

The Loss of Auxiliary Selection in English

Mimi Lipson

1 Introduction

At some point in its history, English ceased to be a language that, like most Germanic and Romance languages, uses the HAVE auxiliary with some verbs and the BE auxiliary with others to form the Perfect. This project is a corpus-based investigation of the decline of auxiliary selection in English. A large corpus of sentences was assembled from literary texts available on the Internet through Project Gutenberg¹, covering the years 1560–1875 (= year of author's birth). The results show that, during the long period of variable auxiliary selection, the choice of auxiliary reflected a previously un-encoded contrast in meaning.

Evidence for semantic structure is always indirect, being more a matter of what entailments go with a particular construction than what constructions are possible or grammatical in a language. This fact makes the diachronic study of semantics through corpora particularly difficult. In this case, both assembling a usable corpus and interpreting the pattern of variation required some basis for characterizing the relevant semantic qualities of the sentence-level tokens. As far as I know, projects such as these are uncommon; therefore, the second and third sections focus on the methodological issues that came up when assembling and analyzing the corpus.

A major theme of contemporary sociolinguistics is the linguistic interpretation of variation, and specifically, how the distribution of new and old forms can be used to answer questions about how innovative forms embed themselves in the grammar. Toward that end, I propose a linguistic trajectory of change for the present case in Section 4. The implications of this embedding mechanism will be discussed in Section 5, particularly with respect to the *competing grammars* model of linguistic change developed in the work of Kroch and others. Finally, there is a growing body of synchronic work on the variable distribution of HAVE and BE in the Germanic and Romance languages (Platzak 1987, Kayne 1993, Freeze 1992, Iatridou 1995), and in the last section I will say a few words about what I believe to be the relevance of this work to the diachronic situation in English. (Readers interested in a more formal semantic treatment of this particular case are referred to Lipson 1999).

¹ e-texts: <http://www.promo.net/pg/list.html>

2 Auxiliary Selection

In auxiliary selection languages, the potential BE-selecting verbs are unaccusatives—that is, verbs in which the single argument shows object-like properties, as in (1), as opposed to subject-like properties, as in (2).

- (1) a. Han *har/er rejst
 *he *has/is gone* (Danish)
- b. Si *ha/è arrivati
 *they *have/are arrived* (Italian)
- (2) a. Han har/*er sovet.
 *he has/*is slept*
- b. Si ha/*è mangiati
 *they have/*are eaten*

There are many internal lexical-semantic distinctions within the general class of unaccusative verbs (for a thorough discussion of unaccusative verb classes, see Levin and Rappaport-Hovav 1996), and auxiliary-selecting languages differ in how widespread BE-selection is among the unaccusative verb types.

3 The Gutenberg Corpus

3.1 Sampling Procedures

In order to conduct a constrained and systematic search of texts from various historical periods, I restricted my search to the class Levin and Rappaport-Hovav call ‘verbs of inherently directed motion,’ a class which includes verbs such as *come*, *go*, *fall*, *rise/arise*, *depart*, *arrive*, *land*, *return*, *gather*, and *meet*. It was necessary to maintain some degree of comparability between verbs, since we have no other assurance that the individual verbs within the class are not behaving differently (through some form of classificational drift or lexical diffusion). Of the verbs sampled, only *come*, *go*, and *fall*, and *become*² occur in sufficient numbers for cross-verb comparison, and so only these verbs were included in the corpus.

The object in assembling the corpus was to collect all tokens of Perfects within this verb class in order to study the decline over time of the use of BE + past participle as a Perfect. However, contemporary usage suggests that not

² Though *become* is transitive, it is universally classed as a BE verb in auxiliary-selecting languages.

all such strings are actually Perfects. In particular, consider the case with *gone*, which occurs in construction with both auxiliaries:

- (4) a. John has gone
b. John is gone

The alternation in (4) does not mean that there is (variable) contemporary BE selection when forming a Perfect of the verb *go*. There are three types of evidence against treating (4b) as an alternative Perfect construction. First and most obviously, there are many types of modification which, in contemporary English, are restricted in their use with *gone*, several of which are listed in (6); specifically, the construction does not combine comfortably with modification that delimits the path, goal, beginning time, or manner of going.

