1-1-2015

# When GET Got Noticed: The Emerging Salience of GET-Passives

Betsy Sneller
*University of Pennsylvania*, esnell@sas.upenn.edu

Sabriya Fisher
*University of Pennsylvania*, sabriya@ling.upenn.edu

# When GET Got Noticed: The Emerging Salience of GET-Passives

# When GET Got Noticed: The Emerging Salience of GET-passives

Betsy Sneller and Sabriya Fisher*

## 1 Introduction

This paper examines the emergence of GET-passives as a stigmatized variant used in passive constructions in American English. A number of studies (Feagin 1979, Labov 1983) provide evidence that GET is increasing in use at the expense of BE as an auxiliary in passive constructions. These studies and others (Herold 1986, Fisher and Sneller 2013) additionally attribute this rise to social correlates like age, gender, and social class.

In the research presented here, we analyze the frequency of passive auxiliary use in the Philadelphia Neighborhood Corpus using data from 162 speakers. This represents the first time a large corpus including social information has been examined with regard to variation between GET and BE in passive constructions. Based on this data as well as supplemental data from Google Ngrams, we confirm an overall increase over time in the use of GET as a passive auxiliary, with a slight reversal in this increase after 1950. We propose that this reversal is due to the emerging salience and stigmatization of the GET variant. Teasing apart the social factors involved in this change reinforces our proposal. Consequently, this work allows predictions to be made about how social salience affects the S-curve of linguistic change.

## 2 Background

### 2.1 Defining the Variable

The variable under consideration in this paper is the choice between GET or BE as a passive auxiliary in eventive statements, as shown in Table 1.

| GET eventive passive | | BE eventive passive |
|---|---|---|
| He got hit. | vs. | He was hit. |
| He gets hit. | vs. | He's hit. |
| He was getting hit. | vs. | He was being hit. |
| He'll get hit. | vs. | He'll be hit. |
| He didn't get hit. | vs. | He wasn't hit. |

Table 1: Examples of semantically equivalent GET and BE passives.

This research focuses exclusively on eventive passives since only they show semantic equivalence in GET-BE variation (see Labov 1975). The BE variant of stative passives, such as *We were married for forty years*, describes a long lasting state of being, whereas the GET variant is eventive: *\*We got married for forty years*. Likewise, resultative passives like *We got worried when you didn't call* fall outside the envelope of variation since the BE variant, *We were worried when you didn't call* can only take a stative reading. As a result, all non-eventive passive constructions were excluded from the analysis.

### 2.2 Variation in Passive Auxiliary Use

Variation between GET- and BE-passives has been attributed to social factors such as gender, age, and socioeconomic class. Feagin (1979) finds the highest frequency of GET-passive use among working

class teens and working class adult men. The preference for the GET-passive is described as a change from below, spreading from the working to upper classes due to the lack of stigma attached to its use in Anniston. Weiner and Labov (1983) find gender stratification whereby GET is used more by men in a sample of speakers from Philadelphia, PA. They also show that increased use of GET is a change in progress based on apparent time observations: GET is higher among adolescents than among adults. This change is confirmed in the Corpus of Historical American English (Kim 2012).
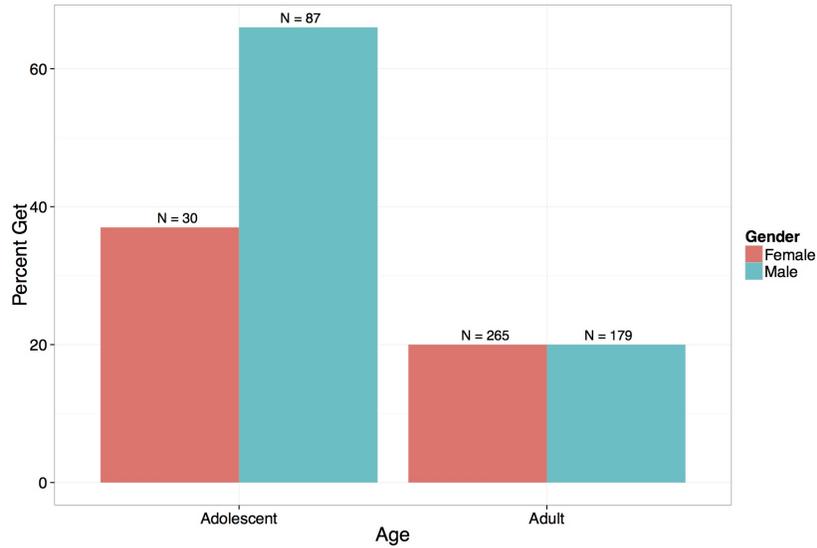


