(23)



Let us now try to understand the quantitative patterns in the negative declarative and negative imperative data. This data can be shown to be well behaved once we address the question of what it means to “lose V-to-I movement” in a grammatical theory that encompasses an split INFL. Consider the fact that modern English does not allow negation to appear between the verb and its complements in infinitives, just as it doesn’t allow this in tensed sentences. Given the structure in (23), this means that the modern language has lost all movement of the verb out of the VP. How did it do so? One possibility is that all movement was lost together, but this would be rather surprising if the splitting of INFL is really justified. Another possibility is that the loss is sequential, with movement to higher positions lost before movement to lower positions. Adopting this second possibility as a hypothesis provides a way of understanding the quantitative data. If V-to-Tense movement is lost first, this will have the largest impact on questions, since in that construction the verb must move through tense to COMP. In negative imperatives, the loss will have no effect since the tense projection is absent, so that movement to COMP need not pass through Tense. In negative declaratives, the situation is intermediate. When negation is in the high position, it intervenes between Tense and Mood, so that loss of verb movement from Mood to Tense will trigger the use of *do*. On the other hand, when negation is low, the verb will move across it even when movement from Mood to Tense is lost, so that the use of *do* will not be triggered. How often *do* occurs will depend on how often the negation is generated in the high versus the low position. What is more, if we had a way to calculate the frequency with which negation occurs in the two positions, we could correct for the amount of low negation in the data and would predict that the corrected curve for negation would coincide with the curve for questions.

Consider the data in Table 14. It shows the frequency of high and low negation in Late Middle English infinitives based on data from the Penn-Helsinki Parsed Corpus of Middle English.

Dates	not-to-verb	to-verb-not
1350-1420	10	4
1420-1500	4	10

Table 14: *not-to-verb* and *to-verb-not* order in negative infinitivals.

Apparently the two positions were equally likely. If we assume that the same equal distribution holds in tensed clauses, where ordinarily we cannot directly determine whether the position of negation high or low, and we perform the just-mentioned correction, we get the results in Figure 4, where the curve for negatives clearly does coincide with the curve for questions.

This analysis leads to the conclusion that there was a reanalysis in the middle of the 16th century. That would be the point when V-to-Tense was lost from the grammar of English. Subsequent to that point verb movement to the lower positions must have been lost and as it was, the frequency of *do* in negative declarative sentences would have risen. As imperatives do contain the lower projections, Mood and Aspect, loss of movement to these positions would have forced the use of *do* even in imperatives. In fact, the jump in the imperative curve signals the point when the loss of lower verb movement begins.

Clearly, the analysis we have proposed leaves many questions unanswered and depends on assumptions that may prove untenable in the long run but it does allow us to account for several grammatical and quantitative patterns and to reconcile the continued existence of apparent V-to-I raising with the claim rise of *do* traces the competition between grammars rather than merely a drift in usage frequencies. What we see is that the competition not only leads to the loss of one competitor at the end of the change but also to a loss in the middle of the change, which alters the grammatical character of the competition and the way that it evolves quantitatively.

## 2.4 The loss of verb-second word order in English.

Old English exhibits the verb-final cum verb-second (V2) word order pattern found in the other West Germanic languages, with some interesting variations, some of which we have discussed above in connection with the shift from OV to VO. Let us now turn to the second great syntactic change that occurred in the course of the transition from Old to modern English, the loss of the V2 constraint. To understand that development, it is necessary to appreciate a peculiar feature of V2 in Old English; namely that pronouns at the left edge of IP generally did not invert with the verb in the presence of a topicalized constituent, as would be expected on the basis of the other West Germanic languages. This exceptional behavior is most often seen subject pronouns, which are always found at the left edge of IP but it is also found with other pronouns, whenever they are scrambled leftward. The behavior of pronouns has led

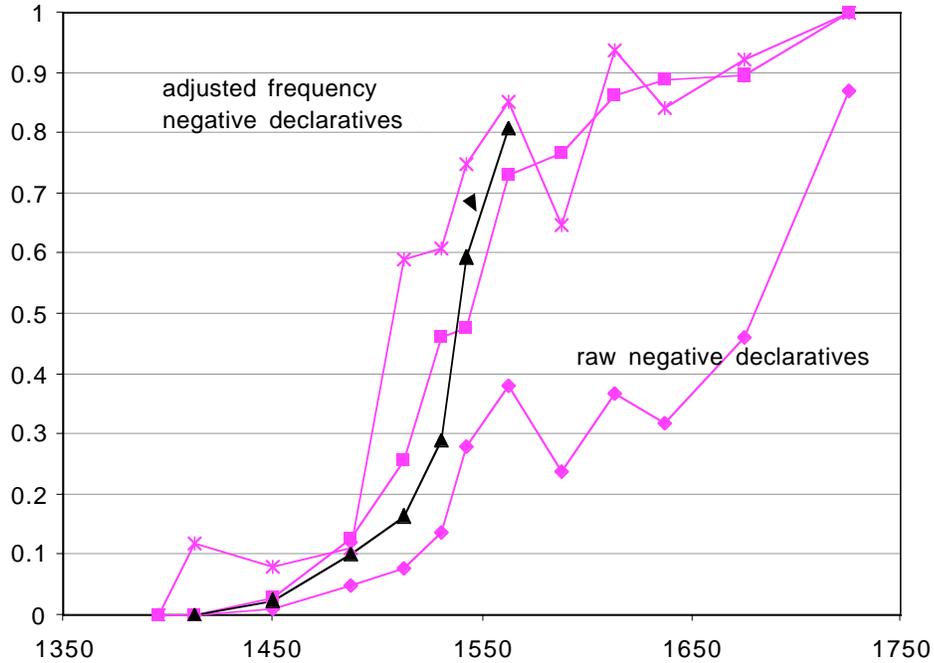


Figure 4: Adjusted frequency of *do* in negative declaratives compared to questions.

many students of Old English syntax to conclude that the verb in an Old English declarative does not move to the complementizer position but to a lower position, perhaps INFL and there is an extensive literature on the various options (see Kroch and Taylor (1997), Haeberli (2000), and Kemenade (1999), among others), which we will assume as background for this discussion.

Kemenade (1987), the first modern generative treatment of Old English syntax, notes that the Old English V2 constraint, including the special behavior of pronouns, survives intact into the first period of Middle English. It is not until after 1250, that we begin to see a real decline in V2 word order. This fact was, however, obscured until recently by the failure of traditional scholars to appreciate the grammatical character of the different behavior of full noun phrase and pronoun subjects. The picture is made more complicated by the apparent existence of a dialect difference within Old English between the language of the Scandinavian-dominated North and the Southern areas where the Vikings never established settlements. Kroch, Taylor and Ringe (2000) present evidence that the northern dialect lacked the special pronoun behavior found standard Old English, which was based on the dialect of the southern kingdom of Wessex. We also argue that the northern dialect's word order was the result of intimate contact between the invading Scandinavian and the native Anglo-Saxon populations in the North of England after 800 C.E. The pattern as we think it must have existed at the beginning of the Middle English period is illustrated in Table 15. The Benet text is the oldest prose text from the North and from an isolated area of the Yorkshire countryside. Though quite late, it shows the dialect difference with the

south very clearly. See Kroch et al. (2000) for further discussion.

Texts	NP subjects			Pronoun subjects		
	Number inverted	Number uniniv.	% inverted	Number inverted	Number uniniv.	% inverted
West Midlands (ca. 1225)	50	4	<b>93</b>	4	84	<b>05</b>
Northern St. Benet (ca. 1400)	7	0	<b>100</b>	58	3	<b>95</b>

Table 15: Frequency of subject-verb inversion in sentences with topicalized objects.