- (6) a. John had gone to the store
b. ?John was gone to the store (*goal*)
c. John had gone on the interstate
d. ?*John was gone on the interstate (*path*)
e. John had gone as soon as possible
f. ?John was gone at 12 sharp (*beginning time*)
g. John had gone effortlessly/by car
h. ?John was gone effortlessly/by car (*manner*)

Also, BE + *gone* and HAVE + *gone* have different entailments with respect to the current status of the underlying eventuality: (4b) entails that John is still gone, whereas (4a) carries no such entailment: John may have gone and then returned. In this sense, *gone* resembles past participles like *drunk* or *finished*:

- (7) a. John had finished (with) his homework
b. John was finished with his homework
(8) a. John had drunk
b. John was drunk

This distinction turns out to be of little help in the sampling process, since it is not usually possible to deduce these sorts of temporal entailments from a written text. However, I discuss below a way in which the sample seems to reflect the entailment facts of contemporary HAVE/BE alternations.

The most serious objection to analyzing (4b) as a BE-selecting Perfect is the fact that a Perfect may be formed from BE + *gone*:

- (9) John *has been gone* for ages.

There is evidence that this sort of stative adjective used to be more widespread within the verb class in question, for instance:

- (10) And with that Diabolus gave back, thinking that more aid *had been come*. (Bunyan)

For this reason, it would not have been sufficient to simply remove the *gone* tokens from the sample. I needed some way of ensuring that these stative adjective uses were not counted in the sample—i.e., that all the tokens included were unambiguous Perfects. Essentially, I used as a guide my own intuitions about what can currently modify BE + *gone* in contemporary speech. I included in my sample tokens occurring with any hint of agentivity—mainly purpose, as in (11), delimiting PPs such as those in (12), and certain kinds of adverbials, as in (13), as well as tokens coordinated with other verbal past participles, as in (14):

- (11) a. They are come to give us joy.
 b. I am come on purpose to quarrel with you
 c. She was come in herself for the Stilton cheese,
 (12) a. They were gone to Hartfield. (*goal*)
 b. I was really gone from Randalls. (*source*)
 c. Is Mr. Elton gone on foot? (*instrument*)
 d. I had not gone three steps... (*extent*)
 e. No sooner had I come to the point... (*result*)
 f. ...after so many hazards as I had gone through... (*path*)
 (13) She had come as fast as she could (*rate/pace*)
 (14) When they had been all walking together, he had so often come and walked by her, and talked so very delightfully!—

All tokens that did not have some type of disambiguating modification were excluded from the sample.

3.2 Sample Overview

In order to get an idea of when the HAVE Perfects first began to appear, I looked at the Penn-Helsinki corpus of Middle English, which is comprised of texts dated from AD 1150 to AD 1500. The entire corpus contained only fourteen examples of HAVE + *come*. Of these, twelve occurred in irrealis (i.e., non-factive) environments—either in counterfactuals, as in (15), or as complements of modal verbs, as in (16):

(15)...and *had I not that tyme comen* he sholde haue taken his lyf from hym...

(16)...for south ye *myght have comen* to my counter, ...

The Middle English corpus established two things: 1. that an adequate sample would have to cover a later period of the language, and 2. that irrealis uses were the first places where BE was replaced with HAVE in this verb class. Indeed, looking at the Early Modern English data, it appears that have was categorically used with modals and counterfactuals, except where be was used to denote futurity. Since they did not show variability, the modal and counterfactual uses were also eliminated from the sample.

Table 1 shows the token distribution of the four verbs in the whole sample as a whole, divided by the period during which the author was born. Table 2 shows the percentages of be Perfects for the four verbs, and for a composite class of the three transitive verbs (*come*, *go*, and *fall*).

When plotted graphically, the decline in the percent of BE Perfects in the sample follows an S-shaped curve³, as in Figure 1.

All Perfects	1560- 1575	1608- 1625	1660- 1675	1710- 1725	1760- 1775	1810- 1825	1860- 1875	Totals
<i>come</i>	39	96	30	19	54	144	23	405
<i>gone</i>	76	38	72	17	59	83	33	378
<i>fallen</i>	21	11	17	7	25	32	16	129
<i>become</i>	28	9	10	9	27	37	14	134
Total	164	154	129	52	165	296	86	1046

Table 1: Project Gutenberg texts (by author's year of birth)

BE Perfects	1560- 1575	1608- 1625	1660- 1675	1710- 1725	1760- 1775	1810- 1825	1860- 1875
<i>come</i>	32	93	20	5	19	26	0
<i>gone</i>	49	25	40	6	19	15	2
<i>fallen</i>	15	8	7	2	0	2	0
<i>become</i>	28	9	10	7	11	2	0

Table 2: BE Perfects in the Gutenberg corpus

³ The period of initial increase is probably due to stylistic differences between the earliest texts, which drew heavily from the morality plays of Bunyan, and the second-earliest category, in which the plays of Shakespeare are most heavily represented.

