

Language Acquisition in German and Phrase Structure Change in Yiddish

To Appear in *Parameter Theory and Dynamics of Change*, Oxford: OUP.

Joel C. Wallenberg
joel.wallenberg@gmail.com

September 22, 2011

Introduction

There is a long tradition stretching back into the 19th century of implicitly assuming a relationship between language change and child language acquisition in the notion of “reanalysis”¹. Recently, studies such as Yang (2000) have developed formal models of language acquisition and expanded them to model how new syntactic variants can arise among children and be maintained in adult speech communities, formalizing the notion of “grammar competition” Kroch (1989). However, there have been very few empirical studies of language acquisition that can be linked to specific, well-documented cases of grammatical change.

This chapter investigates the relationship between acquisition and change in a study of a major phrase structure change in the history of Yiddish, the change in the structure of TP from a German-like Tense-final grammar to its modern Tense-medial grammar (Santorini 1992, 1993). In particular, this study will ask: was the direction of this change predetermined? Additionally, this paper explores the question of exactly what parameter was changing when the position of the tensed verb changed in the history of Yiddish, and I will suggest that an antisymmetric approach to head-

¹I would like to particularly thank Charles Yang for helping me to think about change in terms of acquisition in general, as well as for a number of helpful discussions regarding this paper. Beatrice Santorini and Anthony Kroch were also particularly helpful in working out the ideas in this paper and in interpreting the Early Yiddish data. I would also like to thank all of the attendees of DIGS XI at Unicamp, and one anonymous reviewer, for many helpful questions and comments. All errors are, of course, my own.

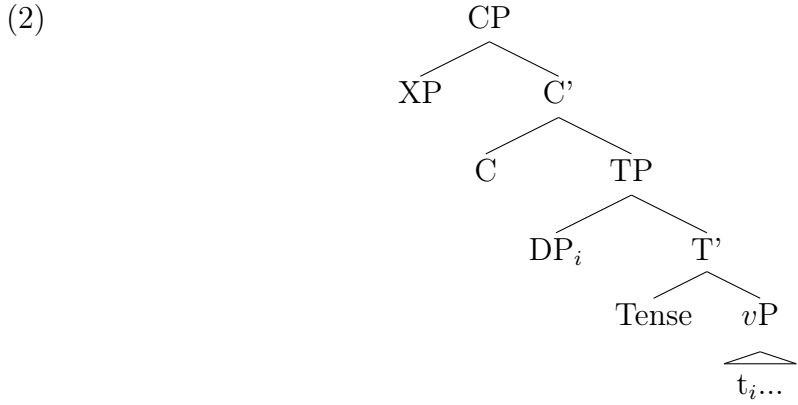
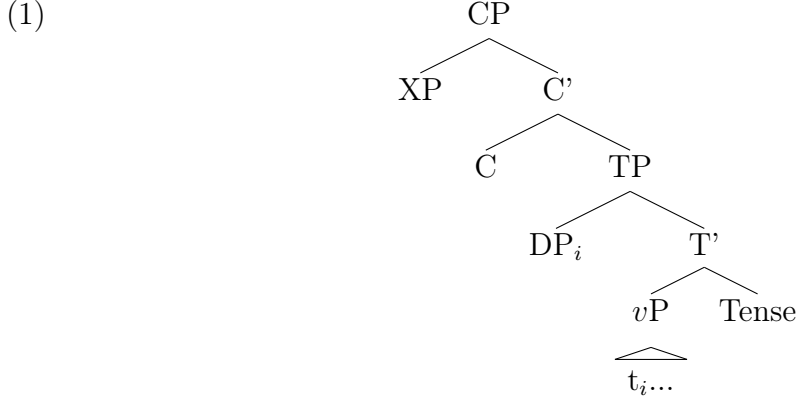
finality allows for a more precise understanding of how this historical change took place.

The rest of the chapter is organized as follows. In the section below I briefly summarize the historical change in Yiddish phrase structure that is under discussion, and frame the problem it poses for acquisition in terms of Yang (2000)'s model. In section 3, I show that Yang (2000)'s model for how children acquire syntactic variation in the input language cannot account for the data from Early Yiddish, at least in the most straightforward way. Section 4 discusses some data from childhood acquisition of modern German varieties which suggests that the previously assumed division of the Early Yiddish data into Tense-final and Tense-medial data may be flawed. Section 5 suggests a different understanding of the Early Yiddish data based on an antisymmetric account of the Tense-final grammar and the West Germanic Verb (Projection) Raising construction, and shows that the observed historical change is now predicted under Yang (2000)'s model. Finally, in section 6 I offer some conclusions and directions for further study.

1 Yiddish and Yang's acquisition model

As Santorini (1992) and Santorini (1993) showed, Yiddish gradually changed from a Tense-final language, like modern German or Dutch, into a Tense-medial language (or, a left-headed TP language, under classical X-bar theoretic assumptions) roughly between the years 1400-1800. After the change was initiated, Tense-medial TPs were introduced into the Yiddish speech community, and a period of variation began during which there was a mixture of both phrase structures in the speech community. That is simply another way of saying that while the change was underway, there was a state of "grammar competition" among Yiddish speakers (in the sense of Kroch 1989 and much subsequent work): there were both Tense-final and Tense-medial TPs evident in the performance of the community and produced by individual speakers, who could alternate between the two structures even from sentence to sentence within the same text (see Santorini 1992, where this fact is established beyond doubt).

Ultimately, the frequency of Tense-medial TPs advanced at the expense of the older Tense-final system, until this natural evolutionary process resulted in the uniformly Tense-medial modern Yiddish. Under a classical X-bar phrase structure (where headedness is a matter of linearization in accordance with a "head parameter" setting), this would mean that Yiddish changed from 1 to 2 below. Note that I will be assuming this classical, non-Kaynian (Kayne 1994, *inter alia*) view of head-initial and head-final phrase structure for the first three sections of this paper; this view will be revised in section 5.