The dialect difference is further illustrated in Table 16. Here we have word order statistics from two manuscripts of the same 15th century text, the *Mirror of St. Edmund*, one northern copy and one from the southwest. The copies are generally identical, indicating that the manuscripts are not far apart in the stema but they do differ in word order in topicalized sentences, as the table shows.

Texts	NP subjects			Pronoun subjects		
	Number inverted	Number uniniv.	% inverted	Number inverted	Number uniniv.	% inverted
Northern St. Edmund	8	0	<b>100</b>	16	9	<b>64</b>
Southern St. Edmund	12	1	<b>92</b>	5	13	<b>28</b>

Table 16: Frequency of subject-verb inversion in sentences with topicalized objects in two manuscripts of the *Mirror of St. Edmund*.

The tendency for northern texts to show more V2-type subject-verb inversion with pronouns than southern texts, even during the loss of V2 further research has shown to be quite general. However, there is also a rather surprising tendency for southern texts to show more inversion with pronouns over time, even as the rate of inversion with noun phrase subjects declines. By the early modern period, there is not much difference between noun phrase and pronoun subjects. Neither invert very often but both do so on occasion, as can be seen in the late 16th and early 17th century examples in (24) – (27), taken from the Helsinki corpus:

- (24) a. This doute I not ... (BOETHEL,64.136)  
 b. These things said hee in the Synagogue, as he taught in Capernaum. (AUTHNEW,VI,40J.501)
- (25) a. This I doubt not ... (BOETHPR,138.153)  
 b. But these things I say, that ye might be saued. (AUTHNEW,V,20J.387)
- (26) a. This beginning of miracles did Iesus in Cana of Galilee, (AUTHNEW,II,1J.120)

- b. and loke whatsoever he lerned by day, that did Simon learne of him alwaies at nighte ... (FORMAN,9.171)
- (27)
- a. This lawyers disputed ... (EVELYN,903.124)
  - b. But this Disease Providence doth cure by applying the powerful Medicine of Money. (BOETHPR,199.426)

Daniel Johnson and Laura Whitton (2002) recently carried out an extensive study at Penn of the rates of inversion in sentences with topicalized direct objects as found in the texts of the Middle English and early Modern English parsed corpora that we have been building. Their results for the non-Northern texts in the corpora are shown in Figure 5:

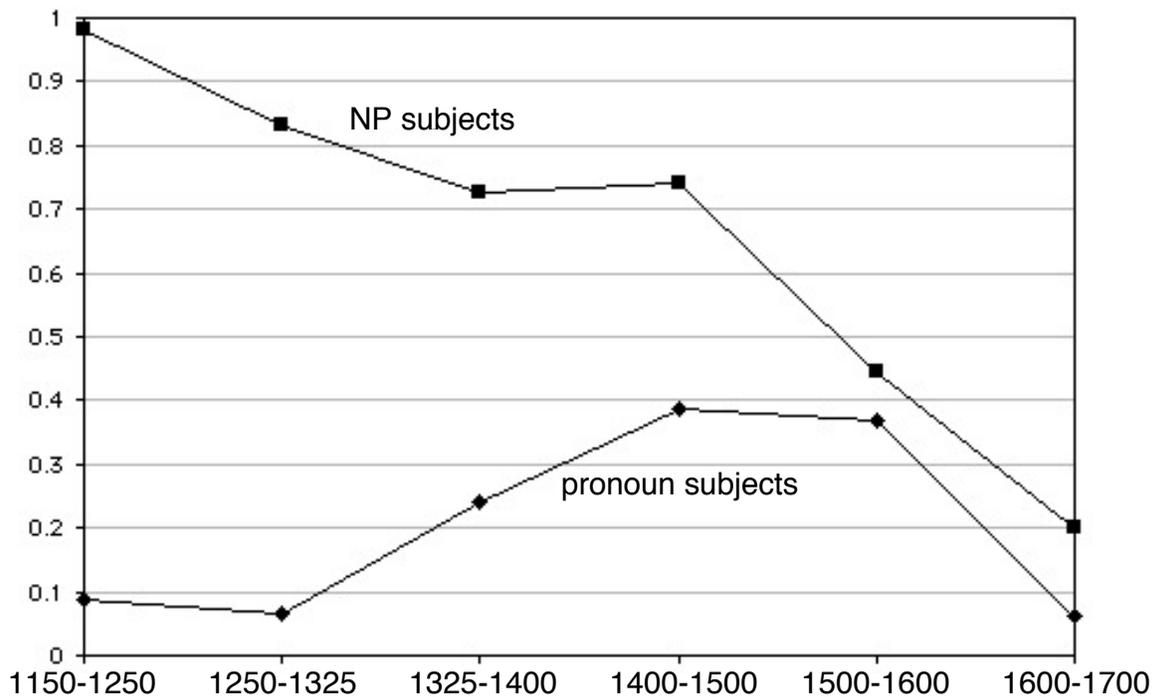


Figure 5: Frequency of subject-verb inversion in the PPCME2 and PPCEModE – full noun phrase versus pronoun subjects (from Johnson and Whitton 2002).

As the figure shows, the rise in the frequency of inversion in sentences with pronominal subjects between periods 1250 and 1500 is substantial and does not follow the slightly downward trend found with noun phrase subjects. As the figure again shows, the difference between pronoun and noun phrase subjects has largely disappeared by the 16th century. As far as we can tell, this behavior is hard to account under a drift model since we would not expect the frequencies of these two contexts to drift in opposite directions, given that they wind up both losing inverted word order at the same time. On the other hand, it is quite easy to build a model based on contact and competition between grammars that yields an evolution of frequencies quite close to the observed data.

Here is a scenario that yields the desired result. Suppose, following a suggestion in Kroch, Taylor and Ringe (2000), that contact between the Northern and Southern dialects led to the creation of the modern non-V2 grammar of English out of the misunderstandings by Northern and Southern speakers of the grammars underlying each other's speech. This contact would, of course, have been extremely extensive in the Middle English period since the Scandinavian-influenced and native Anglo-Saxon areas were in contact at least all along the border of the Danelaw, which cut through the whole of the midlands. Northern speakers might have misunderstood the non-inversion with pronoun subjects that they heard in the speech of Southerners as resulting from a non-V2 grammar and they might have learned this non-V2 language as a second dialect. Both child and adult learners might have engaged in this learning but the mislearning of the second dialect as a non-V2 language instead the correct learning of the Southern dialect indicates that learning by adults was the determinant factor. Because of the mislearning, northerners who used this second dialect would have failed to invert with full noun phrases as well as with pronouns. Southerners, hearing this non-inversion with noun phrase subjects, might in turn have learned it as a second dialect, leading to a situation where everyone in England had access to a non-V2 grammar. Note that the scenario assigns a crucial role to Northerners. Southerners who learned the Northern dialect might have learned to invert pronoun subjects but there is nothing in either the Northern dialect or their own system that would have led them to a grammar without inversion for full noun phrase subjects.