Figure 1: Young males use more GET. Recreated from Weiner and Labov 1983.

In contrast to Feagin (1979), Weiner and Labov (1983) claim that GET is a stigmatized passive variant. They note that, when asked, speakers downplay their use of GET, considering it more colloquial. In addition to this, Herold (1986), as reported in Givon and Yang (1993), finds that GET-passives are more prevalent among the working class. In her sample of speakers, GET-passives make up 49% of total passives in the speech of working class adults and only 17% of total passives in the speech of upper class adults.

In more recent work, Fisher and Sneller (2013) also show social class differences in eventive passive auxiliary use among a sample of nine speakers from three Philadelphia neighborhoods that are stratified for social class: Donegal Street (lower income), Huxley Heights (middle income), Moore Street (upper income). Figure 2 shows higher rates of GET among the lower income neighborhood of Donegal Street. These results are situated within a larger context of overall change toward increased use of GET-passives within the greater Philadelphia speech community.

Figure 2: Social stratification in GET-passive use across three Philadelphia neighborhoods.

## 2.3 Summary

As shown above, previous research suggests that there is a change in progress toward increased use of the GET-passive at the expense of the BE-passive, and that this change is advancing among younger speakers, men, and speakers of lower socioeconomic status. Previous work is inconclusive as to whether GET is a stigmatized passive variant or not, although the social profile for GET-passives fits with the profile for covert prestige (Trudgill 1972, Conn 2005).

In this paper, we find a similar social profile for GET-passives emerging, leading to our hypothesis that the GET-passive is in fact a stigmatized variant with possible covert prestige. This emergent social profile will be further explored in Section 4.

## 3 Methods

For this study, we used both spoken data from the Philadelphia Neighborhood Corpus as well as written data from American English fiction texts. In this section, we discuss these two sources of data as well as our coding methodology.

### 3.1 Spoken Corpora

The primary data for this study come from sociolinguistic interviews from the Philadelphia Neighborhood Corpus (1972–2013). We include only speakers who produced at least five tokens of eventive passives in their interviews, which resulted in 162 speakers ranging in date of birth from 1901–1999. All tokens of all conjugations of GET and BE were extracted from the transcripts of interviews using a shell script, which resulted in a total of 31,513 tokens from the entire corpus. The extraction process included the utterance in which each token occurred, so that context could be used to determine whether the passive was eventive, stative, or resultative.

Tokens were then manually coded for passivity (non-passive and passive), as well as for passive type (stative, resultative, eventive, ambiguous). A 97.9% agreement rate was reached between coders, who were the two authors of this paper. All tokens of stative passives, resultative passives, and ambiguous tokens were excluded from analysis, as were all tokens of "born," which surfaced categorically in BE-passive constructions. This exclusion process left 986 tokens of unambiguous eventive passives to analyze.

Social class was coded using income data from interview reports. These numbers reflected not the actual income of each speaker, but rather the average income from the census tract in which each speaker lived. Income values were adjusted for inflation to reflect 2013 rates. In cases where the income data was not included in the interview reports, we obtained the average income from the census tract in which the interviews were conducted.