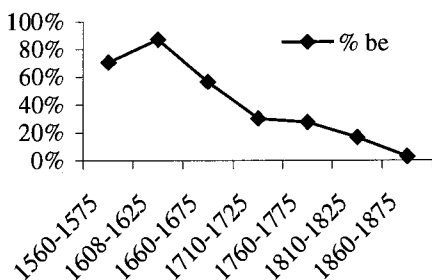


Figure 1: Rate of be-perfects with *come*, *go*, and *fall*.

This trajectory is a typical quantitative picture of the loss of an outgoing variant: in case after case of syntactic, morphological, or phonological change, the rate of change is slow at the beginning and end of the process, and slower in the middle. It is interesting to note that the pattern holds in a linguistic change which, as we will see, involves meaning differentiation.

When the verbs are plotted individually, as in Figure 2, it is apparent that they pattern as a class—with the exception of *become* which, being a transitive verb, is arguably not a real member of the class.

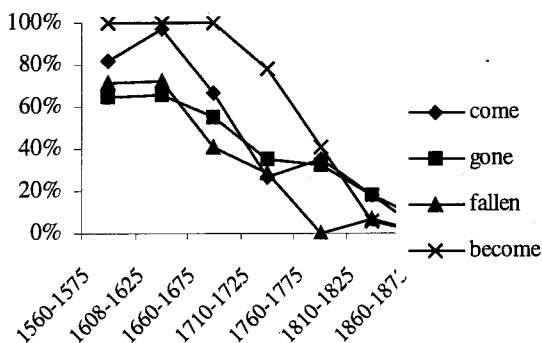


Figure 2: Rates of BE perfects for individual verbs

4 Semantic Differentiation of the Auxiliaries

The earliest data show that modal/counterfactual use of HAVE Perfects were the first to appear in this verb class. In other words, these irrealis contexts seem to have been the wedge that introduced the HAVE Perfect to this verb class. The question then becomes: how did the HAVE Perfect spread to non-irrealis contexts? As I worked with the Gutenberg texts, I kept noticing examples such as those in (17).

- (17) a. For *after* that the children of God had gone in vnto the daughters of men and had begotten them childern, the same childern were the mightiest of the world and men of renowne (Bunyan).
 b. Thus *twice before*, and jump at this dead hour, with martial stalk hath he gone by our watch (Shakespeare).
 c. For a mile or thereabouts my raft went very well, only that I found it drive a little distant from the place where I had landed *before* (Defoe).
 d. ...and my conscience, which was not *yet* come to the pitch of hardness to which it has since, reproached me with the contempt of advice, and the breach of my duty to God and my father (Defoe).
 e. Am I come near ye *now*? (Shakespeare)

I began to strongly suspect that, during the period when there was variation in the auxiliary used to form the Perfect in this verb class, the morphological variation was exploited semantically. Roughly, the two forms seemed to show a tendency toward the entailment difference pointed out above in the discussion of contemporary *have gone/be gone*: BE kept appearing in contexts where the eventuality persists, as in (17d and e), where the sentence is modified by *yet* and *now*. Conversely, HAVE seemed to appear with unusual frequency where the eventuality is more certain to have ended or taken place entirely in the past, as in (17a, b, c), where it coincides with the modifiers *after*, *twice*, and *before*.

4.1 Semantic Heterogeneity of the Perfect

The meaning distinction alluded to is roughly in line with the distinction between what are sometimes referred to as the Universal Perfect and the Existential Perfect (cf. Dowty 1979). In pseudo-Reichenbachian terms, where T=topic time (time during which the event or state referred to took place or held), R=reference time (e.g. now for present Perfects; some time in the past for past Perfects), and S=speech time (the time of utterance), the distinction is whether or not T and R are ever contemporaneous: the Universal Perfect means that the reference time is included in the interval during which the eventuality holds. The distinction is illustrated in (18,19), with the box notation indicating the duration of the topic time.