Further evolution would depend on how these grammars diffused. It is well-known that Northern linguistic features diffused southward during the Middle English period including such grammatical features as the third singular present tense *-s* suffix, which replaced the southern and Old English *-th*, and the third person plural pronoun *they*, which replaced the southern and Old English *hi*. Given the general southward diffusion of Northern features, it would not be surprising if the Northern V2 pattern also diffused southward. This would account for the rise in inversion with pronouns. On the other hand, it does not explain why the frequency of inversion with noun phrase subjects declines, since both V2 systems license it. The decline can only be the result of the spread of the modern English non-V2 grammar at the expense of both V2 grammars. In other words, we have here a case of three-way grammar competition. If we model the data in Figure 5, with a trinomial logistic regression, the mathematical form that corresponds to such a three-way competition, the best fit to the data assigns a fitness measure of .4 to the modern English grammar, of .35 to the Northern V2 grammar and of .25 to the Southern V2 grammar, where .33 corresponds to no fitness advantage. Over the long term, a system evolving in accord with these numbers will drive out both the Northern and the Southern grammars but over a shorter run, the Northern grammar will spread at the expense of the Southern grammar, leading to a temporary increase in its use. The temporal evolution of the logistic model, as fit to the data, is shown in 6:

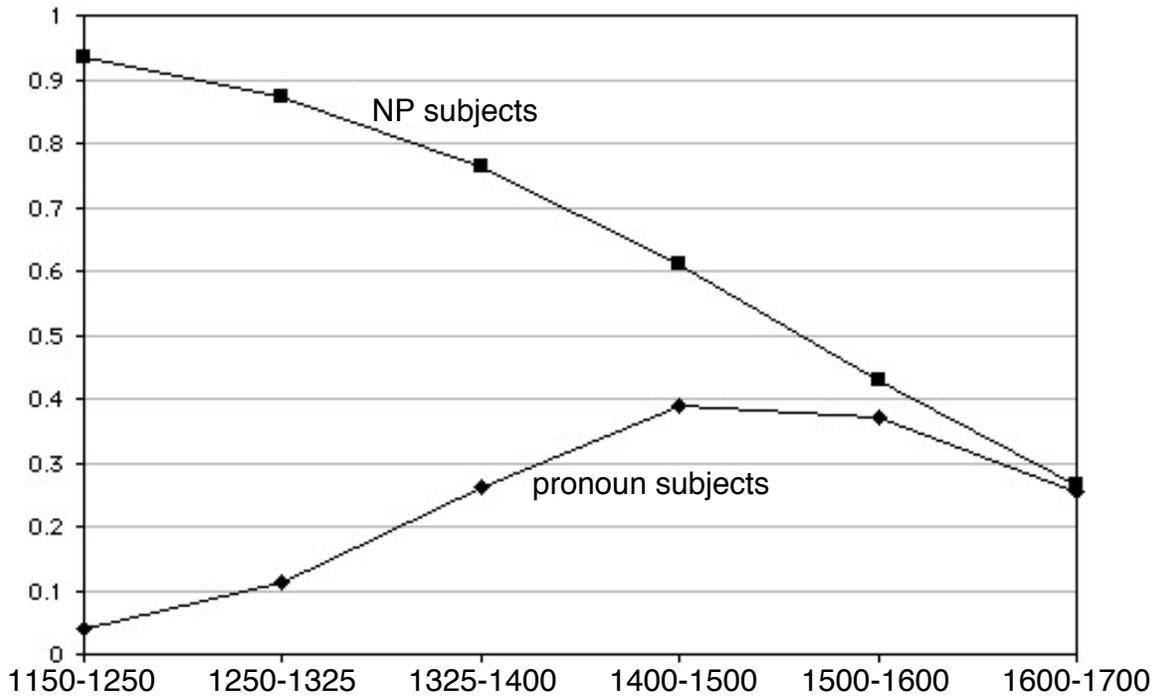


Figure 6: Model of the frequency of subject-verb inversion in the PPCME2 as three-way logistic competition between Northern V2, Southern V2 and Modern English grammars.

We can see by inspection that the model fits the data quite well, supporting, though of course indirectly and mediated by a number of assumptions that would benefit from further support, our general claim that the frequency evolution found in texts in cases of syntactic change reflects not drift but competition among grammars, whose origin is a function of variability in language learning outcomes.

### 3 Imperfect learning

We now move on to the problem of motivating the claim that imperfect language learning, even by children acquiring their first language can be a source of language change. It is, of course, well known that adult second-language learners make many errors and that the end state of their learning remains error-prone. The result of this is that communities with a high proportion of non-native speakers may present children with primary data that is highly skewed relative to what they would hear in an entirely native environment. This skewing will affect what these children learn. In some cases, the errors of the parents will not be corrected in the subsequent generation resulting in a changed language in that community. The type of these errors is likely to reflect the native language of the adult learners and preserved errors will appear to subsequent analysts as “substratum effects.” If these substratum effects spread, they

can become part of the standard language, no longer recognized as foreign. Such cases are difficult to document in the historical record because by the time that a grammatical change appears in the written language, its connection to the social context in which it arose is no longer demonstrable. It is possible, however, that such changes as the rise of INFL-medial word order in Yiddish and even of VO word order in English are due to such effects. In the case of Yiddish, the scenarios that support this hypothesis are various because the Jews who first learned German – this being the source of Yiddish – may have spoken either a Semitic or a Romance dialect. In either case, however, the language they spoke had a left-headed IP, which learners would have tended to impose on the German they were learning. The fact that Early New High German exhibited a high frequency of DP and PP extraposition would have reduced the evidence that IP was right-headed in German and might have contributed to the maintenance of left-headed IP as an option in subsequent generations. Among Jews who subsequently moved to Slavic-speaking territory, the possibility of convergence with German would have been lost and, indeed, further divergence occurred (Santorini 1992, 1995). In the English case, the evidence for contact as the source of VO word order is thin, but the dating of the first evidence for VO is consistent with an account based on contact with the Old Norse of the Viking invaders of the 9th century. It may become possible to pursue this case further as our knowledge of the quantitative patterns in Old English texts improves.

While the error rate in second-language acquisition by adults is high, however, the opposite is true for first-language acquisition by children. The relative stability of languages over time is enough to demonstrate this, even if we didn't have the obvious evidence of contemporary child language acquisition before us. It would be wrong, however, to say that children are error-free learners. The acquisition literature is full of descriptions of child errors. Unfortunately, the discussion is often not as helpful for our purposes as one would like. Necessarily, the primary concern of developmental psycholinguists is accounting for the *success* of children in learning their native language. Errors are viewed either as evidence for a specific linguistic or psycholinguistic theory – for example, anti-nativism or a maturational theory of language acquisition – or as a consequence of the lag between learning and fully skilled production. In thinking about language change, however, we should be interested in a different feature of learner errors; namely, the possibility that they might be heard by other learners and treated by them as possible target forms. To the extent that learner errors are incompatible with universal grammar or other powerful constraints on adult language, such a development is, of course, impossible; but many learner errors are perfectly compatible with universal grammar, reflecting grammatical properties that hold of other languages, though not the target. If the evidence in the target language for the grammatical property at issue is weak, then there might be a significant probability of a child's error spreading to other members of his/her speech community. And if this error is common in child language, then the probability of it leading to language change might be substantial.

Some years ago, H. Thraínsson (personal communication) sketched out a possible scenario for the development of the so-called “dative sickness” in Icelandic that can

serve to illustrate the perspective I am laying out. Icelandic has quirky case subjects, which are most often in the dative case. There are, in addition, a few verbs that have quirky accusative subjects, and there is a tendency in the present-day language, strongly opposed by prescriptive grammarians and teachers, for the dative to spread as a quirky case for subjects at the expense of the accusative. The dative does not spread at the expense of the nominative, the default subject case. The highly marked character of quirky case accusatives is clear, and in one sense the current development is unsurprising. What is interesting is that the quirky accusatives apparently existed quite stably for a long time and that the dative sickness is a twentieth-century phenomenon. Thraínsson's speculation as to how this might have happened is that it is associated with the increasing urbanization of Icelandic society over the past century. The mechanism would be as follows: It is known that young Icelandic learners often substitute the dative for the accusative on quirky case subjects and that they learn not to do this only over many months. This pattern is generally seen as a case of overgeneralization, as happens with the past tense in the acquisition of English. In traditional Icelandic society, children lived in small family groups with a large amount of contact with adults in the course of daily life. In modern society the children spend many hours in pre-school and school settings where they are exposed to the speech of other children of their age to a much greater extent than before. This environment has the potential to cause one child's errors, like the dative sickness, to be treated as a model by other children and so to spread rapidly, especially where the error is a common one. The point of this scenario is that quirky accusative is a highly marked option for subjects and that the evidence for it is weak in frequency terms. For these reasons, it is acquired with some difficulty. When the conditions of acquisition are changed, it is no longer reliably acquired by the whole speech community, and a process of change is initiated.