### 3.2 Written Corpora

Spoken data from the PNC was supplemented with data from Google Ngrams, compiled from American English fiction texts dating back to 1800. This provided a more complete picture of variation between GET and BE in two ways. First, the written data increases the time scale of the research by a century. This allows age-grading to be ruled out as a possible interpretation of the increase of GET in the PNC. Second, examining written data allows the comparison of spoken and written rates of GET. Because the writings included in Google Ngrams are published texts, a higher degree of formality and commensurate decrease in stigmatized variants is expected. If GET-passives become salient around the 1950s as we suggest in this paper, then we expect to find a reaction to this emergent salience in the rates of GET in both spoken and written data.

To view the trajectory of eventive GET and BE tokens over time in Google Ngrams, ten unambiguously eventive passives were examined (AUX + 'caught', 'fired', 'shot', 'chased', 'killed', 'arrested', 'hit', 'promoted', 'paid', 'sent'). Collocations of these passive constructions were entered as search terms that included a calculation of the percentage of collocations that appeared in the corpus as GET rather than BE (e.g., ((got caught)/(got caught + was caught))).

## 4  Results and Analysis

### 4.1  Results from the PNC

#### 4.1.1  Date of Birth

Previous work (Feagin 1979, Weiner and Labov 1983) has suggested that increased use of GET-passives represents a change in progress. As shown in Figure 3, data from the PNC provides clear evidence of a rise in the use of the GET-passive between 1901 and 1999 such that speakers with earlier dates of birth show a preference for BE, while speakers with later dates of birth prefer GET.
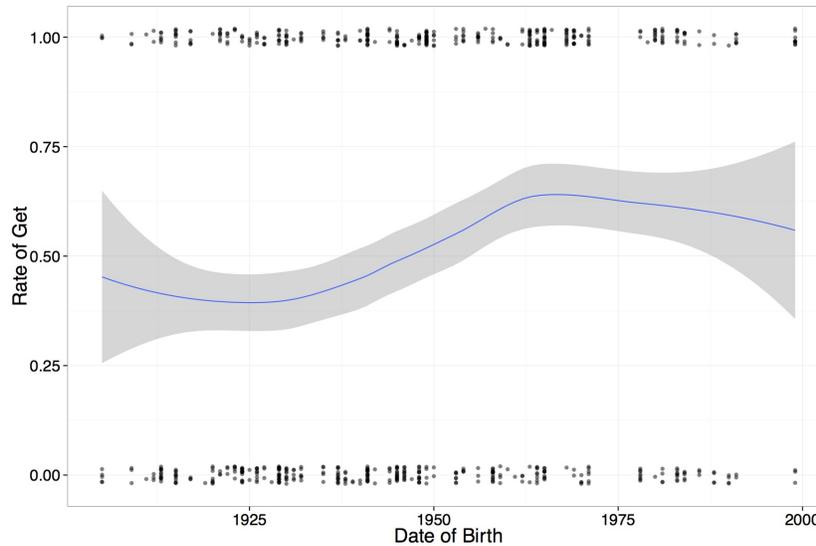


Figure 3: Rate of GET by date of birth.

Importantly, there is a steady increase in the use of GET until after 1950, when we see that rates of GET stop increasing. This time period is consistent with other movements away from stigmatized variants in the Philadelphia dialect (Labov et al., 2013). We propose that this reversal occured when the GET-passive reached salience and stigmatization in Philadelphia. For the time being, we will remain agnostic as to the cause of this emergent salience and stigmatization.