- (17) Mary had always lived there.

Universal:

R...	S
Mary lives there	

(18) Mary had read Valley of the Dolls.

Existential:

Mary reads

R...S

I use the terms Universal and Existential with caution, since the Universal Perfect is said to be available only with stative verbs (but see below for a sense in which the BE Perfect behaves like a Perfect formed with a stative verb); however, it does seem that during the period of auxiliary variation, the auxiliaries were (or at least tended to be) used to encode the distinction in terms of the topic time/ reference time relation. This is a difficult point to argue when dealing with historical texts, since the evidence for these sorts of temporal entailments is quite indirect. However, as suggested above, there are some differences in the way these two Perfects can be modified; for instance, the Universal Perfect cannot be modified by iterative adverbials, like *many times* or *twice*. Table 3 shows that HAVE is more prevalent than would be expected (given the overall distribution) with modification that contraindicates the Universal Perfect reading, whereas BE is more likely with *now*.

Year of birth:	1560- 1575	1608- 1625	1660- 1675	1710- 1725	1760- 1775	1810- 1825
Now + BE	20	22	13	7	17	4
Now + HAVE	0	1	0	2	2	10
Times* + BE	5	4	0	0	0	0
Times + HAVE	8	8	1	6	3	4

Table 3: Effects of temporal modification

* = X times, before, since, after

When plotted as percentages of Perfects selecting BE (Figure 3) and HAVE (Figure 4), the effect is quite clear: though the direction and general rate of change is consistent in every category, the overall rates are dramatically different.

The HAVE/BE distribution is reminiscent of the *be gone/have gone* alternation in another way: because only *be gone* is durative in nature, *have gone* cannot occur in a *while* clause:

(20) Mary read a book while John was/*had gone

The problem is not that Perfects cannot be temporal antecedents for *while*, but rather that the underlying eventuality must be stative, as in *know* and *have*:

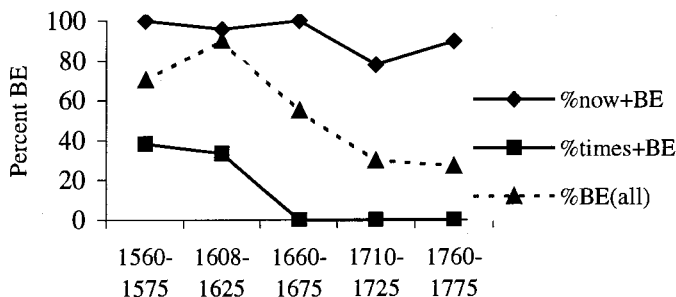


Figure 3: % BE with temporal modifiers

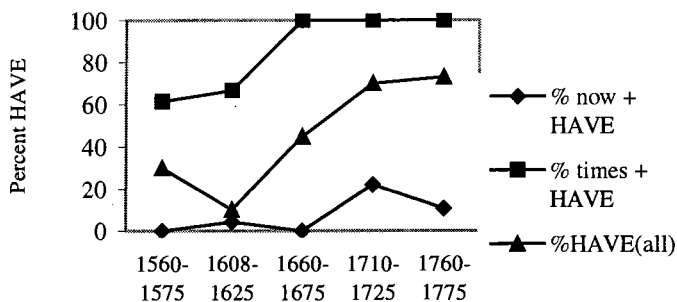


Figure 4: % HAVE with temporal modifiers

- (21) a. I've read *Valley of the Dolls* four times while I have lived here.
 b. John has read *Valley of the Dolls* four times while I have known him.

There are several examples of BE + past participle *while* clauses in the corpus, and some with modification that identifies the construction as a Perfect, for instance:

- (22) a. We found her a ship of Bristol, bound home from Barbadoes, but had been I own out of the road at Barbadoes a few days before she was ready to sail, by a terrible hurricane, *while the captain and chief mate were both gone on shore.* (Defoe)
 b. The dwarf, watching his opportunity, *while Glumdalclitch was gone to the side-board,* mounted the stool that she stood onto take care of me at meals, took me up in both hands, and squeezing my legs to-

gether, wedged them into the marrow bone above my waist, where I stuck for some time, and made a very ridiculous figure. (Swift)

Examples like these suggest that the class of verbs has been reclassified in terms of the possible aspectual interpretations they support.