### 3.1 The acquisition of case forms

It is, of course, impossible to test Thraínsson's scenario in any direct way, but it is plausible enough that we would like to be able to explore the general model that it presupposes and to look for indirect evidence linking change with acquisition errors. The model requires that two characteristics hold of language acquisition: first, that children make errors in acquisition which take time to correct and second, that errors propagate from one child to another. The latter property has not been investigated systematically, as far as I know, but anecdotal evidence for it abounds. For the former property, however, there is considerable evidence in the literature. This is particularly evident in the domain of morphology, where children clearly overgeneralize regular patterns. Beside the often cited case of the English past tense, we have studies like that of Clahsen (1984), who shows overgeneralization of the nominative case in accusative contexts and of the accusative in dative contexts in the speech of three German children between the ages of 26 and 42 months. The use of nominative for accusative is illustrated by the example in (28) and the use of accusative for dative in (29):

- (28) hab **der** keubeu noch nich auf  
 ich habe den Keubeu noch nicht aufgesetzt  
 “I haven’t put the cowboy on yet.”
- (29) die haben wir von **den** Christkind  
 die haben wir von dem Christkind gekriegt  
 “These, we got from the Christ Child.”

The degree of deviation of one three-year-old child’s language from the adult norm in the use of the dative case is summarized in Table 17:

	Syntactic function		Syntactic category	
	Verbal object	Prepositional object	Pronoun	Noun phrase
Accusative contexts	18	4		
Accusative forms			4	18
Dative forms			0	0
Dative contexts	7	21		
Accusative forms			4	8
Dative forms			3	13

Table 17: Accusative and dative case marking in the speech of a three-year-old German child.

Clahsen’s study shows that the accusative and dative are in competition in dative environments after the age of 30 months. This competition indicates that the process of replacement is underway and that it is not instantaneous. If for any reason this process of replacement does not take place for a given child, his/her adult grammar will exhibit a change relative to that of others in the community and this change is potentially a starting point for change in the community. Of course, in standard German the evidence for case marking is so strong that there is no prospect of overgeneralization leading to change in the adult grammar. At the same time, Tracy (1986) has given evidence that the course of acquisition of the German case system is influenced by the surface morphophonological patterns, with the more easily discriminable and more frequent distinctions being correctly learned first. Since the evidence for case marking in German is found primarily in the suffixes on determiners, one might expect change to be possible if these inflections become more difficult to discriminate, perhaps as a consequence of independent phonological changes. Such a process seems to have been responsible, at least in part, for the collapse of the inflectional system in the history of English and for its weakening in the history of Yiddish.

Russian is an especially interesting language in which to study errors made by children. Its morphology is rich and contains many minority morphological patterns. In the nominal domain, to choose just one area of interest for syntax, there is a striking difference between the acquisition of the case system as such and the acquisition of

the whole range of endings. The case system seems to be acquired in large part by the age of two, the earliest point at which it could manifest itself in production. But with minority pattern endings and in cases where there is great phonological similarity across morphemes, error-free performance can take years. Zakharova (1958) reports the results in addtocountertable1 Table 17 for two common noun classes. The table is taken from an English language summary of the research (Adshead 1979):

	Age in years			
	3.0	4.0	5.0	6.0
Declension in <i>-a</i>	13%	4%	2%	–
Hard masculine declension	15%	8%	3%	3%

Table 17: Percentage of incorrect case forms by age for 160 Russian children for two common declension classes.

For the oblique case in two less common noun classes, she reports the results in Table 18:

	Age in years			
	3.0	4.0	5.0	6.0
Soft feminine obliques	32%	25%	13%	5%
Soft masculine obliques	25%	16%	3%	–

Table 18: Percentage of incorrect case forms by age for 160 Russian children for oblique noun phrases in two less common declension classes.

## 3.2 Acquisition of syntax: competing grammars in child language

### 3.2.1 default case in Russian

There are many studies in the domain of syntax which show children making errors that subsequently undergo correction. Babyonyshev *et al.* (1994) have shown that Russian children often incorrectly assign nominative case, as in (30a), instead of the expected genitive, as in (30b), to the subjects of negated unaccusative verbs.

- (30) a. studenty v klasse okazalos'  
 students.NOM.PL in class turned-up  
 “The students turned up in class.”
- b. studentov v klasse ne okazalos'  
 students.GEN.PL in class NEG turned-up

“The students did not turn up in class.”

Their figures for children between the ages of 3;5 and 4;7 show this case assignment 40% of the time.

Polinsky (to appear), in her summary of work on the acquisition of Russian, reports the common use of the nominative for the dative with verbs that take dative subjects, though the age at which she observed this usage was much lower (1;5 - 2;0). Babyonshev’s work on the genitive of negation with unaccusative verbs has a possible connection with the decline of the genitive negation with transitive verbs. Polinsky reports that 34% of the direct objects of negated verbs appear in the genitive in 19th century texts but that in modern spoken Russian the frequency is only 17%. Children also use the accusative for the genitive in this environment, and further research might reveal this to be a case of acquisition errors leading to language change.

What is notable in these Russian cases, and also in the Icelandic dative sickness case mentioned above, is that children are failing to apply consistently a marked morphological requirement and fall back on the default case that would apply to the structure in the absence of the requirement. Thus, we get nominative for genitive with unaccusative verbs and accusative for genitive on the objects of negated verbs. The Icelandic case is more striking since Icelandic children substitute one quirky case for another, but they do it in a way that respects the hierarchy of markedness, using the less marked form in place of the more marked one. Clearly, these morphosyntactic examples are cases of overgeneralization, like those found in the acquisition of morphological forms, and reflect the structure of morphological defaults in a language. Children seem not to wait for information regarding the markedness of given linguistic forms or contexts, but rather to jump to the conclusion that the default rules apply. Perhaps the most interesting property of this behavior is that it takes so long for the child to retreat from use of the default. It does not seem to be the case that children use the unmarked form only until they have learned the marked form. Instead, there is a long period of coexistence between the two forms, with the adult-like form replacing the default over a period of months or even years. This process looks very much like the grammar competition that occurs in the course of language change, though the time scale is, of course, very different.

### 3.2.2 *English wh-* question inversion

It is not the case that all children go through the same phases of overgeneralization or that all children take similar lengths of time to approach the target. In one case of the acquisition of syntax by an English-speaking child that was documented in great detail, William and Teresa Labov (1976) recorded thousands of *wh-* questions produced by one of their daughters. They found that the child’s earliest *wh-* questions did not systematically exhibit inversion of subject and verb, as required in the adult language, and that this pattern, which was not found with *yes/no* questions, persisted for more than a year, gradually declining in frequency over time and disappearing completely only after 18 months. The examples in (31) and (32) below illustrate

the variation. They and others were all uttered within minutes of each other on a day when the child was 3 years and 10 months old. The child showed no symptoms of language impairment, and the course of her language acquisition was otherwise rapid and normal. Another child in the family, however, whose usage was monitored subsequently, did not exhibit an extended non-inversion stage.

- (31) a. Where this comes from?
- b. What *not* starts with?
- c. How them buy their tents?
- d. Why we can't wear sneakers?
- (32) a. Where's Philadelphia?
- b. What does the sun do to snow?
- c. How do babies get inside the mommies?
- d. Why are we going down?

The graph in Figure 7 exhibits the time course of acquisition based on a sample of approximately 3000 *wh*- questions.