The change in the position of Yiddish Tense is a perfect test case for Yang (2000)'s model of syntactic acquisition: the change is as simple as a syntactic change can be, as it involves only one parameter, and the state of variation is well-documented throughout almost the entire time course of the change (i.e. nearly all but the actuation of the change is attested). Once a new variant has been introduced into the linguistic environment of learners, the model in Yang (2000) is intended to evaluate the “fitness” (in an evolutionary sense) of the two variants and predict whether the new variant will deterministically win out over the older variant (given enough time, and holding other environmental factors equal). Yang adapts the well-known learning model from Bush and Mosteller (1951) and Bush and Mosteller (1958) to syntactic acquisition in a situation of grammar competition, and evaluates the evolutionary fitness of the competing variants in terms of the amount of evidence each grammar makes available to the learner for acquiring it. A brief description of Yang's syntactic learning model for the case of grammar competition is as follows (for details and the relevant mathematical proofs, see the full description in Yang 2000).

First, given a mixture of 2 grammars in the input to the learner, G_1 and G_2 ,

a child is expected to learn both grammars; this is simply another way of defining “grammar competition”, and it is the situation that santorini1992 observed in the intraspeaker variation of Yiddish speakers during the period of variation (see also Kroch 1989; Pintzuk 1991; Kroch 1994, and many studies building on those foundational studies). In acquiring the two grammars, the hypothetical child assigns some probability (weight) to each, and then continue to update these weights dynamically throughout the learning process, depending on the input data the child is exposed to in the speech community. Both grammars G_1 and G_2 generate some sentences that unambiguously identify them to the learner; if they did not, then the learner would not be able to know in the first place that both G_1 and G_2 are present. However, the two grammars most likely also generate some ambiguous sentences, sentences which could be the product of either grammar; e.g. an OV grammar and a VO grammar will both generate some short intransitive sentences of the form “The dog barked”, and these sentences will be string-wise identical no matter which of the two grammars generated them.

Yang's model states that when the child hears a sentence, she picks a grammar to analyze the sentence, choosing blindly based only on the preexisting weights associated with each grammar (i.e. the weights based on sentences the child heard prior to the current one). If the child hears an unambiguous sentence, e.g. only G_1 could have produced the sentence, then if the child picked G_1 beforehand and used it to try to analyze the sentence, G_1 will be rewarded; otherwise, G_2 will be punished and G_1 will be indirectly rewarded. Either way, G_1 ends up with an augmented weight. However, if the child encounters an ambiguous input, i.e. either G_1 or G_2 , can analyze the string the child hears, then the child will reward whichever grammar she happened to be using at the time. Ultimately, as the task is iterated many times, the child will end up assigning a higher probability at the end of the learning to the grammar that was most successful in analyzing unambiguous inputs. And, of course, this process can be iterated over a number of generations of learners as well, so the winner of a diachronic competition between two grammars over a long period of history will also be the one that can analyze the most unambiguous sentences in the input to learning.

This is the same thing as saying that the most successful grammar is the one which generates the highest frequency of unambiguous outputs. When the learning process is iterated over a number of generations, the first generation of learners becomes the second generation's parents (or adult speech community, more generally) and determines the composition of the linguistic input to the second generation's learning. The grammar which can analyze the most unambiguous inputs, from the learner's perspective, is by definition also the one producing the most unambiguous

outputs, from the adult speaker's perspective. As the weights are updated over and over again by the learner and, by extension, generations of learners using G_1 and G_2 alternately to analyze the outputs of G_1 and G_2 in the ambient linguistic environment, the grammar which produces more unambiguous sentences of its own type will have its weight augmented more often. Yang (2000) also shows mathematically that the proportion of unambiguous sentences a grammar generates is decisive even independently of the initial weights of G_1 and G_2 and the initial frequencies of G_1 and G_2 in the linguistic environment when the learning process begins. Thus, even if a given grammar begins as an extreme minority variant, if it is detectable to the learner at all, it will eventually win out over the majority variant if it generates/analyzes a higher proportion of unambiguous sentences of its own type than its competitor does.

In selectional, evolutionary terms, each grammar has a "fitness" which determines how likely it is to "reproduce" itself in a given learner's probability weights and in the acquisition process of future generations of learners:

$$\text{Fitness}(G) = \frac{\text{proportion of unambiguously "G" clauses it generates}}{\text{the clauses it generates.}}$$

If a grammar, G_1 , has a higher fitness than another grammar, G_2 , i.e. it generates more unambiguous clauses which signal, "I'm a G_1 clause", then G_1 has an "advantage" over G_2 :

$$\text{Advantage}(G_1 \text{ over } G_2) = \text{Fitness}(G_1) - \text{Fitness}(G_2)$$

Yang argues that if $\text{Fitness}(G_1) > \text{Fitness}(G_2)$, then G_1 must win in the long run (and vice-versa). Thus the outcome of any syntactic change is entirely fixed, once the change begins.

The goal of the remainder of the paper is to test the model in Yang (2000) against the empirical facts of the Yiddish change in the position of Tense, and make sense of the results of this experiment. We already know the result of the change: the Tense-medial (left-headed Tense) grammar won out and became the modern language. So, reasoning backwards, Yang's model hypothesizes that the Tense-final grammar was less fit than the Tense-medial grammar; the Tense-medial grammar should generate a higher proportion of unambiguously Tense-medial sentences than the Tense-final grammar generates of unambiguously Tense-final sentences. As we will see in the next section, under standard definitions of the two grammars, this prediction is not borne out.