To further understand the reversal in the trajectory of the change toward increased GET, we will turn to a more fine-grained analysis of social factors with regard to date of birth as a diagnostic of the social meaning of GET-passives.
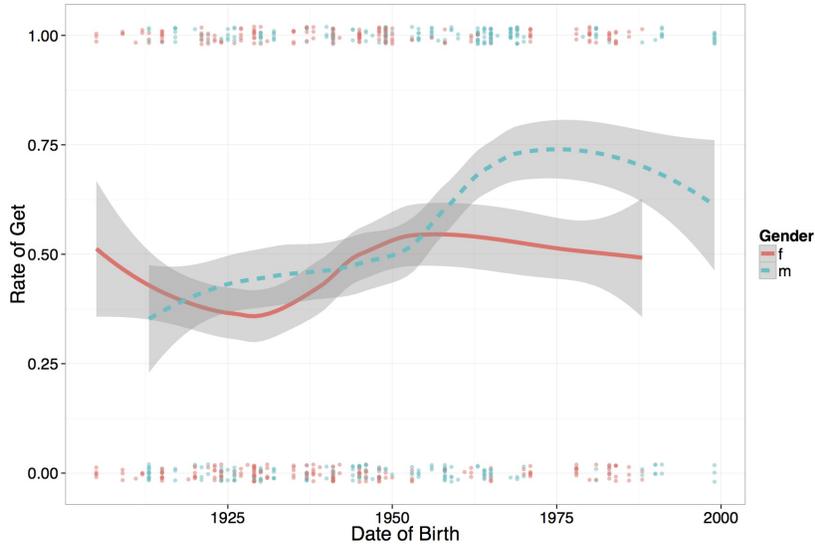
#### 4.1.2  Gender

Figure 4: Rate of GET by gender.

In Figure 4, an interesting pattern emerges. From 1900-1955, men and women pattern together with respect to GET-passives: both genders show an increase in use of GET-passives at similar rates. After 1955, however, the behavior of men and women diverges, with men continuing to advance the change uninterrupted, while women halt their increase in use of the GET-passive.

This pattern is not inconsistent with Labov's Principles of sound change (Labov, 1990):

| | |
|---|---|
| Principle 1: | For stable sociolinguistic variables, men use a higher frequency of nonstandard forms than women. |
| Principle 1a: | In change from above, women favor the incoming prestige form more than men. |
| Principle 2: | In change from below, women are most often the innovators. |

In other words, we can posit an analysis of Figure 4 as depicting a change in progress from below which becomes stigmatized around 1950. Women, in response to this stigmatization, suspend their participation in this change. Men, as a whole, continue uninterrupted by the emergent salience of the GET-passive.

| Predictor | Coefficient | P-value |
|---|---|---|
| Date of Birth | 0.11 | 0.75 |
| Gender (male) | -0.09 | 0.7 |
| Date of Birth(young):Gender(m) | 1.12 | .01* |
| glmer(Code ~Date of Birth ∗ Gender + (1|Speaker)) | | |

Table 2: Results of the best-fit model.

Table 2 depicts the results of the best-fit logistic regression mixed-effects model, with an random intercept for Speaker. In this model, Date of Birth is binned into "older" (speakers born before 1950) and "younger" (speakers born after 1950). All models included a random intercept for speaker. Neither Date of Birth nor Gender are significant in this model, but the interaction of Date of Birth and Gender is.[1] This means that for younger speakers, men are significantly more likely to use GET than

---

[1]It is worth noting that Date of Birth and Gender are not significant predictors, but that this lack of significance is likely due to the small amount of data distributed over a large timespan.

BE. This serves as a confirmation of the trend depicted in Figure 4, which shows no meaningful difference between men and women for the older speakers, but a difference between men and women for the younger speakers where men are more likely to use GET.