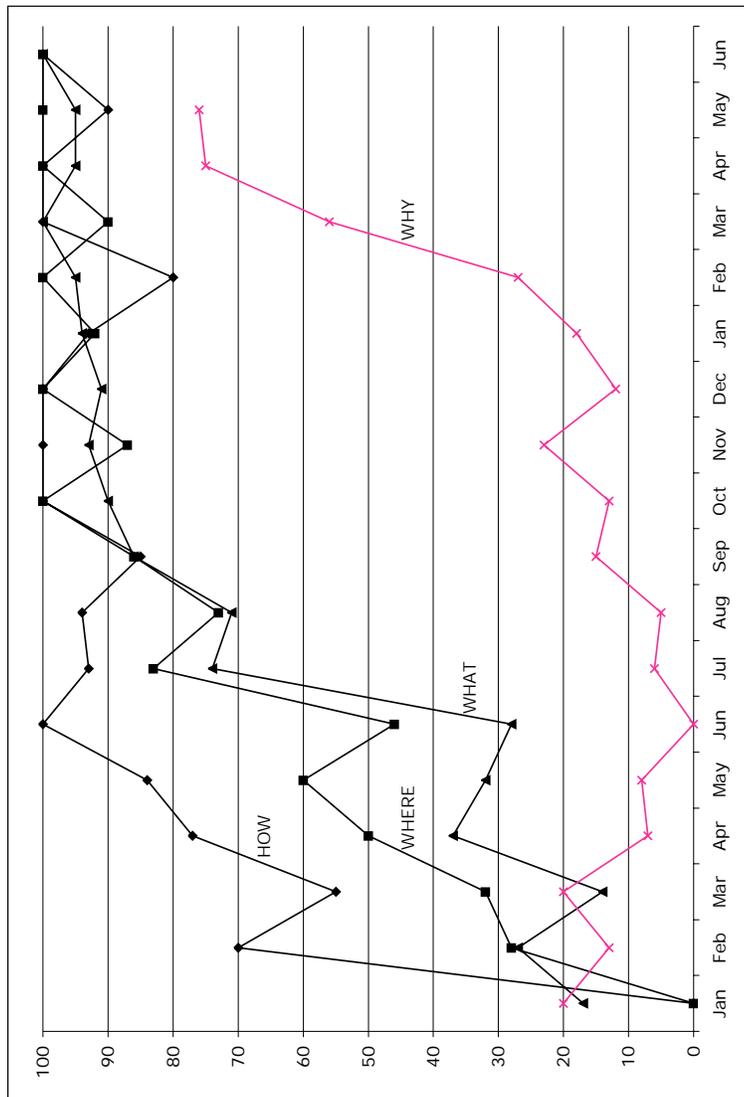


Figure 7: Acquisition of *wh*- questions by Labov daughter

Interestingly, it is one question word, *why*, that is responsible for most of the deviation from adult English and the graph shows that the Labov child's *why* questions actually move away from the target for a period of six months while other *wh*- questions are approaching the adult target. Indeed, at the end of these six months, *why* questions almost never invert, suggesting that the child has acquired a grammar different from

the English adult target. This conclusion is strongly supported by another diary study of *wh*- question inversion carried out by Thornton (citealpthornton04), who also observed her own daughter. Thornton’s daughter, like the Labovs’, distinguished *why* questions from others and also took many months to acquire systematic inversion with *why*, as illustrated in the graph in Figure 8.

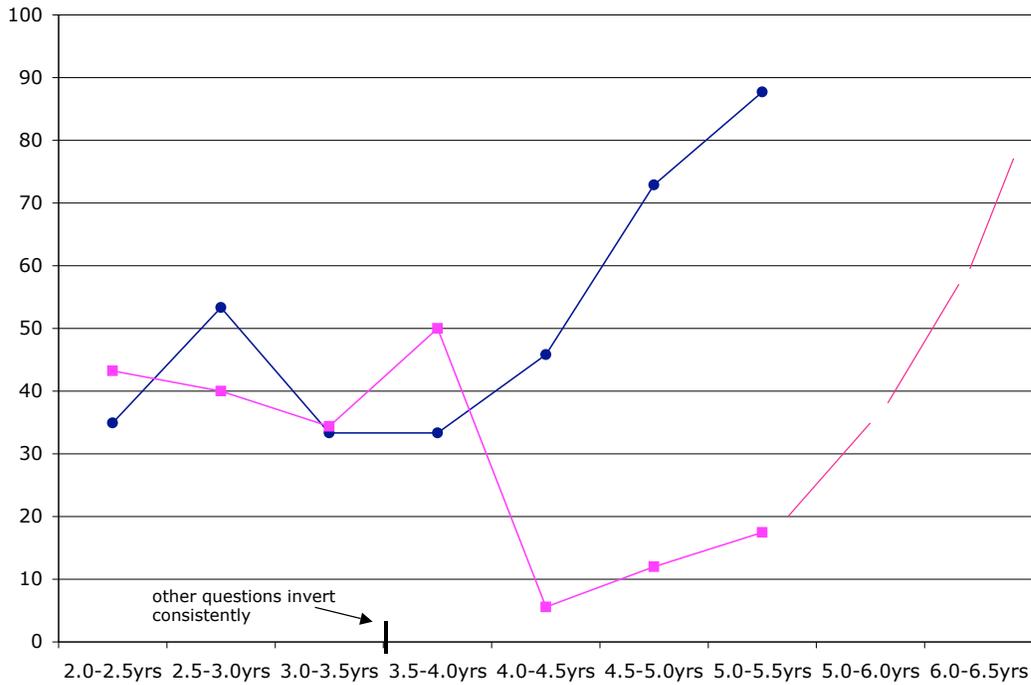


Figure 8: Acquisition of inversion in *wh*- questions by A. L.

Here are some examples of non-inversion with *why* question in the Thornton child data:

- (33) a. Why that boy is looking at us?  
 b. Why you gotta went to a conference?  
 c. Why we can't go in the upstairs bed?  
 d. Why anyone's not sitting in that seat?
- (34) a. Why you didn't bring me to Margie's house?  
 b. Why she doesn't like bananas?  
 c. Why Daddy's never coming to watch?

Thornton makes an important contribution to our understanding of the behavior of these children by demonstrating that her child pattern of inversion and non-inversion follows very closely a target pattern in another language, namely Italian as described by Luigi Rizzi (Rizzi 1999). Like the English of the two children we are describing, Italian requires subject-verb inversion with most *wh*- words, but not with *perché* (“why”):

- (35) a. Che cosa ha fatto Gianni?  
 what thing has done Gianni  
 b. \*Che cosa Gianni ha fatto
- (36) a. Perché Gianni è venuto?  
 why Gianni is come  
 b. Perché è venuto Gianni?

Thornton shows that her daughter's *why* questions pattern in detail like adult *why* questions in Italian. She lists the following detailed syntactic parallels:

- (i) compatibility of focus phrases immediately following *why*, (ii) use of topic phrases [immediately preceding and following *why*], (iii) the possibility of a preceding subordinate clause, (iv) obligatory inversion in long-distance questions with tensed embedded clauses, and (v) optional inversion in questions with infinitival complements.

Even leaving aside the precise syntactic analysis of the Italian case, we can agree with Thornton that her child's behavior reflects, not immaturity, but a syntactic parameter set to the same value as that of adult Italian. Presumably, the Labov child's grammar is similar, though the Italian facts were not available at the time of their study so that we cannot directly compare their description to Thornton's. Thornton does not answer the question of how a child in a monolingual English-speaking environment could wind up with an Italian parameter setting but she does note that only some children go through an Italian-like stage. She suggests that the Italian setting may represent some kind default, due presumably to the semantics of *why*, which is an operator on entire CP's and not semantically associated with any lower position inside the clause. Even so, the English-speaking child must ignore the input data to accept the default setting and, having made this wrong choice, s/he cannot retreat from it except through a long period of competition.