2 A Hypothetical Early Yiddish Learner

We can test the hypothesis from Yang’s learning model using a parsed diachronic corpus of Yiddish (the Penn Yiddish Corpus, Santorini 1997/2008) to estimate the fitness of the Yiddish Tense-final and Tense-medial grammars, which were in competition during the period of phrase structure variation in the 15th through 18th centuries. Note that all of the data below is taken solely from subordinate clauses, as in Santorini (1992) and Santorini (1993), since even Tense-final Yiddish matrix clauses were V2, exhibiting general tensed-verb movement to C as in German (following the standard analysis of West Germanic V2 going back to den Besten 1983). Thus, all matrix clauses in Yiddish are ambiguous between Tense-final and Tense-medial phrase structure, from the point of view of both the analyst and the learner.

In order to compare the relative fitnesses of the two grammars, it is necessary to estimate the frequency of unambiguous (subordinate) clauses each grammar generates from some kind of representative sample of each grammar’s performance. For the Tense-medial grammar, the natural choice is a sample of text from some time period after the Tense-final-to-Tense-medial change had gone to completion. Any Yiddish text from the 19th century or later is essentially uniformly Tense-medial, and so the frequencies of unambiguous vs. ambiguous clause-types generated by the Tense-medial grammar were estimated from the texts in the *Penn Yiddish Corpus* from the years 1848-1947. (Note also that I have verified that no unambiguously Tense-final subordinate clauses occur in this sample.)

For the Tense-final grammar, I adopted the same method, taking as my sample the West Yiddish texts in the Penn Yiddish corpus from the year 1507 or earlier. There were no unambiguously Tense-medial subordinate clauses in this group of texts (which form the oldest time period of the most conservative dialect region in the corpus), and so this sample is as close to pure Tense-final Yiddish as one is likely to find.² The majority of relevant clauses in this period come from one early Western Yiddish text, Elia ha-Levi ben Asher Ashkenazi’s (also called Elia Levita or Elia Bachur) *Bovo Bukh* (Joffe 1949).

Having chosen the two sample corpora to represent the Tense-final and Tense-medial Yiddish grammars, the unambiguous clauses of each type are those that

²It is not possible to be completely certain that this sample is 100% Tense-final. However, any error in this regard would only artificially inflate the estimated frequency of ambiguous clauses generated by the Tense-final grammar (because some truly Tense-medial clauses would be incorrectly counted as ambiguous Tense-final clauses). However, this potential error is not problematic to the argument or results presented in the rest of the paper: as you will see below, it would only strengthen the main result of the paper if we corrected the estimated frequency of ambiguous clauses downward for the Tense-final grammar.

contain certain diagnostic elements, elements which can unambiguously diagnose a clause's underlying structure. Following Santorini (1992) and Santorini (1993), these diagnostic elements include sentential negation and verbal particles (i.e. also known as Germanic separable prefixes) including *loshen koydesh* elements³, both of which only occur preceding a finite verb in Tense-final clauses and following a finite verb in Tense-medial ones (See also, for the use of diagnostic elements in analyzing Old and Middle English: Pintzuk 1991; Kroch and Taylor 2000; Pintzuk 2005; Pintzuk and Taylor 2006; Pintzuk and Haeberli 2006). Pronominal objects, which do not participate in extraposition constructions due to their light prosodic status, are diagnostic of a Tense-medial clause when they occur following a finite verb (see references above, *inter alia*). Similarly, Wallenberg (2008) and Wallenberg (2009) showed that no objects scramble leftward across the finite verb in Yiddish (and in other languages; see Wallenberg 2009 for details), and so all objects preceding the finite verb are diagnostic of Tense-final structure. These diagnostics are illustrated in the examples below:

Particle, object, and negation diagnosing Tense-final:

- (3) zeyt gibetin d[a]z **mir** eyer fatr **moykhl** iz
 be prayed that me-DAT your father forgiving is
 'Hope/ask that your father is forgiving me'
 (letter in Weinryb 1937, date: 1588)
- (4) ...das ikh **im ab** zag
 that I **him off** spoke
 'that I refused him'
 (G"otz fun Fiderholtz's Complaint, date: 1518)
- (5) vau keyn fleysh nakh keyn blut **nit** iz
 What no flesh nor no blood **not** is
 'What neither flesh nor blood is...'
 (Preface to *Lev Tov*, date: 1620)

In 3 and 4 above, the the particles *moykhl* and *ab* precede the tensed verb, indicating that that the verb has not moved leftward to a medial Tense head in a left-

³Certain elements that were borrowed from Hebrew and form predicates when combined with German light verbs, such as *khason* *hobn* (= "to get married", lit. "marriage have") or *moykhel zeyn* (= "to forgive", lit. "forgive be"). These have the syntax of native Germanic particles in Yiddish; see Santorini (1989), Santorini (1992)

headed TP structure, and so these clauses are Tense-final. Similarly, the preverbal positions of negation in 5 and of the objects in 3 and 4 also show that Tense appears at the right edge of the structure in these examples.

In contrast, the examples below are Tense-medial clauses, showing particles, negation, and a pronominal object following the finite verb. In examples 6-9, the position of each of the boldface diagnostic elements shows that the finite verb has moved leftward past it to a medial Tense head.

Particle, negation, and pronominal object diagnosing Tense-medial:

- (6) un dernokh hot zi im gefregt, tsi er hot **lib** tsimes
 and afterwards has she him asked Q he has love tsimes.
 ‘Afterwards, she asked him if he likes tsimes.’
 (Olsvanger 1947, *Royte Pomerantsen*, date: 1947)
- (7) ...vi me ruft **oyf** dem rov tsu der toyre
 how one calls up the rabbi to the Torah
 (Olsvanger 1947, *Royte Pomerantsen*, date: 1947)
- (8) az men hot **ihm** friher keyn mahl **nist** dem emes gzagt
 that one has him earlier no time not the truth said
 ‘that he hadn’t earlier been told the truth even one time’
 (Vos iz dos azuns geshehn in Vien un in Lemberg?,
 revolutionary proclamation by Judah ben Abraham, date: 1848)
- (9) der tate hot im gefregt, vi azoy men ruft **im**
 The father has him asked how one calls him
 ‘The father asked him what his name was.’
 (*Grine Felder*, date: 1910)