### 4.1.3 Social Class

In addition to gender, we expect social class to have an effect on rates of GET after the proposed point of emergent salience.
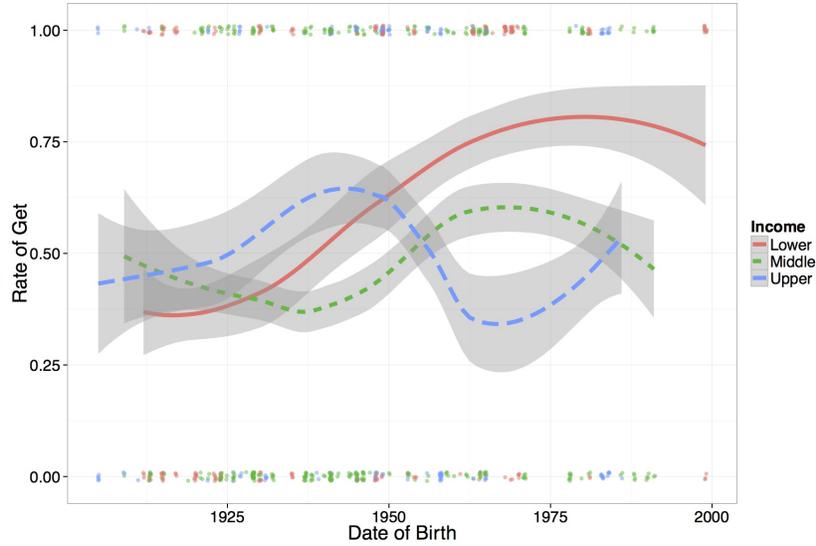


Figure 5: Rate of GET by income.

Figure 5 provides further justification for the claim that the GET-passive became socially salient post-1950. All three social classes pattern together until the 1950s, at which time a pattern of social stratification emerges. The lower class leads the change towards increased use of GET, while the middle class follows behind at lower frequencies, and the upper class appears to reverse the change. This social stratification provides additional evidence that the GET-passive is not socially stigmatized before 1950, but becomes so after this point.

| Predictor | Coefficient | P-value |
|---|---|---|
| Date of Birth | 1.49 | $< .001$*** |
| Income (mid) | -0.08 | .8 |
| Income (upper) | 0.25 | .48 |
| Date of Birth:Income(m) | -0.86 | .1 |
| Date of Birth:Income(up) | -1.55 | .016* |

glmer(Code ~Date of Birth ∗ Income + (1|Speaker))

Table 3: Results of the best-fit model.

Table 3 depicts the results of the best-fit logistic regression mixed effects model, with Speaker as a random intercept. Date of Birth remains a significant predictor of GET usage. However, it should be noted that the data is slightly skewed towards lower-income speakers having later dates of birth and upper-income speakers having earlier dates of birth. While rates of GET increase overall (see Figure 3), upper-income younger speakers are moving away from this trend in comparison to their

lower-income peers. Again, this is consistent with an analysis of the salience of the GET-passive emerging after 1950.

Income groups are not significant predictors on their own, but the interactions of Date of Birth and Income are. In this model, Date of Birth(older) and Income(lower) are taken as the baselines. The interaction of Date of Birth and Income(middle) is not significant. This means that for younger speakers, middle income speakers are not significantly different than lower income speakers. However, it should be noted that middle income speakers do trend toward using more BE than their lower income counterparts. It is likely that the lack of significance is due to the relatively small Ns in the dataset. Importantly, the interaction of Date of Birth and Income(upper) does emerge as significant. This means that for younger speakers only, upper income speakers are significantly more likely to use BE than their lower income counterparts. This social stratification found only in the younger speakers is consistent with our claim that GET-passives are not stigmatized before 1950 but become salient and stigmatized after this point.

To summarize, we have confirmed that increased use of the GET-passive is a change in progress in the PNC. We have additionally found a significant Gender effect for speakers born after the 1950s such that men use more GET than women as well as an Income effect for those same speakers such that those with lower incomes use more GET. These findings are consistent with GET as a stigmatized passive auxiliary variant that rose to social salience post-1950.