### 3.2.3 German V2

Natascha Müller (1996) reports on Ivar, a child who was videotaped every two weeks between the ages of 1 1/2 and 5 years. Beginning at the age of 3;0, Ivar exhibits verb-second word order in subordinate clauses for a period of two years, with a gradual shift toward the adult pattern. The examples in (37) and (38) illustrate the pattern, which occurs both with subjects and non-subjects in the topic position:

- (37) a. erst wenn wir sind fertig mit das (3;4,9)  
 first when we are finished with that  
 b. guck mal wie das ist groß (3;8,1)  
 look once how that is big

- (38) a. wenn da kommen andere schiffe dann gehen die dagegen (3;4,23)  
 when there come other boats then go these against  
 b. ... daß dann sagt er ... (3;5,7)  
 that then says he

Like other German children, Ivar learned that the language was verb-final by the age of 2 1/2 so that even example (37) must involve verb movement to a left-edge functional projection. Müller finds that correct and deviant placement of subordinate-clause finite verbs coexist in Ivar's speech with the same subordinating complementizer and the same finite verb. Thus, his behavior cannot in any obvious way be reduced to simple stage-by-stage learning. Although Ivar is bilingual in German and French, any effect of French on Ivar's learning of German would have to be indirect since his French never exhibits V2 word order. Perhaps Ivar has wrongly generalized V-to-I movement from French to German. Again, not all German-French children deviate from the target in this way. Müller points out that the deviations from the target are all consistent with UG. In particular, Ivar's subordinate clauses look similar to those of adult Yiddish, just as the Labov and Thornton children's deviant *wh*- questions look like modern Italian.

Müller believes that Ivar and other children in her sample initially adopt incorrect parameter settings and/or grammatical categorizations of functional projections, which on a Borer-type notion of parameter may amount to the same thing. She notes that it is not at all obvious how the children manage to recover from this error, given the the generally assumed insensitivity of language acquisition to negative evidence. The slow rate of change in Müller's children and the Labov child look like what happens in children's recovery from overgeneralization in morphology. Again, the pattern resembles, on a different time scale, the competition between grammars found in historical change. We have argued elsewhere (Kroch 1994) that such competition will normally resolve in favor of one form or another, due to the operation of Aronoff's *Blocking Effect*. In the case of first-language acquisition, the form that is reinforced by the speech community surrounding the child will normally win out. Nonetheless, since the errant grammars induced by children are consistent with UG, they are potential sources for change under favorable circumstances.

### 3.2.4 English *tough*-movement

There are interesting cases where the initial acquisition by children of a parameter setting different from that of the adult grammar is the rule rather than the exception. As in the cases discussed above, the adult target is approached slowly over a period of many months. One striking case of this type is the acquisition of *tough*-movement. Beginning with Carol Chomsky's 1969 study, it has been reported that children often interpret a sentence like (39) as meaning (40a) rather than (40b), the interpretation that adults would give to the sentence:

- (39) The king was hard to draw.

- (40) a. The king<sub>i</sub> was hard [ \_\_\_<sub>i</sub> to draw ]  
 “It was hard for the king to draw (something).”  
 b. The king<sub>i</sub> was hard [ to draw \_\_\_<sub>i</sub> ]  
 “It was hard (for someone) to draw the king.”

Initially, students of the phenomenon hypothesized that children went through stages in the acquisition of *tough*-movement, first assigning subject-gap readings and then later shifting to the adult object-gap reading. More recent work (McKee 1997, Anderson 2001a,b), has shown that children understand and use the adult reading from the earliest ages at which they can be tested, so that there is no direct evidence for the stages account. Since, however, the object-gap reading is the only one that occurs more than sporadically in the input language, it still seems necessary to say that the child initially misanalyzes the input. Apparently, the misanalysis occurs at a very young age, before the child uses the construction in its own speech. By the time that the child is old enough to test, it has learned the adult analysis, but it does not immediately abandon its initial hypothesis. Hence, the ambiguity that shows up in all studies of the phenomenon. Indeed, all work on the construction agrees that the two readings alternate in child usage for many months, with the adult usage slowly replacing the child’s initial erroneous analysis, as in the studies of Ivar and the Labov daughter.

Also as in the earlier cases, the child’s usage here is compatible with Universal Grammar, a point made explicitly by Anderson. She points out that child usage corresponds to what was possible in earlier English, citing examples like the following one from Jonathan Swift to show that early Modern English allowed the subject-gap reading:

- (41) He found the natives very hard to believe that the fact was possible.  
 (1726-7 Swift *Gulliver’s Travels*, book III.x)

Some modern languages as well, for example Brazilian Portuguese (Galves 2001), allow both subject-gap and object-gap readings, reinforcing Anderson’s point.

It appears that the subject-gap interpretation in child English must again be an overgeneralization of a default, as happens in the acquisition of morphology. Only this assumption allows us to understand why children so consistently make the error. Perhaps the default involves the control of PRO, which is assumed to be by the next higher subject in the absence of positive counterevidence. This hypothesis also helps to account for the widely noticed fact that adult second-language learners of English often use *tough*-movement predicates with a subject-gap interpretation, just as children do.

How slow the move toward the target is in the *tough*-movement case can be seen in the following graph from Anderson. McKee has similar data.

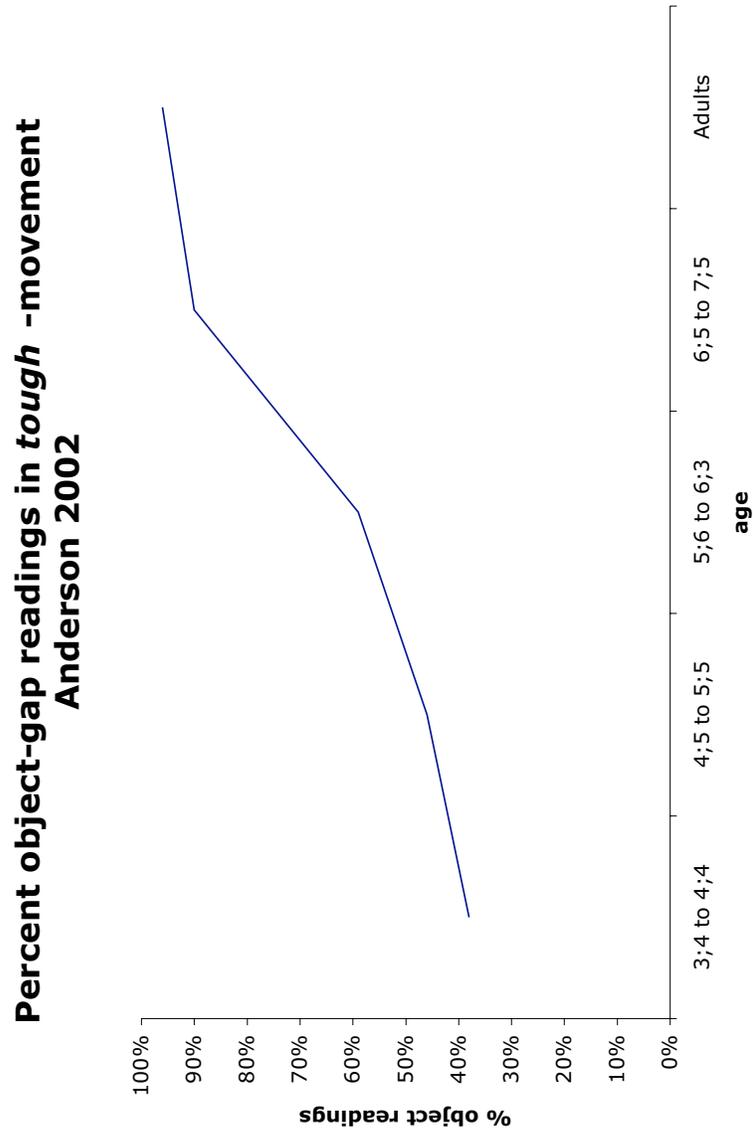


Figure 8: Time course of children’s usage of *tough*-movement.