Finally, in clauses containing a finite auxiliary and a nonfinite verb, the order nonfinite-verb > finite-auxiliary is diagnostic of Tense-final structure, as that order only occurs in undisputed Tense-final languages (e.g. modern German, Dutch).⁴ 10

⁴With the exception of Stylistic Fronting in languages like Icelandic. I do not pursue this point further here, since there is no clear evidence for SF in the history of Yiddish and V > FINITE-AUX orders do not persist into the modern Yiddish period. However, it is worth noting that SF may in fact arise from a change from Tense-final to Tense-medial in the modern languages that exhibit it, which have arguably undergone such a change.

and 11 both contain the Tense-final diagnostic of a nonfinite verb preceding the finite auxiliary, and in addition, 10 also shows pre-finite-verb negation and a preceding object, and 11 shows the additional diagnostic of the pre-tensed verb particle, *oyz* (all diagnostics are in boldface below).

V > Aux order (among other things) diagnosing Tense-final:

- (10) ...d[a]z **mir** yusf di h' zhubim **nit gebn vil**
 that me Joseph the five guilders not give wants
 'that Joseph doesn't want to give me the five guilders'
 (court testimony, date: 1465; also cited in Santorini 1992)
- (11) d[a]z es unzr her gut **oyz ginumn hut** far an
 that it our lord good out took has presently
 '...that our good Lord has made a success of it presently'
 (Leib bar Moses Melir's *Book of Esther*, date: 1589)

In addition to subordinate clauses with the above diagnostic elements, here is a great deal of ambiguous data in the Early Yiddish written record, and so presumably, also in the input to the hypothetical learner of Early Yiddish. In fact, the majority of clauses in the corpus are ambiguous with respect to the Tense-medial/Tense-final parameter. Some of these are ambiguous because there is simply not enough material available in the clause to diagnose the underlying structure, such as intransitive clauses containing only a Subject and finite lexical verb. In other cases, clauses with more material are nonetheless amenable to more than one analysis in terms of underlying TP structure because of the existence of other syntactic operations/constructions which are known to habitually obscure the underlying position of Tense. Following the numerous studies cited above with respect to the set of diagnostic elements in Germanic, the two most important of these obscuring constructions in studying the Tense-final to Tense-medial change are the West Germanic Verb (Projection) Raising Construction (VPR), as found in modern Dutch, West Flemish, and South/Swiss German (see Wurmbrand 2004, for an overview of the phenomenon and relevant literature) shown in 12, and various types of extraposition processes (e.g. PP extraposition as in 13 and DP extraposition or Heavy NP Shift as in 14). Note that the boldface diagnostic elements in the three sentences below demonstrate conclusively that these constructions occur in Tense-final Early Yiddish clauses.

- (12) dz **es** di mtsreym **nit** zaltn zehn
 that it the Egyptians not should see.

‘...that the Egyptians shouldn’t see it’
(Leib bar Moses Melir’s *Book of Esther*, date: 1589)

- (13) d[a]z ikh **reyn** verde fun der ashin
that I clean become from the ashes
‘that I become clean from the ashes’
(Johann Jakob Christian’s *Eyn sheyn purim shpil*, 1004, date: 1697
also cited in Santorini 1992: 607)
- (14) ven er **nit** veys eyn guti veyd
if he not knows a good pasture
‘if he doesn’t know a good pasture’
(Abraham Apotheker Ashkenazi’s *Sam Hayyim*, 41, date: 1590
also cited in Santorini 1992: 607)

In this way, many very frequent strings in the corpus provide no evidence to the learner as to whether the target language is Tense-medial or Tense-final; e.g., the configuration SUBJECT > FINITE-LEXICAL-VERB > OBJECT could represent a Tense-medial clause or a Tense-final clause with extraposition. Or, if a situation of competing grammars obtains in the input to the learner, these strings do not help the learner determine what the target frequencies of the two grammars are. Two examples of ambiguous subordinate clauses appear below. Note the difference in the dates of the two examples: these strings are easily generated by either the Tense-final or the Tense-medial grammar.

- (15) vu er vust di altn hern
how he knew the old lords/knights.
(Elia Levita’s *Bovo Bukh*, date: 1507)
- (16) az der rov hot a toes
that the rabbi has a mistake
‘that the Rabbi made a mistake’
(*Royte Pomerantsen*, date: 1947)

The table below shows counts and frequencies of ambiguous vs. unambiguous strings for both the Tense-final and Tense-medial grammars, as estimated by the sample corpora of pre-1507 Western Yiddish and post-1848 (Eastern) Yiddish, respectively. Clauses containing only a single finite verb are termed “Simplex”, and

clauses containing a finite auxiliary and at least one nonfinite verb are labelled “Complex”.⁵

Table 3.1: Tense-final and Tense-medial Sample Corpora

The comparison of the two sample corpora in the table above and the chart below pose a clear problem for our understanding of the phrase structure change in Early Yiddish. (Chart: dark-blue = Simplex, light-blue = Complex)

Figure 3.1

According to the sample corpora, the Tense-final grammar produces a higher frequency of unambiguously Tense-final clauses than the Tense-medial grammar produces of unambiguously Tense-medial clauses. $\text{Fitness}(\text{Tense-Final}) = .414$ and $\text{Fitness}(\text{Tense-Medial}) = .337$, giving the Tense-final grammar an advantage of .077 over the Tense-medial grammar according to the model from Yang (2000). This result predicts that the Tense-final grammar should have won out over the Tense-medial grammar over time, and that the change from Tense-final to Tense-medial should never have gone to completion. Indeed, if Tense-medial had come to occur at an appreciable frequency in the Yiddish speech community for some non-obvious reason, then the change should have reversed after that point.