## 4.2  Results from Google Ngrams

We now turn to the data from Google Ngrams to examine how our proposal of the emergent salience of GET-passives is reflected in written data. Figure 6 shows a steady increase in the use of GET-passives between the publishing years of 1800 and 2000.
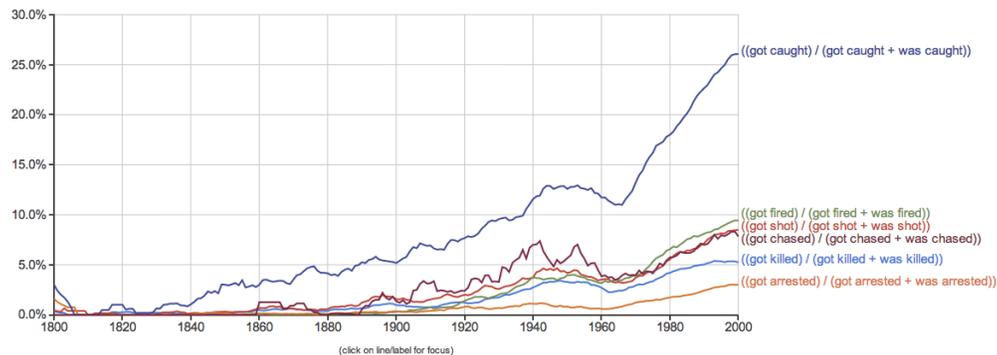


Figure 6: Percentage of GET in written texts of American English.

These results from published texts are clearly similar to the spoken data found in the PNC and allow us to eliminate age grading as a possible explanation of our results. There is an overall increase in the use of GET-passives up until around the 1950s, at which point we see a similar retreat from the variant. Notably, the rates of GET usage are much lower in the published texts than in the spoken data (under 30%), but this is to be expected. Published texts represent a more formal register, which is less likely to include innovative variants like the GET-passive than colloquial spoken data is. After the dip, rates of GET for the lexical items studied increases to nearly 30%. This is still relatively low, but raises the question of why a stigmatized variant would rapidly increase in written formats after 1950. Answering this question would require further research, but one avenue to pursue would be looking at the formality of texts included in Google Ngrams at each time period.

Crucially, we see in the data from Google Ngrams what looks to be the beginning of the S-curve of language change, but with a slight retreat from the incrementation in GET occurring around the 1950s. This result is in line with an interpretation of the GET-passive as reaching stigmatization at that point in time. Importantly, we do not see a complete reversal, but rather a slight hiccup in the

trajectory of the change. This data thus provides an example of what can happen to the S-curve of language change when the change in question becomes socially salient. This finding is discussed further in Section 5 below.

## 5  Discussion: the Effect of Salience on Syntactic Change

We find a change in progress toward increased use of GET-passives in spoken data, which corroborates similar claims made by Feagin (1979) and Weiner and Labov (1983). Additionally, we find that this change is also reflected in published texts between the years 1800 and 2000. The GET-passive is believed to be a stigmatized passive variant and its increase has been driven by lower class men since the period following 1950.

 Our results further speak to the effect that social salience and stigmatization can have on the process of syntactic change, namely that the S-curve becomes momentarily interrupted by this emergent stigmatization. This finding is similar to that of Warner (2005), who examined the rise of DO-support in Middle English.
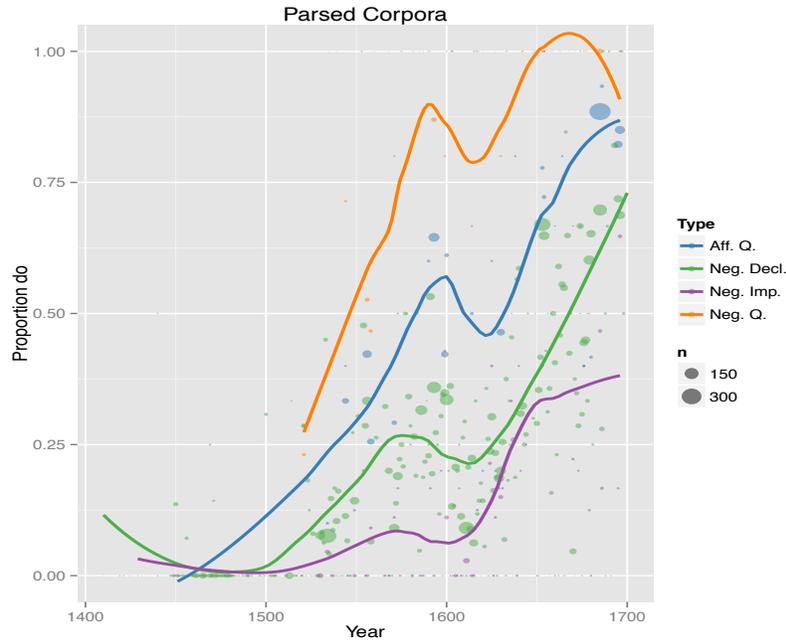