Anderson notes that the move toward the target is a retreat from a system that allows a superset of grammatical sentences to a subset, just what ought to be impossible according to the subset principle (Berwick 1985). However, it is just this situation that we always find in retreats from overgeneralization and in grammar competition, where the decline and loss of one of two competing forms is necessarily a move to a

subset language, extensionally viewed. The length of time that it takes for the child to approach the target, as well as the formal character of the learning involved, argues that we are looking at competition among competing grammatical options, namely subject- and object-gap only grammars. The object gap option eventually wins out because it is the one that is reinforced by the linguistic behavior of the rest of the community, just as in the cases of parameter missetting that we have discussed. Here the relevant parameter is the syntactic structure associated with the class of *tough* predicates. If we accept an approach to parameters like that of Borer (1983), under which parameters are properties of individual items in the grammatical inventory of a language (see also Kroch 1994), then the *tough*-movement case is entirely parallel to other cases of grammar competition.

The process of correction through grammar competition need not lead to the complete loss of the disfavored form from a speech community, depending on factors that have not been studied; for example, whether the disfavored form receives any reinforcement in speech and how sensitive different individuals are to the language environment. A search on Google will reveal that subject gap *tough* sentences occur with a certain frequency in American English. Here is one example, which as far as one can tell, was produced by a native speaker:

- (42) From a maneuverability perspective, the airplane is very easy to go to regimes that other aircraft can't go. Lt. Col. Jeffrey Harrigian, USAF, 43rd Fighter Squadron commander quoted in the AETCNS web newsletter #081803266, Aug. 18, 2003.

### 3.2.5 Swedish negation

A final case in which children initially adopt a parameter setting not supported by target language input is the order of auxiliary verbs with respect to adverbs and negation in Swedish subordinate clauses (thanks to Dianne Jonas for directing me to this case). In main clauses the tensed verb always comes before negation due to the verb-second constraint but that constraint does not hold in subordinate clauses, whose word order, therefore, might be more difficult to learn. Based on several empirical studies, Håkansson and Collberg (1994) show that Swedish children, when they first begin to produce subordinate clauses at the end of their third year, consistently distinguish main verbs from auxiliaries in their placement relative to negation in subordinate clauses, though they never do so in main clauses. In their behavior the children are unlike adults. The adult language consistently places the tensed verb *after* negation in a (non-CP recursion) subordinate clause whether or not it is an auxiliary:

- (43) a. Det var han som inte kommer.  
           it was he that not comes  
       b. Det var han som inte ska komma.  
           it was he that not shall come

The children, in contrast, place tensed main verbs after negation but tensed auxiliaries before it:

- (44) a. Det var han som inte kommer.  
it was he that not comes  
b. Det var han som ska inte komma.  
it was he that shall not come

Håkansson and Collberg propose that the reason for the child behavior is that there is a default parameter setting in Universal Grammar under which auxiliaries are generated in INFL, perhaps for semantic reasons (Tomaselli and Schwartz 1990), and that parallel behavior of main verbs and auxiliaries must be learned from the input. From our point of view, this case, like the others, shows the children ignoring input data in making their initial parameter settings. It seems to take about six months for the child to switch completely to the adult system.

The Swedish child language case provides a useful perspective on the historical evolution of the English auxiliary. It has long been thought (Lightfoot 1979, Roberts 1985, 1993, Warner 1993) that the morphological peculiarities of the English modals and the loss of infinitival morphology played a decisive role in the development of a syntactic class of auxiliary verbs in English. However, although none of the morphosyntactic developments that affected the English modals are found in Swedish, Swedish children initially conclude that auxiliaries and main verbs are syntactically distinct. Apparently, the semantics of the auxiliaries is responsible for the default parameter setting, which makes the English auxiliary class, not, as is often thought, a special case but the default case. The syntactic collapse of auxiliaries with main verbs generally found in the languages of Europe becomes the marked case syntactically. At the same time, the markedness hierarchy seems to go in the other direction in the morphology, where parallel treatment of verbal inflection across main verbs and auxiliaries is unmarked. When the morphology is sufficiently rich, the syntax is apparently forced to follow it and to treat main verbs and auxiliaries alike but when the morphology weakens, the syntactic default appears.

## 4 Conclusion

In the first part of this paper. I presented evidence that language syntax does not change due to drifts in usage frequencies. On the contrary, in several cases that we have studied, there is evidence of grammar change at the beginning of a period of shift in these frequencies. This evidence led me to the claim that changes in grammar cause changes in frequencies through the playing out over time of competition between innovative and conservative grammars. If this proposal is correct, however, a puzzle arises: Where do grammar changes come from? The model appears to require that they arise out of nowhere. The solution to the puzzle, I have suggested, lies in a more complex understanding of the nature of language acquisition.

The studies of the acquisition of syntax that I described in the second half of this paper show clearly that the learning of syntax by children is not a simple process of closer and closer approximation to the adult target. Defaults provided by Universal Grammar or universal semantic/conceptual categories compete with the actual linguistic facts of the target language in determining learners' syntactic analyses and sometimes lead them to acquire non-target grammatical features. In addition, there are certainly random effects of differences in exposure to input data and differences in individual learners that lead to non-target learning, though these effects, unless extremely gross, are difficult to detect with current methods. In any case, once non-target features are acquired, whatever the reason, a secondary process of adaptation to the target system takes hold and over time guides the learner back toward the target. This secondary learning, however, seems to be of a different character than the initial acquisition. The target grammatical feature replaces the non-target one only over a fairly extended period of time and the replacement shows the characteristic course of grammar competition in historical change.

The fact that linguistic properties are often diachronically stable indicates that the process of acquisition, complex though it may be, is largely homeostatic but, as with complex systems generally, only within limits and only probabilistically. The complexity of the acquisition process with its interplay of competing factors leaves room for non-target features occasionally to survive the process of correction to the target and to spread. We do not presently know how to measure the strength of the various factors that influence the likelihood that deviations from the target will spread and, of course, we will never be able to predict the course of language history, given the complexity of the system. Still, we are making progress in understanding what factors are in play as languages move through time.

## References

- David Adshead. 1979. The Russian child's acquisition of morphology. *Journal of Russian Studies*, 37:21–24.
- Cynthia Allen. 2000. Obsolescence and sudden death in syntax: the decline of verb-final order in early middle english. In Ricardo Bermudez-Otero, David Denison, Richard Hogg, and C. B. McCully, editors, *Generative Theory and Corpus Studies: a Dialogue from 10th ICEHL*, pp. 3–25. Mouton de Gruyter, Berlin.
- Deborah Anderson. 2001a. The acquisition of null operator structures in English. In João Costa and Maria Freitas, editors, *Proceedings of GALA 2001*. Associação Portuguesa de Linguística.
- Deborah Anderson. 2001b. Structural ambiguity in early English tough constructions: Are child grammars deficient or simply different from adult grammars? In *Working Papers in English and Applied Linguistics*, volume 8, pp. 1–24. Research Centre for English and Applied Linguistics, University of Cambridge.