4.2 A Semantic Trajectory

The question now becomes: how did the HAVE Perfect come to be used in non-irrealis contexts? It seems to me, intuitively, that it is not a coincidence that the HAVE Perfects, which started out in irrealis environments, next occurred preferentially in Existential Perfects. So, what do counterfactuals and modals have in common with the Existential Perfect sketched in (19) but not the Universal Perfect sketched in (18)? In both cases, the eventuality is *not* asserted to persist into the present. We will see if we can make some use of this fact.

Iatridou (1996) accounts for the contribution of past-tense morphology in the interpretational contrast between normal conditionals (23a) and counterfactuals (23b):

- (23)a. If you live here, you are home now.
- b. If you lived here, you would be home now.

The difference, essentially, is that (23b) is anti-factive: it cannot be felicitously uttered if in fact you *do* live here—whereas (23a) is factively neutral. Iatridou proposes that the link between the use of past morphology for *non-present* denotation and for *counterfactual* denotation is the relational quality of EXCLUSION (or, more accurately, non-inclusion). This relation can be asserted of time intervals (as in the simple past denotation) or of sets of worlds (the modal interpretation). In other words, the use of past tense in a conditional like (23b) triggers the inference that no world in which the proposition “you live here” is true is included in the set of worlds that are consistent with what we know to be consistent with the present one.

There is a similar sort of interpretational side-effect connecting irrealis contexts to the punctual, Existential interpretation of the Perfect (and to the interpretation of *have gone* as opposed to *be gone*): In the Existential interpretation, the set of time intervals in which the eventuality holds (i.e., the Topic Time) is not included in the time of reference; this is not the case with the Universal Perfect. In the case of the irrealis context, this relation holds between the possible worlds evoked by the modal/counterfactual and the actual world at the time of speech: *John might/could/should have gone* does not entail that John did in fact go.

The connection sketched above assumes that a relation can be extended from one semantic domain to another. In the case of counterfactuality and past marking, a morphological-semantic mapping applied in the unmarked case in temporal structures is superimposed upon a modal structure. The trajectory I am suggesting is an extension of the environments selecting the HAVE auxiliary from the modal domain to the aspectual domain. Note that this would not be unprecedented: there are some other examples of languages where the Perfect morphology is used to denote special epistemic positions, such as the Macedonian Dubitative Perfect (Friedman 1988) and the Bulgarian Perfect of Evidentiality (Izvorski 1997).

5 Competing Grammars?

There is a steady accumulation of research on the synchronic facts of auxiliary selection. Attempts to explain the distribution of auxiliaries, either within a language or crosslinguistically, have tended toward syntactic explanation—such as the accounts of Kayne (1993) and Freeze (1993), both of whom propose that the HAVE auxiliary is the spellout of a movement-driven incorporation of BE with an abstract pronoun-like feature, or Platzack (1987) who argues that auxiliary selection among the Scandinavian languages is an epiphenomenon of the null subject parameter. Whatever the merits of the syntactic approach to auxiliary selection, it is difficult to see how such analyses could explain the diachronic situation in English: the decline in auxiliary selection did not coincide with any change in pro-drop or word-order possibilities. The next obvious place to look for an explanation of the loss of this (putatively) syntactic feature is the work of historical syntacticians.

In his writing on historical syntax, Kroch (1994) invokes the Blocking Principle from morphology as a cause of change: *Stable linguistic systems do not permit equivalent morphological (or syntactic) doublets* (i.e., dived/dove). He argues that all morphosyntactic change indicates the existence of competing grammars, with each member of the doublet represented in one or the other grammar. The argument for competing grammars has been convincingly made in cases such as the evolution of periphrastic *do* in Middle English (Kroch 1989) or the change from V-to-C to V-to-I movement in Yiddish (Santorini 1993), and on the evidence of cases such as these, Kroch (1994) makes some explicit claims about the underlying mechanism of linguistic change: “*We have seen that the historic evolution of competing variants in syntactic change is similar to the evolution of morphological doublets. In both cases, the coexistence of variant forms is diachronically unstable: one form tends to drive the other out of use and so out of the language*” (Kroch 1994: 17). The old grammar is replaced by a new one at the

same rate in all environments, since each environment is representative of an underlying grammar which is the actual locus of change. In some cases, it can appear that certain environments "lead" in the change, since the input rate of application of the new form may be greater or less, but this is not to be taken as evidence that change progresses independently in individual linguistic environments. In other words, Kroch argues that it is *not* the case that forms in specific environments change by a process of analogy with other environments—i.e., "generalization" of the application of a new rule.