Of course, the historical record and modern Yiddish show that the change did indeed go to completion in favor of the Tense-medial grammar, and so we are left with a puzzle that has one of 3 possible solutions:

1. Yang’s (2000) model of acquisition and change is wrong.
2. There is some inherent (UG or processing) bias in favor of Tense-Medial (or left-headedness generally), which causes learners to reward that grammar more when it analyzes a sentence.
3. I have not grouped the data in a way that accurately represents how the competition was actually perceived by the learner.

The first possibility would reject a model which is simple an adaptation of a very simple and well-established model of statistical learning. The second possibility is

⁵The “CP-recursion” row refers to clear embedded V-to-C movement, which will always be ambiguous between Tense-final and Tense-medial. The cases counted in this row are embedded contexts in which some non-subject XP has been topicalized and the finite verb has inverted with a subject pronoun. Due to a minor oversight the query counting these did not separate the Simplex and Complex cases.

a serious one, but it is not clear how to explore it in the context of the present study. Therefore, I will adopt the third hypothesis for the time-being as it is the most restrictive one, and argue that it leads to a reasonable analysis of the historical Yiddish case as well as a productive line of research in general.

3 The Acquisition of Modern German

Two observations concerning the childhood acquisition of modern German, a uniformly Tense-final target grammar, provide a clue to the resolution of the paradox with which we ended the previous section. First, the data show that the modern Yiddish Tense-medial grammar is a natural innovation for acquirers of Tense-final, even when those acquirers receive no direct evidence for a Tense-medial grammar. And secondly, learners of Tense-final varieties of German which also show the West Germanic Verb (Projection) Raising (VPR) construction (which in fact is all varieties of German, but to varying degrees) do not straightforwardly interpret this construction as an instance of the target Tense-final grammar. Instead, there is evidence to suggest that they relate them to a Tense-medial hypothesis.

Fritzenschaft et al. (1990) and Gawlitzek-Maiwald (1997) show that children acquiring German make a variety of errors in both matrix and subordinate clause syntax on the way to their ultimately successful acquisition of the target grammar. In particular, they record acquisition errors in which the children produce what appear to be Tense-medial subordinate clauses; they produce subordinate clauses containing unambiguous diagnostics for Tense-medial of the type that occur in the modern Yiddish Tense-medial grammar, and were not available in Yiddish before that grammar was innovated. The sentences shown below were uttered by children acquiring German monolingually from two German-speaking parents.

- (17) *dass du hast **net** die meerjungfrau*
 that you have not the mermaid
 ‘...that you don’t have the mermaid’
 (Benny, 3 years 1 month old, Fritzenschaft et al. 1990, 76)

- (18) *wenn des dreht **sich** was tut ’s dann*
 if it turns REFL what does it then
 ‘if it turns, then what does it do’
 (Benny, 3 years, 2 months, and 26 days old, Gawlitzek-Maiwald 1997, 137)

In 17 above, the finite verb has moved leftward across the negation *net*, leaving it to the right in the modern Yiddish order. Similarly, the weak pronoun *sich* appears

to the right of the finite verb in 18, which is another of the Tense-medial diagnostics I enumerated in the section above. Note that the overt complementizer *dass* in 17 and the overt subordinator *wenn* in 18 show that these contexts are clearly embedded and make it less likely that the verb movement in these clauses is due to embedded V-to-C movement (i.e. CP-recursion).

Penner (1990) records similar examples from his study of the acquisition of Bernese Swiss German, such as 19 and 20 below.

- (19) *nei eine, wo tuet **nid** hokche*
 no someone who does not sit
 (S., 3 years old, Penner 1990, 178)

- (20) *we me tuet 's lige la*
 when one does it lie let
 '...when one lets it lie'
 (M., 3 years old, Penner 1990, 178)

Like the sentences from the child Benny above, these two subordinate clauses from two different acquirers of Bernese Swiss contain the unambiguous diagnostics for Tense-medial of post-tensed-verb negation and weak object pronoun. Additionally, 19 is the clearest example of Tense-medial of all since relative clauses are a well-established non-CP-recursion environment across Germanic, and so the tensed verb must have moved to a lower head than C here (Iatridou and Kroch 1992).

In addition, Penner (1990) makes the second observation which relates directly to the study of Early Yiddish: he finds that learners of Swiss German are slower to acquire the target syntax of their variety than are learners of standard German, and he suggests that this is due to the frequent use of the VPR construction in adult Bernese Swiss German, which obscures the Tense-final grammar for the children. In fact, it may be no coincidence that most of the Tense-medial subordinate clauses cited in Fritzenschaft et al. (1990) and Gawlitzek-Maiwald (1997) come from the child Benny, who has a Swabian mother, and so is likely to have been exposed to the VPR orders that are common throughout the Southern varieties of German (including Swiss varieties). The idea that VPR slows down the acquisition of Tense-final, coupled with the fact that children exposed to VPR sometimes innovate a Tense-medial grammar, suggests that children do not interpret VPR as a mere variant of Tense-final. They may be positing a relationship between VPR and Tense-medial which has the following result: if the children happen to entertain the hypothesis that they should acquire a Tense-medial grammar, they do not consider VPR clauses in the

input as counting against this hypothesis. Or, possibly, these clauses even reinforce the Tense-medial hypothesis. In either case, the presence of VPR in the input would slow the acquisition of Tense-final, as Penner found, because it would take longer for the children to reject their Tense-medial hypothesis under either scenario.

The results from German acquisition prompt a possible answer to the paradox of the Yiddish historical development. If children acquiring German do not straightforwardly count VPR clauses as Tense-final, then the hypothetical acquirer of early Yiddish would also have difficulty doing so. This is especially plausible given that the acquirers of Early Yiddish were also confronted with actual positive evidence for a Tense-medial grammar, unlike the acquirers of modern German: the goal of the German and Swiss children is to reject the Tense-medial hypothesis, while the Yiddish speaking children had the more subtle task of correctly acquiring the frequencies of two competing grammars, once the Tense-medial grammar had attained any appreciable frequency in the population.