- Maria Babyonyshev, R. Fein, J. Ganger, and S. Avrutin. 1994. Maturation of A-chains: preliminary new evidence from the acquisition of Russian unaccusatives. *MIT Working Papers in Linguistics*, 23:1–20.
- Robert Berwick. 1985. *The Acquisition of Syntactic Knowledge*. MIT Press, Cambridge, MA.
- Jonathan Bobaljik. 2002. Realizing Germanic inflection: Why morphology does not drive syntax. *Journal of Comparative Germanic Linguistics*, 6.2:129–167.
- Hagit Borer. 1983. *Parametric syntax. Case studies in Semitic and Romance languages*. Studies in generative grammar 13. Foris, Dordrecht.
- William Michael Canale. 1978. *Word order change in Old English: base reanalysis in generative grammar*. PhD thesis, University of Toronto.
- Carol Chomsky. 1969. *The Acquisition of Syntax in Children from 5 to 10*. Number 57 in Research Monograph. The MIT Press, Cambridge, MA.
- Harald Clahsen. 1984. Der Erwerb von Kasusmarkierungen in der deutschen Kindersprache. *Linguistische Berichte*, 89:1–32.
- Harald Clahsen. 1991. Constraints on parameter setting: A grammatical analysis of some acquisition stages in German child language. *Language Acquisition*, 1(4): 361–391.
- Robin Clark and Ian Roberts. 1993. A computational model of language learnability and language change. *Linguistic Inquiry*, 24:299–345.
- Alvar Ellegård. 1953. *The auxiliary do: the establishment and regulation of its use in English*. Almqvist & Wiksell, Stockholm.
- Charlotte Galves. 2001. *Ensaio sobre as gramáticas do português*. Editora da UNICAMP, Campinas, Brazil.
- Eric Haeberli. 2000. Adjuncts and the syntax of subjects in Old and Middle English. In Susan Pintzuk, George Tsoulas, and Anthony Warner, editors, *Diachronic syntax: models and mechanisms*, pp. 109–131. Oxford University Press.
- Gisela Håkansson and Sheila Dooley Collberg. 1994. The preference for modal + neg: and I2 perspective applied to Swedish 11 children. *Second Languages Research*, 10 (2):95–124.
- Chung-hye Han and Anthony Kroch. 2000. The rise of do-support in English: implications for clause structure. In M. Hirotani, A. Coetzee, N. Hall, and J-Y Kim, editors, *Proceedings of the 30th Meeting of the North East Linguistics Society*, pp. 311–325. Graduate Linguistic Student Association, Amherst, MA.

- Thorbyörg Hróarsdóttir. 2002. Explaining language change: A three step process. In Susann Fischer Alexiadou, Artemis and Melita Stavrou, editors, *Papers from the workshop “Language Change from a Generative Perspective”*, volume 19 of *Linguistics in Potsdam*, pp. 103–141. Institut für Linguistik.
- Thorbyörg Hróarsdóttir. 2003. Language change and language acquisition. *Nordlyd*, 31:116–131.
- Thorbyörg Hróarsdóttir. 2004. Cues and expressions. *Nordlyd*, 32:135–155.
- Daniel Johnson and Laura Whitton. 2002. A corpus-based quantitative analysis of subject-verb inversion in Middle and Early Modern English. Ms. University of Pennsylvania.
- Ans van Kemenade. 1987. *Syntactic case and morphological case in the history of English*. Foris, Dordrecht.
- Ans van Kemenade. 1999. Sentential negation and clause structure in Old English. In Ingrid Tieken-Boon van Ostade, Gunnel Tottie, and Wim van der Wurff, editors, *Negation in the History of English*, pp. 147–165. Mouton, New York.
- Anthony Kroch. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1:199–244.
- Anthony Kroch. 1994. Morphosyntactic variation. In K. Beals, editor, *Proceedings of the thirtieth annual meeting of the Chicago Linguistics Society*, volume 2, pp. 180–201. Chicago Linguistics Society.
- Anthony Kroch and Ann Taylor. 1997. Verb movement in Old and Middle English: Dialect variation and language contact. In Ans van Kemenade and Nigel Vincent, editors, *Parameters of morphosyntactic change*, pp. 297–325. Cambridge University Press, Cambridge.
- Anthony Kroch, Ann Taylor, and Donald Ringe. 2000. The Middle English verb-second constraint: A case study in language contact and language change. In Susan C. Herring, Pieter van Reenen, and Lene Schoesler, editors, *Textual parameters in older languages*, Current issues in linguistic theory 1950, pp. 353–391. John Benjamins, Amsterdam/Philadelphia.
- William Labov and Teresa Labov. 1976. Learning the syntax of questions. In Robin Campbell and Philip Smith, editors, *Recent Advances in the Psychology of Language*, volume 4B of *NATO Conference Series, III: Human Factors*, pp. 1–44. Plenum, New York.
- David Lightfoot. 1979. *Principles of diachronic syntax*. Cambridge studies in linguistics 23. Cambridge University Press, Cambridge.

- David Lightfoot. 1993. Why UG needs a learning theory: triggering verb movement. In C. Jones, editor, *Historical Linguistics: Problems and Perspectives*, pp. 190–214. Longman, London.
- David Lightfoot. 1999. *The development of language: Acquisition, change, and evolution*. Blackwell, Malden, MA.
- Cecile McKee. 1997. Some adjectives are ‘easy’ and some are not. *Lexicology*, 3(1): 59–87.
- Natascha Müller. 1996. V2 in first-language acquisition: early child grammars fall within the range of universal grammar. *Linguistics*, 34:993–1028.
- Zvi Penner. 1992. The ban on parameter resetting, default mechanisms, and the acquisition of V2 in Bernese Swiss German. In Jürgen Meisel, editor, *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Development*, pp. 245–281. Kluwer, Dordrecht.
- Susan Pintzuk. 1997. From OV to VO in the history of English. Ms., University of York.
- Susan Pintzuk. 2002. Verb-object order in Old English: variation as grammatical competition. In David Lightfoot, editor, *Syntactic Effects of Morphological Change*, pp. 275–299. Oxford University Press, Oxford.
- Susan Pintzuk and Anthony Kroch. 1989. The rightward movement of complements and adjuncts in the Old English of Beowulf. *Language Variation and Change*, 1: 115–143.
- Maria Polinsky. to appear. Language acquisition: full and incomplete scenarios. In George Fowler, editor, *Slavic Linguistics in the Twenty-First Century*. Slavica, Bloomington, In.
- Luigi Rizzi. 1999. On the position “int(errogative)” in the left periphery of the clause. University of Siena.
- Ian Roberts. 1985. Agreement parameters and the development of the English modal auxiliaries. *Natural Language and Linguistic Theory*, 3:21–58.
- Ian Roberts. 1993. *Verbs and diachronic syntax*. Studies in natural language and linguistic theory. Kluwer, Dordrecht.
- Beatrice Santorini. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory*, 10:595–640.
- Beatrice Santorini. 1993. The rate of phrase structure change in the history of Yiddish. *Language Variation and Change*, 5:257–283.

- Beatrice Santorini. 1995. Two types of verb second in the history of Yiddish. In Adrian Battye and Ian Roberts, editors, *Clause Structure and Language Change*, pp. 53–79. Oxford University Press.
- Robert P. Stockwell. 1977. Motivations for exbraciation in Old English. In N.C. Li, editor, *Mechanisms of syntactic change*, pp. 291–314. University of Texas Press, Austin, TX.
- Ann Taylor. 1994. The change from SOV to SVO in Ancient Greek. *Language Variation and Change*, 6:1–37.
- Alessandra Tomaselli and Bonnie Schwartz. 1990. Analysing the acquisition stages of negation in l2 german: support for ug in adult sla. *Second Language Research*, 6:1–38.
- Rosemarie Tracy. 1986. The acquisition of case morphology in German. *Linguistics*, 24:47–78.
- Anthony Warner. 1993. *English auxiliaries, structure and history*. Cambridge University Press, Cambridge.
- A. V. Zakharova. 1958. Usvoenie doshkol'nikami padezhnykh form. *Doklady APN RSFSR*, 3.