Since auxiliary selection has so often been analyzed as a syntactic phenomenon, we would expect that this sort of analysis, and specifically the invocation of the Blocking Principle, would apply to a case where auxiliary selection is lost in a language. However, note that the Blocking Principle, as stated, can be satisfied in more than one way: either one or the other morpheme may be lost *or* their status as doublets may change via a process of semantic differentiation. In the present case, it seems clear that the process involves more than a meaning-neutral transition from one morpheme (BE) to another (HAVE). Along the way, there seems to have been a language-internal redistribution of grammatical distinctions. Of course this begs the question: why *did* the BE Perfect go out of use? I don't have an answer for this, but possibly it was a problem of dual interpretations of past participles like *fallen* and *become*. In other words, perhaps having the resources to distinguish Existential Perfects from Universal Perfects in only one verb class was too cumbersome and asymmetrical. Note again that meaning one entails the other: *I am come* entails *I have come*. Therefore, any situation in which *I am come* is felicitously uttered also supports the statement *I have come*. Therefore, if one or the other interpretation had to survive, it makes sense that it would be the more broadly applicable.

6 Conclusion

There is an accumulating body of examples where morphosemantic change involves language-internal morphological shifts from one semantic domain to another. The existence of such a path of change presents a challenge to the universal applicability of a competing grammars analysis of language change and suggests that a general model of morphosemantic change will require a somewhat different outlook—one more sensitive to the possibilities of endogenous change.

This particular case study leaves many unanswered questions. The details of auxiliary selection may be of limited interest to sociolinguists, but I believe the more general questions are worth asking: how can morphosemantic change be most efficiently investigated in a corpus? Is morphose-

mantic change like other sorts of linguistic change? What can diachronic variation tell us about synchronic variation, and vice-versa?

References

- Anagnostopoulou, Elena, Sabine Iatridou and Roumyana Izvorski. 1997. On the Morpho-Syntax of the Perfect and how it relates to its Meaning. Unpublished MIT/Penn manuscript.
- Dowty, David. 1979. *Word Meaning in Montague Grammar*. Dordrecht: Reidel.
- Freeze, Richard. 1992. Existentials and Other Locatives. *Language*
- Friedman, Victor A. 1988. The Category of Evidentiality in the Balkans and the Caucasus American Contributions to the Tenth International Congress of Slavists. Schenker, Alexander M. (ed.). 121-139.
- Iatridou, Sabine. 1995. To Have and Have not: On the Deconstructionist Approach. WCCFL 14.
- Iatridou, Sabine. 1996. The Grammatical Ingredients of Counterfactuality. ms. University of Pennsylvania.
- Izvorski, Roumyana. 1997. The Present Perfect as an Epistemic Modal. Paper presented at SALT 7.
- Kayne, Richard. 1993. Towards a Modular Theory of Auxiliary Selection. *Studia Linguistica*.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1: 199-244
- Kroch, Anthony. 1994. Morphosyntactic Variation. In K. Beals et al., eds. *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasessions on Variation and Linguistic Theory*.
- Labov, William. 1994. *Principles of Linguistic Change. Vol. 1: Internal factors*. Oxford & Cambridge MA: Blackwell.
- Levin, Beth and Malka Rappaport-Hovav. 1996. *Unaccusativity*. Cambridge, MA: MIT Press
- Lipson, Mimi. 1999. *Co-variation of form and meaning in the loss of auxiliary selection in English*. Paper presented at DGFS, Workshop 1 (Meaning Variation-Meaning Change)
- Platzack, Christer. 1987. Have/Be Variation in the Scandanavian Languages.
- Santorini, Beatrice. 1993. The Rate of Phrase-structure Change in the History of Yiddish. *Language Variation and Change*, 5:257-283.
- Weinreich, Uriel, William Labov and Marvin I. Herzog. 1968. Empirical Foundations for a Theory of Language Change. In W. P. Lehmann and Yakov Malkiel, eds. *Directions for Historical Linguistics*, 95-188. Austin & London: University of Texas Press.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104-6305
lipson3@ling.upenn.edu