4 A Kaynian Solution to the Puzzle

The solution to the problem posed at the end of section 3 above lies in how we divide the evidence to the hypothetical learner of Early Yiddish. Yang's model predicts the outcome of Yiddish becoming a uniformly Tense-medial language to be impossible according to the division of the input I presented in section 3. If we have reason to believe that Yang's model is still correct, then the division of the data in section 3 must have been incorrect in some way. In particular, the data from the childhood acquisition of modern German suggest that we divided the input to the hypothetical Early Yiddish child wrongly with respect to the West Germanic Verb (Projection) Raising construction.

If studies such as Biberauer (2003) and Wallenberg (2009) are on the right track in their analysis of VPR, then the effect of this construction on the childhood acquisition of Tense-final Swiss German may go beyond the mere surface similarity between VPR and Tense-medial languages. Penner (1990) only went so far as to suggest that the seemingly Tense-initial verb-clusters which adult VPR creates in the input to children represent a complication in the Bernese Swiss German system, and so slow down the acquisition of the target Tense-final grammar. VPR clauses distract from the evidence for the Tense-final system, and have to be learned as an additional derivation in the grammar. However, in an antisymmetric approach to head-finality in the tradition of kayne1994, head-final orders are derived from underlyingly head-initial ones, and so by hypothesis, the AUX > VERB order characteristic of VPR represents the underlying left-headed order of verbal heads in the structure, just as

it does in a Tense-medial language, such as modern English, Icelandic, or modern Yiddish.

Following the approach in Biberauer (2003) and Biberauer and Roberts (2005), the difference between Tense-final and Tense-medial languages is not the underlying order of heads, but rather whether or not the vP (or the relevant projection containing the lexical verb) is pied-piped along with the subject as it moves to fill Spec(TP), leaving the tensed verb in final position. Thus, the derivation of the Early Yiddish Tense-final clause in 10 above is as shown in 21 below:

$$(21) \quad d[a]z \left[{}_{vP} \text{ mir yusf di h' zhubim nit gebn} \right] \text{ vil } t_{{}_{vP}}$$

Wallenberg (2009: 6.3) expands on this analysis and proposes that the VPR construction in Tense-final languages is the result of a similar pied-piping of a vP to Spec(TP), but a smaller, remnant vP which the lexical verb has left by head-movement. In this way, the arguments of the lexical verb and any low- vP -attached adjuncts are deposited in a position preceding the finite verb, but leaving the lexical verb in clause-final position, as shown in the following structure for example 12 above (see Wallenberg 2009 for details and extensive discussion):

$$(22) \quad dz \left[{}_{vP} \text{ es di mtsreym nit} \right] \text{ zalt } t_{{}_{vP}} \text{ zehn } t_{{}_{vP}}$$

Thus, the antisymmetric approach analyzes the VPR construction in Tense-final Early Yiddish and the modern Yiddish Tense-medial system as being quite similar. The Tense-final VPR construction is distinguished by one additional layer of structure being pied-piped to Spec(TP) with the subject, but both Tense-final VPR and the modern Yiddish Tense-medial system display the underlying order of verbal heads in the surface syntax.

Figure 3.2

Returning to the data from our sample corpora, the chart below is identical to the chart above in section 3, except that the gray area shows how much of the Tense-final grammar's unambiguous data is composed of subordinate clauses showing the VPR construction, i.e., clauses with a finite auxiliary and a nonfinite verb, in which the verbs are in the VPR (AUX > V) order.

Now let us suppose that we have been incorrect in tallying these clauses up on the side of unambiguous evidence to the learning in favor of the Tense-final grammar. Of course, they are clearly not Tense-medial either, as they all contain some other structure that is considered diagnostic of Tense-final, and so is not generated by the modern Yiddish Tense-medial grammar. However, following the antisymmetric approach described above, suppose that these clauses represent something of a mixed signal to a learner, as the results in Penner (1990) potentially suggest, particularly if the learner is faced with a situation of competing Tense-final and Tense-medial grammars. On the one hand, the clauses provide evidence that the underlying left-headed structure should appear on the surface, just as in the Tense-medial grammar. On the other hand, these clauses provide evidence that more structure should appear preceding the finite verb than the Tense-medial grammar allows, as in the Tense-final grammar.

Perhaps this type of mixed signal can be learned in the context of a homogeneous Tense-final target grammar, as in the Swiss German case, but it cannot be successfully used by children to decide in favor of the Tense-final grammar over a Tense-medial grammar if the target is a situation of competing grammars. The status of VPR examples as a mixed signal is only expected under the Kaynian approach, since VPR represents the underlying order of heads under that account. Under a traditional non-Kaynian analysis of VPR, VPR is a local permutation of an underlyingly right-headed structure, and VPR would only be learned as an additional variant of the Tense-final system. VPR would therefore be straightforwardly counted as Tense-final, as we did above.

If we remove the VPR sentences from consideration as suggested by the Kaynian approach, and hypothesize that learners (and generations of learners) do not use them to decide the outcome of the competition, then the revised relative fitness of the two grammars can be seen in the chart below.

Figure 3.3

The effect of removing the VPR clauses from consideration has changed Advantage ($G_{Tense-final}$ over $G_{Tense-medial}$) such that now the Tense-medial grammar is predicted to take the place of the Tense-final grammar in the population over the generations following its innovation. The conclusion is clear: treating VPR as a minor variant of Tense-final makes the wrong diachronic prediction for Yiddish, but applying an antisymmetric understanding of the VPR structure leads us to the correct prediction.