Figure 7: Dip in the rise of DO-support in Middle English. Figure recreated from Warner 2005.

 Warner (2005) examines rising rates of DO-support in four syntactic categories (affirmative questions, negative declaratives, negative imperatives, negative questions), as well as across texts of varying register. He finds that the use of DO decreased in the first half of the 17th century. He also finds that after this period of time, more informal texts contain higher rates of DO than the higher register texts do. He attributes this dip, as well as the differing rates of DO across text type, to an emerging salience of DO-support.

 Warner's data shows that even after this dive in the rise of DO-support, the change continues to completion. If the data on GET-passives from Google Ngrams is indeed showing a similar dip in the rise of the GET-passive, then we expect to see this change also continue toward completion.

 The same exact pattern does not hold in the spoken data from the PNC, but this may be due to the relatively shallow time depth of the PNC and the dearth of data from speakers born after 1990. While we have spoken data ranging over nearly a century, the data from Warner (2005) as well as from Google Ngrams suggests that several centuries of time depth might be necessary before a clear

picture emerges. Our expectation is that, given enough time, both spoken and written data will show the GET-passive behaving very similarly to DO-support: the S-curve will continue but will show a slight hiccup in its incrementation at the point of emergent salience.

## 6 Conclusion

This research provides evidence that the GET-passive has been increasing at the expense of the BE-passive in both spoken and written American English over the past two centuries. In Philadelphia, the GET-passive became salient for speakers born around the 1950s. This salience seems to have resulted in the stigmatization of the GET variant, which is reflected in the emergent social stratification by both income and gender such that women and upper-income speakers retreat from participating in the change. An additional conclusion of this work is that emergent salience in syntactic change may cause a momentary dip in the trajectory of the S-curve, following Warner (2005).

## References

Conn, Jeffery. 2005. Of "Moice" and Men: The Evolution of a Male-Led Sound Change. Doctoral dissertation, University of Pennsylvania.

Feagin, Crawford. 1979. *Variation and Change in Alabama English: A Sociolinguistic Study of the White Community*. Georgetown University Press.

Fisher, Sabriya, and Betsy Sneller. 2013. Get vs. Be: social class differences in variable passive auxiliary use. Paper presented at NWAV 42, Carnegie Mellon University.

Givon, Talmy, and Lynne Yang. 1993. The Rise of the English GET-Passive. In *Voice: Form and Function*, ed. Barbara Fox and Paul Hopper, 119–150. John Benjamins.

Herold, Ruth. 1986. A quantitative study of the alternation between the BE- and GET- Passives. Paper presented at NWAV 15, Stanford University.

Kim, Jong Bok. 2012. English Get-Passive Constructions: A Corpus-Based Approach. *Studies in Generative Grammar* 22.2:437–457.

Labov, William. 1975. Empirical foundations of linguistic theory. In *The Scope of American Linguistics*, ed. Robert Austerlitz, 77–133. de Ridder.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2:205–254.

Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89:30–65.

Philadelphia Neighborhood Corpus of LING560 Studies. 1972–2013. With support of NSF contract 921643 to W. Labov.

Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1:179–195.

Warner, Anthony. 2005. Why DO dove: Evidence for register variation in Early Modern English negatives. *Language Variation and Change* 17:257–280.

Weiner, Judith, and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19:29–58.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104
*esnell@ling.upenn.edu*
*sabriya@ling.upenn.edu*