5 Directions for Further Research and Conclusions

In this paper I have presented the change in the position of Tense in Early Yiddish as a test case for the learning model in Yang (2000). This model has the property that change is always unidirectional, and in this case, it predicted the wrong direction of change according to the traditional understanding of the Tense-final grammar and the West Germanic Verb (Projection) Raising construction. This clear and precisely defined incorrect hypothesis led us to entertain a new hypothesis about how the learner must have interpreted the syntactic variation in Early Yiddish once the input became mixed. This new hypothesis was in line both with data from studies of the modern acquisition of German and a particular theoretical approach to the patterns found in the Tense-final grammar, an antisymmetric approach.

However, this analysis raises a number of questions because of the prominent place it gives the VPR construction in explaining the change in Yiddish Tense. For example, does a Tense-final language need to exhibit VPR in order to change to Tense-medial? Both Early Yiddish and Old English changed to Tense-medial and both showed high frequencies of VPR (for OE see Pintzuk and Haerberli 2006), but we need more detailed descriptions of this change in order to decide if VPR is always a necessary condition for the change from Tense-final to Tense-medial. If possible, investigating this change in North Germanic would be a potential next step in deciding this question.

This study has shown that the study of language change, language acquisition, and a careful understanding of the linguistic structures involved in variation can interact to inform each other in the development and testing of precise quantitative hypotheses. While there is clearly much more work to be done in relating the acquisition of syntactic variants to morphosyntactic change, it is my hope that the approach taken in this paper will be found useful in future quantitative studies of this relationship.

Table 3.1: Tense-final and Tense-medial Sample Corpora

Early West Yiddish (Tense-Final Sample)			
	Simplex	Complex	All
Ambiguous	50	42	92
CP-recursion: Ambiguous, Simplex and Complex counted together			34
Unambiguous	32	57	89
Frequency of Unambiguous Tense-Final	.390	.576	.414

Late East Yiddish (Tense-Medial Sample)			
	Simplex	Complex	All
Ambiguous	752	841	1593
CP-recursion: Ambiguous, Simplex and Complex counted together			65
Unambiguous	363	479	842
Frequency of Unambiguous Tense-Medial	.326	.363	.337

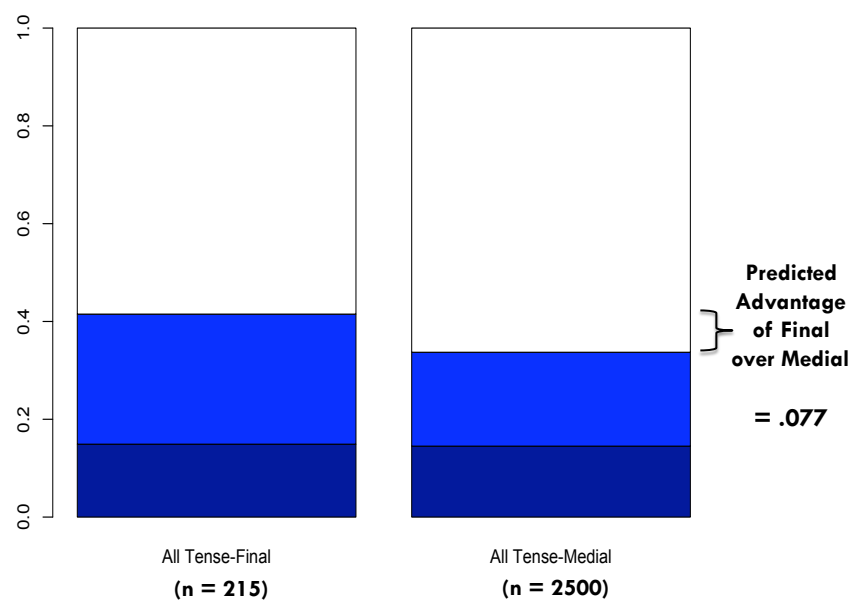


Figure 3.1

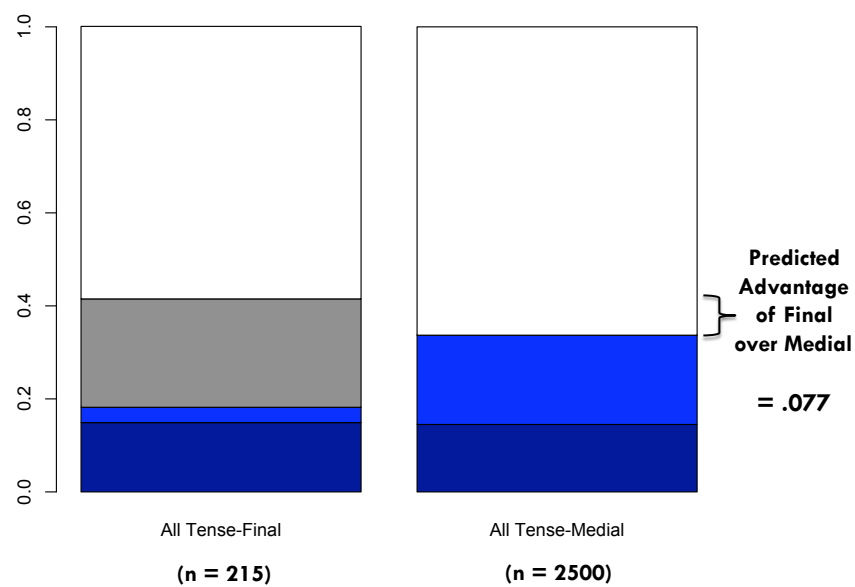


Figure 3.2

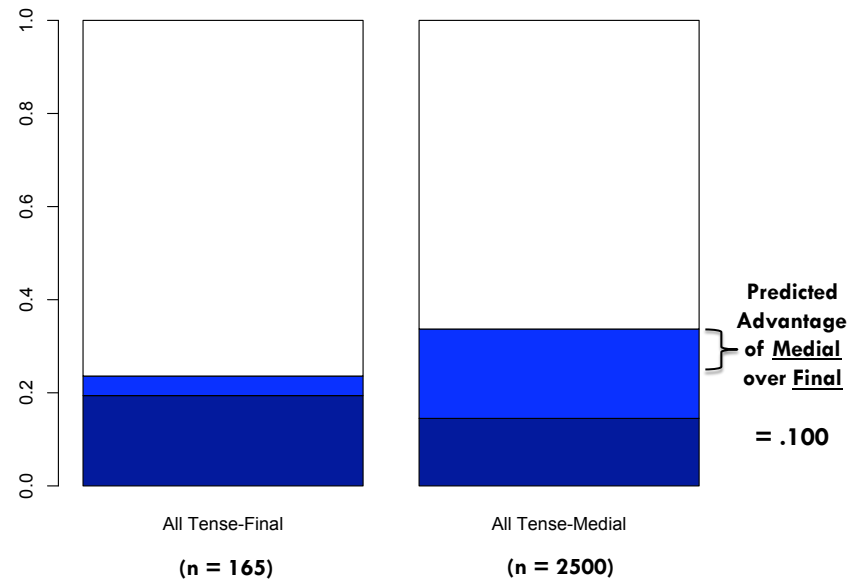


Figure 3.3

References

- den Besten, Hans. 1983. On the interaction of root transformations and lexical deletive rules. In *On the Formal Syntax of Westgermania*, ed. Werner Abraham, 47–131. Amsterdam: John Benjamins.
- Biberauer, Theresa. 2003. Verb second (V2) in Afrikaans: a minimalist investigation of word order variation. Doctoral Dissertation, University of Cambridge, Cambridge, England.
- Biberauer, Theresa, and Ian Roberts. 2005. Changing EPP-parameters in the history of english: accounting for variation and change. *English Language and Linguistics* 9:5–46.
- Bush, Robert, and Frederic Mosteller. 1951. A mathematical model for simple learning. *Psychological Review* 58:313–323.

- Bush, Robert, and Frederic Mosteller. 1958. *Stochastic models for learning*. New York: Wiley.
- Fritzenschaft, Agnes, Ira Gawlitzek-Maiwald, Rosmarie Tracy, and Susanne Winkler. 1990. Wege zur komplexen syntax. *Zeitschrift für Sprachwissenschaft* 9:52–134.
- Gawlitzek-Maiwald, Ira. 1997. *Der monolinguale und bilinguale erwerb von infinitivkonstruktionen: Ein vergleich von deutsch und englisch*. Tübingen: Max Niemeyer Verlag.
- Iatridou, Sabine, and Anthony S. Kroch. 1992. The Licensing of CP-recursion and its Relevance to the Germanic Verb-Second Phenomenon. *Working Papers in Scandinavian Syntax* 50:1–24.
- Joffe, Judah A. 1949. *Elia Bachur's poetical works. with philological commentary, grammar and dictionary of Old Yiddish. three volumes..* Isny im Allgäu: Judah A. Joffe Publication Committee.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.
- Kroch, Anthony S. 1994. Morphosyntactic variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K. Beals et al et al.
- Kroch, Anthony S., and Ann Taylor. 2000. Verb-complement order in Middle English. In *Diachronic syntax: Models and mechanisms*, ed. S. Pintzuk, G. Tsoulas, and A. Warner, 132–163. Oxford: Oxford University Press.
- Olsvanger, I., ed. 1947. *Royte pomerantsen*. Schocken Books.
- Penner, Zvi. 1990. On the Acquisition of Verb Placement and Verb Projection Raising in Bernese Swiss German. In *Spracherwerb und grammatik: linguistische untersuchungen zum erwerb von syntax und morphologie*, ed. M. Rothweiler, 166–189. Opladen: Westdeutscher Verlag.
- Pintzuk, Susan. 1991. Phrase structures in competition: Variation and change in Old English word order. Doctoral Dissertation, University of Pennsylvania.

- Pintzuk, Susan. 2005. Arguments against a universal base: evidence from old english. *English Language and Linguistics* 9:115–138.
- Pintzuk, Susan, and Eric Haeberli. 2006. Head-final structure in Old English root clauses. In *Diachronic Issues in Generative Syntax*, volume 9. University of Trieste.
- Pintzuk, Susan, and Ann Taylor. 2006. The loss of OV order in the history of English. In *Blackwell handbook of the history of English*, ed. A. van Kemenade and B. Los. Blackwell.
- Santorini, Beatrice. 1989. The generalization of the Verb-Second Constraint in the history of Yiddish. Doctoral Dissertation, University of Pennsylvania.
- Santorini, Beatrice. 1992. Variation and change in Yiddish subordinate clause word order. *Natural Language & Linguistic Theory* 10:595–640.
- Santorini, Beatrice. 1993. The rate of phrase structure change in the history of yiddish. *Language Variation and Change* 5:257–283.
- Santorini, Beatrice. 1997/2008. The Penn Yiddish Corpus. University of Pennsylvania.
For details, contact: beatrice@babel.ling.upenn.edu.
- Wallenberg, Joel C. 2008. Scrambling and phrase structure in early and modern Yiddish. Paper presented at the 10th Diachronic Generative Syntax (DIGS) Conference, August 8, 2008, Cornell University, Ithaca, NY.
- Wallenberg, Joel C. 2009. Antisymmetry and the Conservation of C-Command: scrambling and phrase structure in synchronic and diachronic perspective. Doctoral Dissertation, University of Pennsylvania.
- Weinryb, Beier. 1937. A pekl briv in yidish fun 1588, Krake-Prag (a collection of Yiddish letters from 1588, Cracow-Prague). In *Historishe shriftn fun yivo (yivo historical papers)*, ed. E. Tsherikover, volume 2. Vilna: YIVO.
- Wurmbrand, Susi. 2004. West Germanic verb clusters: the empirical domain. In *Verb clusters: A study of Hungarian, German, and Dutch*, ed. Katalin É. Kiss and Henk van Riemsdijk, 43–85. Amsterdam: John Benjamins.
- Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12:231–250.